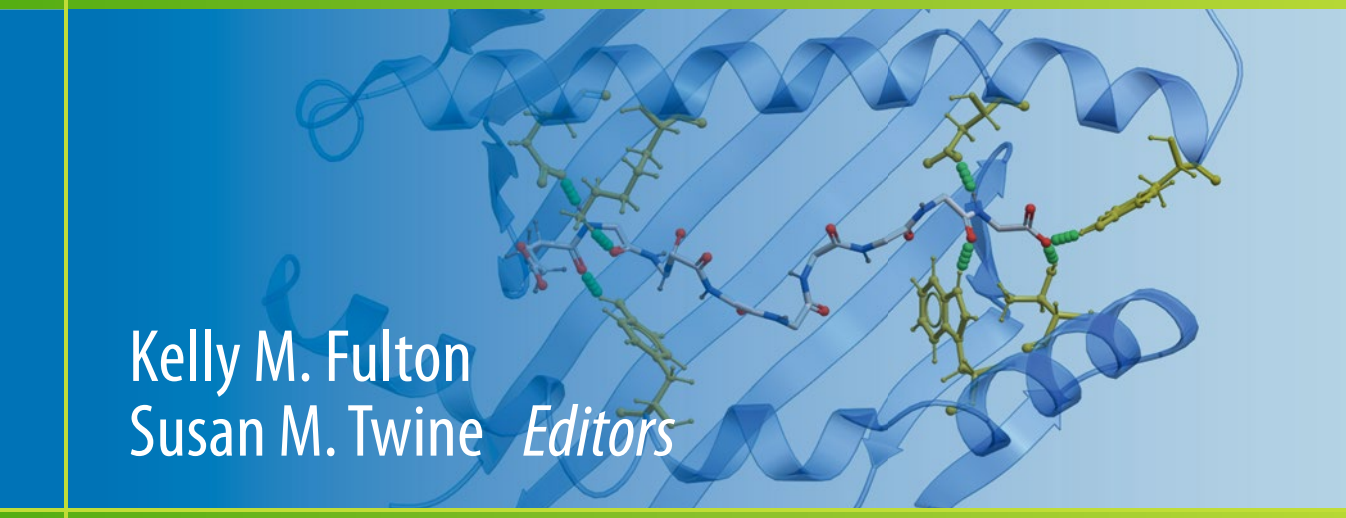
Methods in Molecular Biology 1061

Springer Protocols



Kelly M. Fulton
Susan M. Twine *Editors*

Immuno-proteomics

Methods and Protocols

Humana Press

METHODS IN MOLECULAR BIOLOGY™

Series Editor
John M. Walker
School of Life Sciences
University of Hertfordshire
Hatfield, Hertfordshire, AL10 9AB, UK

For further volumes:
http://www.springer.com/series/7651

Immunoproteomics

Methods and Protocols

Edited by

Kelly M. Fulton

Human Health Therapeutics, National Research Council Canada, Ottawa, ON, Canada

Susan M. Twine

Human Health Therapeutics, National Research Council Canada, Ottawa, ON, Canada

Humana Press

Editors
Kelly M. Fulton
Human Health Therapeutics
National Research Council Canada
Ottawa, ON, Canada

Susan M. Twine
Human Health Therapeutics
National Research Council Canada
Ottawa, ON, Canada

ISSN 1064-3745 ISSN 1940-6029 (electronic)
ISBN 978-1-62703-588-0 ISBN 978-1-62703-589-7 (eBook)
DOI 10.1007/978-1-62703-589-7
Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2013943915

Printed on acid-free paper

Humana Press is a brand of Springer
Springer is part of Springer Science+Business Media (www.springer.com)

Preface

The adaptive immune system is a complex system of cells, tissues, and organs that constantly samples proteins and peptides of the immune landscape as part of defense against disease. Antibodies directed against foreign proteins, or peptides derived from invading pathogens that activate B- and T-cells, are highly amenable to study by protein-based methods. Classical methods to study the immune response to disease have been used for many years, including agglutination, enzyme-linked immunoabsorbent assay, or Western blotting. Rapid advances in genomics and proteomics in the past two decades now allow characterization and quantification of protein and peptide targets of the immune response to disease. The collective study of the complement of proteins and peptides that stimulate an immune response has been termed "immunoproteomics." Describing a broad and rapidly growing field, this umbrella term includes gel-based, array-based, mass spectrometry-based, DNA-based, and *in silico* approaches. Immunoproteomics is yielding an understanding of disease and disease progression, vaccine candidates, and biomarkers. The resulting information has potential to be used in diagnostics, disease progression, and vaccine correlates of protection analysis, to name but a few. This book provides a broad overview of the current diverse approaches and techniques that are being exploited to study the immunoproteome.

The book opens with an introduction to the immune system and a broad overview of some of the major techniques used in immunoproteomics. Following this, authors present techniques used for the study of the antibody targets of bacterial pathogens, viruses, and cancer. These include classical techniques such as serological proteome analysis and expression arrays, in addition to emerging approaches such as *in vivo* microbial antigen discovery, detection of glycoprotein antigens, and mass spectrometry differentiation of immune cells. Later chapters describe mass spectrometry-based approaches to characterize T-cell epitopes followed by the detection and relative quantification of cytokines in serum. *In silico* prediction of epitopes using sequence-based or modeling approaches has been important in immunological research for many years, and protocols are presented here to aid the experimental researcher in utilizing these approaches. Lastly, perspective upon translation and commercialization of immunoproteomic biomarkers, an increasingly important consideration for many researchers, is discussed in the closing chapter.

This book is aimed at scientists new to the field and those with years of experience in immunoproteomics. The variety of techniques presented provides not only an overview of the breadth of the field but valuable hands-on insights from specialists. Success stories will allow readers to transfer these techniques to their laboratories, in addition to providing a reference to guide researchers towards appropriate techniques.

Ottawa, ON, Canada

Kelly M. Fulton
Susan M. Twine

Contents

Contributors

Bob Amess • *Department of Chemical Engineering and Biotechnology, Cambridge Centre for Neuropsychiatric Research, University of Cambridge, Cambridge, UK*

David P. AuCoin • *Department of Microbiology and Immunology, University of Nevada, Reno, NV, USA*

Sabine Bahn • *Department of Chemical Engineering and Biotechnology, Cambridge Centre for Neuropsychiatric Research, University of Cambridge, Cambridge, UK*

Elisa Beghetto • *Sigma-Tau Industrie Farmaceutiche Riunite S.p.A., Rome, Italy*

N. Helena Beyer • *Statens Serum Institut, Copenhagen, Denmark*

Nick Bond • *Department of Biochemistry, Cambridge Centre for Proteomics, University of Cambridge, Cambridge, UK*

Andrew J. Bordner • *Mayo Clinic, Scottsdale, AZ, USA*

Leonia Bozzacco • *Laboratory of Cellular Physiology and Immunology and Chris Browne Center, The Rockefeller University, New York, NY, USA*

Christian Capo • *Unité de Recherche sur les Maladies Infectieuses Tropicales et Emergentes, Aix Marseille Université, Marseille, France*

King C. Chan • *Laboratory of Proteomics and Analytical Technologies, Advanced Technology Program, Frederick National Laboratory for Cancer Research, SAIC-Frederick, Frederick, MD, USA*

Philip D. Charles • *Department of Biochemistry, Cambridge Centre for Proteomics, University of Cambridge, Cambridge, UK*

Sindy J. Chaves • *Department of Microbiology and Immunology, University of Nevada, Reno, NV, USA*

Wangxue Chen • *Human Health Therapeutics Portfolio, National Research Council Canada, Ottawa, ON, Canada*

Andy Christoforou • *Department of Biochemistry, Cambridge Centre for Proteomics, University of Cambridge, Cambridge, UK*

Aurélie Daumas • *Unité de Recherche sur les Maladies Infectieuses Tropicales et Emergentes, Aix Marseille Université, Marseille, France*

Mike Deery • *Department of Biochemistry, Cambridge Centre for Proteomics, University of Cambridge, Cambridge, UK*

Kevin M. Downard • *School of Molecular Bioscience, University of Sydney, Sydney, NSW, Australia*

Scott M. Ferguson • *Human Health Therapeutics Portfolio, National Research Council Canada, Ottawa, ON, Canada*

Simon J. Foote • *Human Health Therapeutics Portfolio, National Research Council Canada, Ottawa, ON, Canada*

Kelly M. Fulton • *Human Health Therapeutics, National Research Council Canada, Ottawa, ON, Canada*

Nicola Gargano • *Sigma-Tau Industrie Farmaceutiche Riunite S.p.A., Rome, Italy*

Arnoud Groen • *Department of Biochemistry, Cambridge Centre for Proteomics, University of Cambridge, Cambridge, UK*
Paul C. Guest • *Department of Chemical Engineering and Biotechnology, Cambridge Centre for Neuropsychiatric Research, University of Cambridge, Cambridge, UK*
Greg Harris • *Human Health Therapeutics Portfolio, National Research Council Canada, Ottawa, ON, Canada*
Arsalan S. Haqqani • *Human Health Therapeutics Portfolio, National Research Council Canada, Ottawa, ON, Canada*
Niels H.H. Heegaard • *Statens Serum Institut, Copenhagen, Denmark*
Svenja Hester • *Department of Biochemistry, Cambridge Centre for Proteomics, University of Cambridge, Cambridge, UK*
Julie Howard • *Department of Biochemistry, Cambridge Centre for Proteomics, University of Cambridge, Cambridge, UK*
Julian A.J. Jaros • *Department of Chemical Engineering and Biotechnology, Cambridge Centre for Neuropsychiatric Research, University of Cambridge, Cambridge, UK*
Theodoros A. Koutroukides • *Department of Chemical Engineering and Biotechnology, Cambridge Centre for Neuropsychiatric Research, University of Cambridge, Cambridge, UK*
Thomas R. Kozel • *Department of Microbiology and Immunology, University of Nevada, Reno, NV, USA*
Lakshmi Krishnan • *Human Health Therapeutics Portfolio, National Research Council Canada, Ottawa, ON, Canada*
Yishai Levin • *Proteomics and Mass Spectrometry Unit, Weizmann Institute of Science, Rehovot, Israel*
Kathryn S. Lilley • *Department of Biochemistry, Cambridge Centre for Proteomics, University of Cambridge, Cambridge, UK*
Guangjin Liu • *Key Laboratory of Animal Bacteriology, Ministry of Agriculture, Nanjing Agricultural University, Nanjing, China*
Chengping Lu • *Key Laboratory of Animal Bacteriology, Ministry of Agriculture, Nanjing Agricultural University, Nanjing, China*
Christina J. Maier • *Division of Molecular Dermatology, Department of Dermatology, Paracelsus Private Medical University, Salzburg, Austria*
Richard H. Maier • *Division of Molecular Dermatology, Department of Dermatology, Paracelsus Private Medical University, Salzburg, Austria*
Shannon S. Martin • *DynPort Vaccine Company LLC CSC, Frederick, MD, USA*
Daniel Martins-de-Souza • *Proteomics and Biomarkers, Max Planck Institute of Psychiatry, Munich, Germany*
Scott McComb • *Department of Biochemistry, Microbiology, and Immunology, University of Ottawa, Ottawa, ON, Canada*
Jean-Louis Mege • *Unité de Recherche sur les Maladies Infectieuses Tropicales et Emergentes, Aix Marseille Université, Marseille, France*
Alexander Nøstdal • *Department of Cellular and Molecular Medicine, University of Copenhagen, København, Denmark*
Kamil Önder • *Division of Molecular Dermatology, Department of Dermatology, Paracelsus Private Medical University, Salzburg, Austria*
Richard Ouedraogo • *Unité de Recherche sur les Maladies Infectieuses Tropicales et Emergentes, Aix Marseille Université, Marseille, France*

Hassan Rahmoune • *Department of Chemical Engineering and Biotechnology, Cambridge Centre for Neuropsychiatric Research, University of Cambridge, Cambridge, UK*
Ugur Sahin • *Translational Oncology (TRON), Mainz, Germany*
Kathleen Schegg • *Department of Microbiology and Immunology, University of Nevada, Reno, NV, USA*
Danica B. Stanimirovic • *Human Health Therapeutics Portfolio, National Research Council Canada, Ottawa, ON, Canada*
Felicity Stark • *Human Health Therapeutics Portfolio, National Research Council Canada, Ottawa, ON, Canada*
Luc Tessier • *Human Health Therapeutics Portfolio, National Research Council Canada, Ottawa, ON, Canada*
Julien Textoris • *Unité de Recherche sur les Maladies Infectieuses Tropicales et Emergentes, Aix Marseille Université, Marseille, France*
Aude Thiriot • *Microbiology and Immunobiology, Harvard Medical School, Boston, MA, USA*
Özlem Türeci • *Translational Oncology at the University Hospital, Mainz, Germany*
Susan M. Twine • *Human Health Therapeutics, National Research Council Canada, Ottawa, ON, Canada*
Timothy D. Veenstra • *Laboratory of Proteomics and Analytical Technologies, Advanced Technology Program, Frederick National Laboratory for Cancer Research, SAIC-Frederick, Frederick, MD, USA*
Hans H. Wandall • *Department of Cellular and Molecular Medicine, University of Copenhagen, København, Denmark*
Timothy J. Waybright • *Laboratory of Proteomics and Analytical Technologies, Advanced Technology Program, Frederick National Laboratory for Cancer Research, SAIC-Frederick, Frederick, MD, USA*
Fei Wen • *Department of Chemical Engineering, University of Michigan, Ann Arbor, MI, USA*
Lawrence Wolfraim • *DynPort Vaccine Company LLC CSC, Frederick, MD, USA*
Zhen Xiao • *Laboratory of Proteomics and Analytical Technologies, Advanced Technology Program, Frederick National Laboratory for Cancer Research, SAIC-Frederick, Frederick, MD, USA*
Haiqiang Yu • *ThemoFisher Scientifc, San Francisco, CA, USA*
Wei Zhang • *Key Laboratory of Animal Bacteriology, Ministry of Agriculture, Nanjing Agricultural University, Nanjing, China*
Huimin Zhao • *Department of Chemical and Biomolecular Engineering, University of Illinois at Urbana-Champaign, Urbana, IL, USA; Departments of Chemistry, Biochemistry, and Bioengineering, Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, IL, USA*

Chapter 1

Introduction to the Immune System

Scott McComb, Aude Thiriot, Lakshmi Krishnan, and Felicity Stark

Abstract

The immune system in a broad sense is a mechanism that allows a living organism to discriminate between "self" and "non-self." Examples of immune systems occur in multicellular organisms as simple and ancient as sea sponges. In fact, complex multicellular life would be impossible without the ability to exclude external life from the internal environment. This introduction to the immune system explores the cell types and soluble factors involved in immune reactions, as well as their location in the body during development and maintenance. Additionally, a description of the immunological events during an innate and adaptive immune reaction to an infection is discussed, as well as a brief introduction to autoimmunity and cancer immunity.

Key words Immune system, Immunity, Vaccines, Adaptive immunity, Innate immunity, Inflammation

1 Introduction

Early immune system discoveries were largely fuelled by a desire to prevent the spread of disease and develop better treatments for the sick (Fig. 1). As far back as the eighteenth century, microbiologists sought to inoculate healthy people against diseases. In fact, vaccines were created to combat illness before anyone could prove that microbes caused illness, or that immune cells could kill microbes. Over 100 years prior to Koch's postulates in 1890, which definitively identified microbes as the causative agent of disease, Edward Jenner had made a crude vaccine from the pus of cow pox lesions to successfully immunize people against small pox [1].

As microbiologists were uncovering the properties of bacteria, it was commonly believed that white blood cells aided the spread of pathogens by transporting them throughout the host. In 1882, Eli Metchnikoff recognized that white blood cells were instead destroying the engulfed pathogen [2]. The word phagocyte, from the Greek words "phagein," to eat and "cyte" cell, was used to describe this cellular action. Metchnikoff had identified what would come to be known as innate immunity. Innate immunity is the evolutionarily older arm of the immune system, composed of

Kelly M. Fulton and Susan M. Twine (eds.), *Immunoproteomics: Methods and Protocols*, Methods in Molecular Biology, vol. 1061,
DOI 10.1007/978-1-62703-589-7_1,

Fig. 1 The term "immune" is derived from the Latin "immunis," literally meaning to be exempt from the public service required of lower classes. While the Roman empire may have bestowed immunity from taxation upon the chosen few, there was no special exemption from the scourge of virulent disease. From (http://en.wikipedia.org/wiki/File:Plague_in_Ashod.jpg)

barriers (skin), small molecules (complement), and cells such as macrophages and dendritic cells. The innate immune system is so called because it provides protection from pathogens without the need for preconditioning from the environment. In other words, when the innate immune system encounters a pathogen, it will react immediately to kill or to remove it from the host.

In addition to Metchnikoff's discovery of cellular immunity, other researchers were examining the ability of bodily fluids (humors) to provide protection against disease. In 1890, Emil von Behring and Shibasaburō Kitasato discovered antibodies when they identified acellular components of the blood that conferred immunity when transferred from one animal to another [3]. Antibodies, along with cytokines and complement are the components of humoral immunity. Interestingly, the discovery of antibodies sparked heated debates that divided scientists about the importance each type of immunity played in overall host immunity. This divide was bridged in 1903 when scientists Almroth Wright and Steward Douglas proved that humoral responses aided the cellular immune response suggesting that both cellular and humoral immune responses played important roles. They observed that antibodies and complement enhanced the phagocytosis of bacteria by binding to the bacteria, an event termed opsonization [4].

In contrast to humoral immunity, the cellular fraction of the immune system is more commonly known as cell mediated immunity. The ability of the immune system to generate a specific cellular response to a pathogen became known as adaptive immunity. How the immune system was able to produce such highly specific immune receptors remained a long standing mystery in immunology until the mid twentieth century. In 1965, Dreyer and Bennet published a speculative paper suggesting that DNA recombination of immune genes could generate immune diversity [5]. Throughout the next decade, the work of many researchers identified that the V(D)J genetic elements of lymphocyte receptors could be shuffled randomly. Through this rearrangement, a library of lymphocytes is created each with a unique gene for its immune receptor. The result is a large pool of highly diverse lymphocytes each capable of binding a unique pathogen associated molecule. The discovery of lymphocyte receptor specificity could explain Emil von Behring and Shibasaburō Kitasato's earlier findings that sera from vaccinated animals could provide protection to other animals challenged with the same pathogen. The antibodies present in the sera were actually soluble immune receptor which is a product of a clone of B lymphocytes.

While the practical application of immunizations may have outpaced our understanding of the immune system in the early years of modern medicine, we now have insight into the machinery behind immunity. This knowledge has translated into better tools to both monitor and manage immune responses, and ultimately improve patient outcomes. In this chapter we will further delve into the mechanisms that allow our bodies to recognize, respond to, and remember pathogens which challenge the body.

2 Organs of the Immune System

While immune cells can be found throughout the body, the immune system also has a specialized network of immune organs (Fig. 2). The organization of the immune system within immune organs allows for a regulated immune response capable of rapidly producing a large number of cells that can halt a spreading infection. From these immune reservoirs, immune cells and molecules can be released to penetrate almost any tissue throughout the body.

While all cells of the blood originate from the bone marrow, their sites of maturation and residence differ. Many different types of innate immune cells typically arise from the bone marrow and take residence in the blood and tissues (discussed further below). For acquired immune cells, T and B cells will recombine their immune receptors in the thymus and bone marrow, respectively; these are known as primary immune organs. After maturation in these primary sites, T and B cells will reside in the lymphatic tissue,

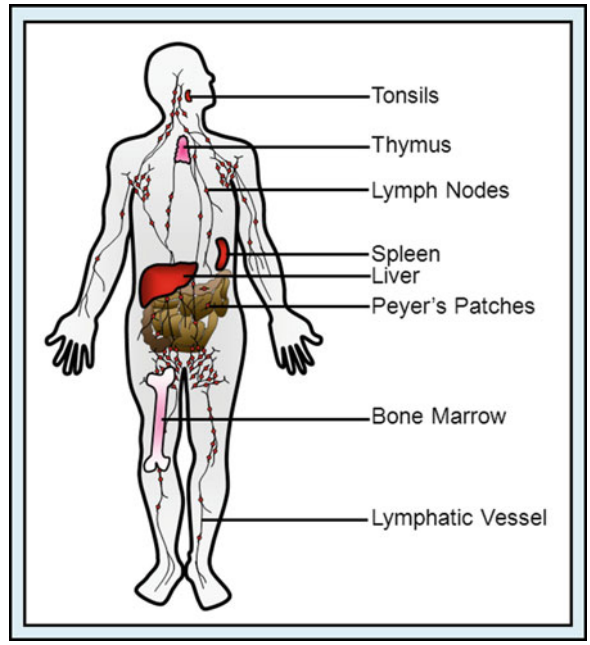


Fig. 2 While immune cells are located throughout the body, the immune system relies on specialized organs to generate immune cells, detect pathogens, and initiate immune responses. Many immune cells arise from precursor cells located in the bone marrow. In the case of T cells, they undergo genetic recombination to form a TCR in the thymus. T and B cells with mature receptors then migrate through lymphatic vessels to lymph nodes where they await activation signals. Large organ systems also have specialized immune sites that harbor immune cells such as the spleen for the circulatory system or Peyer's patches for the gut

and as such they are often referred to as lymphocytes. These sites of lymphocyte residence are known as the secondary immune organs, and they include the lymph nodes, spleen, Peyer's patches, the appendix, tonsils, adenoids, and other mucosal associated lymphoid tissue.

The transportation of immune cells throughout the body occurs both in blood and lymph. Lymph is a clear fluid made up of proteins of the interstitial fluid of bodily tissues; it accumulates in lymph capillaries which are thin walled vessels dispersed throughout all tissues (except the central nervous system (CNS)). Lymph capillaries differ from blood capillaries in that they allow only a one-way passage of lymph into and not out of them. Capillaries lead to the afferent lymph vessels that lead into the lymph node. Unlike the circulatory system of blood, the lymphatic system is not a closed circulatory system and all lymph is moved from lymph nodes to efferent lymphatic vessels towards the heart, unidirectionally. Lymphoid fluid moves through lymph vessels by the passive actions of skeletal muscle contractions and eventually is returned to the blood at the junction of the internal jugular and subclavian veins

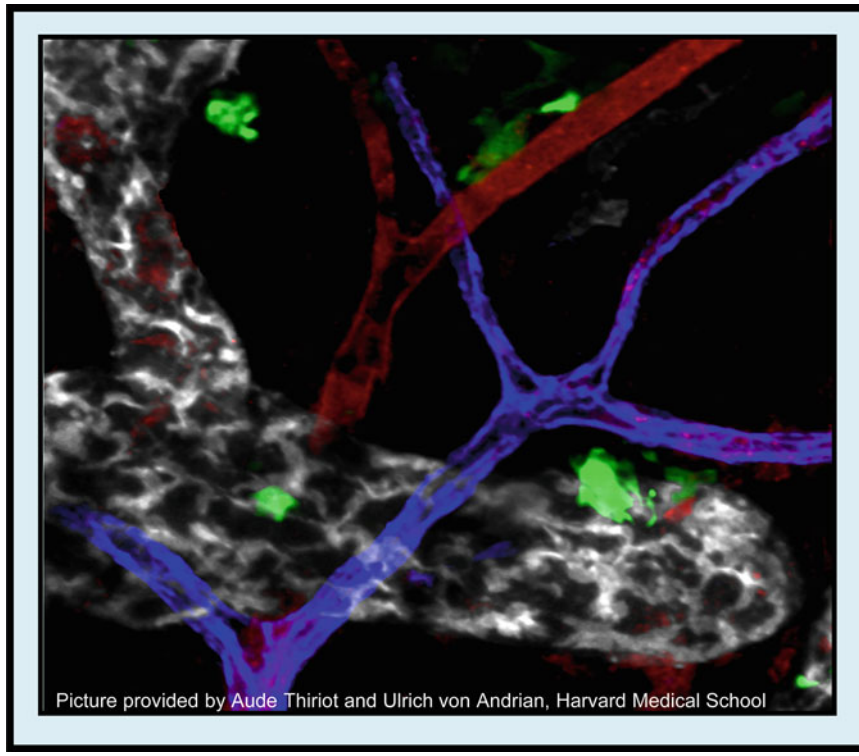


Fig. 3 Immune cells are waiting and watching for invading pathogens throughout the body. Shown here, dendritic cells (CD11c-*green*) within the skin, sit near blood vessels (CD31-*red*) and venules (DARC-*blue*). If these sentinel dendritic cells detect a pathogen, they will trap the invader and use a network of lymphatic vessels (LYVE-1 *white*) to move to a lymph node where an immune response can be activated

at the base of the right side of the neck [6]. Using the bloodstream and lymph, immune cells and pathogen debris can enter lymphoid organs to generate immune responses.

While some innate immune cells are tasked with locating pathogens throughout the body and killing them by various means, other innate immune cells such as dendritic cells have a more complex task to fulfil. Once a dendritic cell has engulfed pathogen, it will migrate in the blood or lymph towards a lymphoid site. The dendritic cell will migrate through the lymphoid tissue with a goal of encountering a lymphocyte and initiating an adaptive immune response which typically occurs within 7 days (Fig. 3). Thus it can be envisioned that the lymphoid compartment is the meeting ground between innate and adaptive immune cells that initiates the activation of the adaptive immune response.

2.1 Cells of the Immune System

The cells of the immune system have been classified into two general groups; these are innate or adaptive immunity (Fig. 4). Innate immune cells react quickly, whereas adaptive immune cells have a delayed response that can take days to fully develop but go on to form immunological memory.

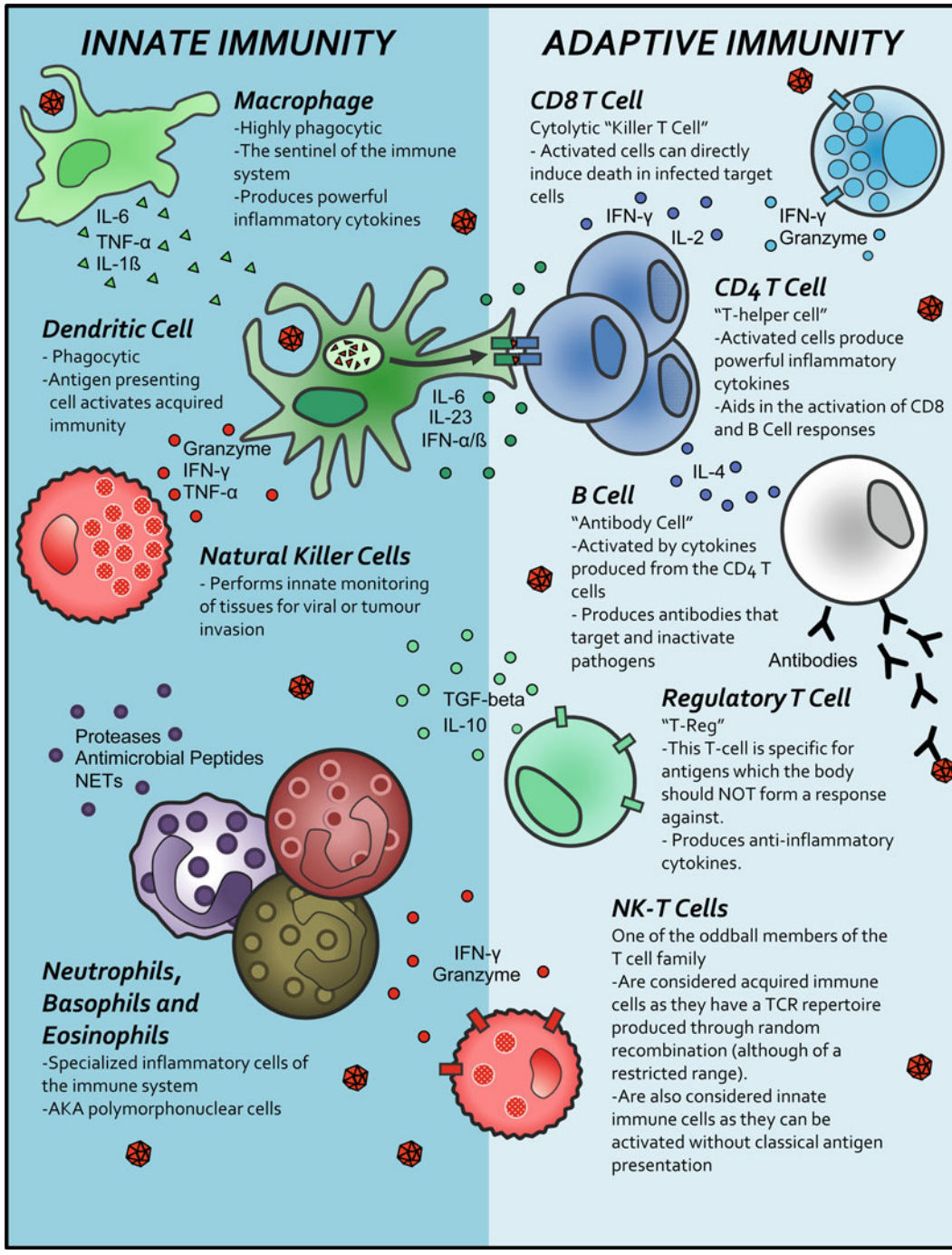


Fig. 4 Common cells of the immune system

The quick responding innate immune cell types include granulocytes (polymorphonuclear cells), mast cells, macrophages and dendritic cells. Mast cells are best known for their ability to rapidly release granules of histamine and heparin in response to an infection.

This rapid response can be important in initiating inflammation and wound healing, but is also involved in allergic responses. Granulocytes encompass a group of three cell types differentiated by the contents of their granules: neutrophils, basophils, and eosinophils. All three are relatively short lived (~5 days) but are important early responders to parasites, extracellular bacteria, and tumors. The early arrival of granulocytes during infection induces acute wound inflammation and dilation of the surrounding blood vessels that allow for the rapid influx of other immune cells. *Neutrophils* are of particular importance as they make up about half of the circulating white blood cells in humans and have a keen ability to phagocytose and destroy invading microbes. In addition to secretion of inflammatory mediators such as cytokines and soluble antimicrobial peptides, neutrophils have recently been found to also undergo a suicidal extrusion of neutrophil extracellular traps (NETs). NETs are primarily composed of DNA and provide a physical barrier by which a pathogen is trapped and prevented from spreading [7].

Similar to neutrophils, macrophages are adept phagocytic cells, capable of ingesting and destroying invading microbes. Unlike neutrophils, which are blood-resident and short-lived, macrophages take up residence in all tissues of the body and are relatively long-lived. Macrophages are also well known to induce inflammation by producing cytokines and chemokines that both draw in and activate other immune cells to the site of infection.

Dendritic cells (DCs) are also capable of internalizing and destroying invading microbes through phagocytosis, however they are better known for their important role of activating the cells of the adaptive immune system. Both DCs and macrophages are often called antigen presenting cells (APCs) due to their efficient ability to internalize pathogens and present pathogen peptides on their cell surface. The term "antigen" refers to a molecule that can evoke an antibody response, and in this case the antigen is the pathogen peptide fragments.

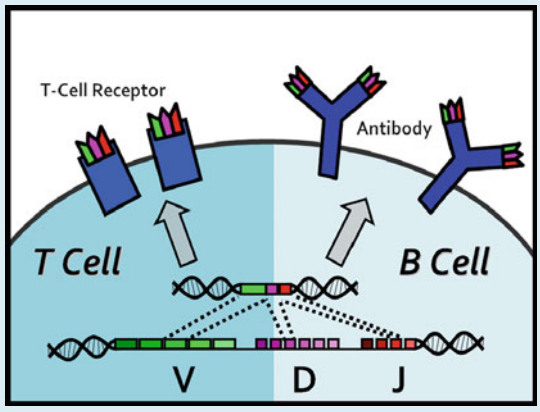
Cells, such as DCs and macrophages, which are capable of antigen presentation can be sub-classified as either professional or non-professional APCs. Both classes of cells process and present pathogen peptides on a receptor called the major histocompatibility complex class II (MHC-II). More specifically, upon APC internalization of the microbe or debris, peptide fragments are generated by proteasomal degradation. The peptide fragments then become bound to the MHC-II receptor and are shuttled to the surface of the cell. Within a lymphoid site, APCs will come into contact with and activate lymphocytes. This occurs via the engagement of the MHC-II with the T cell receptor (TCR); this interaction is often referred to as Signal 1, but it alone is not enough to activate naive T cells. Professional APCs such as Macrophages, DC's and B cells also express cell surface co-stimulatory molecules such as CD80, CD86, and CD40, which can provide "Signal 2" to a naive

lymphocyte. Importantly, a lymphocyte will become activated only if it receives both signal 1 and 2, otherwise it can go into a stasis known as anergy. A non-professional APC such as a fibroblast, thymic epithelial cell, or vascular endothelial cell, only expresses MHC-II in the presence of certain cytokines and does not express co-stimulatory molecules. Thus they are unable to activate naive T cells but can play a role in reactivating memory lymphocytes.

T cells are generally classified into two groups expressing either cell surface CD4 or CD8 receptors. CD4 and CD8 play an important role in the formation of the immune synapse between the TCR and MHC of lymphocytes and target cells respectively. CD8 T cells are most commonly known as cytotoxic T lymphocytes (CTL) because once they strongly engage a target cell they secrete cytotoxic granules and perforin into the immune synapse that penetrates the target cell and induces apoptosis. CD4 T cells are commonly referred to as helper T cells because they play an important role in contributing to the cytokine response that drives either cell mediated immunity by macrophages and CD8 T cells or humoral immunity mediated by B cells. When CD4 T cells are activated in the presence of IL-12 and IFNγ they become of the Th1 phenotype and secrete IFNγ and TNFβ into the environment which induces inflammation and supports the function of macrophages and CD8 T cells to kill pathogens. However, when CD4 T cells are activated in the presence IL-4, they become of the Th2 phenotype which supports the action of B cells to produce antibodies.

B cells provide surveillance to the body for signs of infection by circulating in the blood and the lymph. Their antigen recognition receptor known as the B cell receptor (BCR) is actually a cell bound antibody. Once the B cell becomes activated in the presence of their target antigen, it becomes a plasma cell and begins to produce and secrete large amounts of antibody that can bind to the target protein and neutralize it (Fig. 5). B cells are also classified into two major populations named B1 and B2 lymphocytes, according to their cell surface markers, anatomical location and immunological function. B1 cells are enriched in pleural and peritoneal cavities and are the main producers of natural antibodies, in particular IgM. B2 cells consist of two subsets, the marginal zone B cells (MZ B) and the conventional follicular B cells (FO B). MZ B cells reside in the marginal zone of the spleen and are involved in innate immune response. FO B resides in the follicular zone in the spleen and are present in other lymphoid organs such as lymph node. FO B cells are involved in the adaptive immune response.

While the lymphocytes mentioned thus far predominantly contribute to the adaptive immune response, there exists another lymphocyte termed natural killer cells (NK) that contributes to innate immune responses. The NK lymphocyte performs directed cytolytic killing similar to CD8 T cells by secreting cytotoxic granules into the immune synapse. Their specialized receptors



GENERATING THE ADAPTIVE IMMUNE RECEPTOR LIBRARY

Adaptive immunity is one of the most powerful elements of the immune system because it selects for the most appropriate immune receptor to target the infecting pathogen. Through this selection process, the adaptive immune system can also 'remember' pathogens by maintaining pathogen specific memory cells. Unlike other cells within the body, T and B cells undergo changes at the DNA level during development. Through a mechanism of somatic recombination, gene elements are cut and pasted resulting in random recombination. This creates a staggering diversity of T and B cell receptors.

B cells for example can potentially produce antibodies specific for up to ~10^{18} unique targets, whereas T cells can produce ~10^{13} different receptors [11]. While new T and B cells are produced throughout an organism's lifetime, once a T cell undergoes somatic DNA recombination it maintains this receptor for the remainder of its life. In contrast, activated B cells will undergo further genetic alterations through somatic hypermutation, gene conversion, and class switching. This genetic change further diversifies the specific B cell receptor which is already capable of binding a target. Through successive rounds of mutation and selection, the antibody affinity can be improved several fold.

Fig. 5 Somatic DNA recombination generates lymphocyte receptor diversity

(NKG2D, KIR etc.) scan neighboring cells for signs of infection by recognizing the absence of cell surface MHC. A mechanism by which pathogens evade the immune system, is to prevent the translocation of MHC-peptide molecules to the cell surface. NK cells detect this concealment and kill any cells that do not express a certain amount of MHC molecule. Since NK cells do not require activation to kill, target cells such as tumor or virally infected cells can be killed within 3 days compared to CD8 T cells that require 5–7 days to start killing. NK cells are not thought to express genetically rearranged antigen receptors. However, recent findings have revealed that a subset of liver-resident NK cells can mediate adaptive immune responses. Indeed, this hepatic NK cell population can acquire long-lived and highly specific memory of a variety of viral and hapten-based antigens. The molecular mechanisms by which these NK cells recognize specific antigens remains be discovered [8].

A number of other immune cell subsets perform important roles in an immune response; these include regulatory T cells and suppressor macrophages that play a role in suppressing allergy and autoimmune diseases. There also exists a unique cell subset called the *NK-T cell*, which has both cell surface receptors belonging to NK cells and T cells and are thought to be involved in the

suppression of cell mediated autoimmune responses [9]. The Th-17 subset of CD4 T cells has also been of interest to researchers as its production of IL-17 has been implicated as a causative agent of autoimmunity, however this same cell has been shown to play a protective role in pulmonary infections [10]. While immune cells have been generally classified into groups according to their most prominent roles in an immune response, the constant discovery of novel secreted small molecules and receptors has complicated the categorization process but has also opened new avenues of research to identify new cellular and molecular targets to harness in the prevention of disease.

2.2 Cytokines

Cytokines encompass a wide variety of signaling peptides, proteins and glycoproteins that are used in intercellular communication (Table 2). The categorization of cytokines into tidy groups has been challenging since new properties of existing cytokines are constantly being uncovered. Historically, cytokines were named either by their cellular origin or their cellular target; that is, interleukins are produced by leukocytes and tumor necrosis factor (TNF) inhibits tumorigenesis. Since the discovery that some previously identified interleukins were also produced by other cell types, the term interleukin has shifted to define a broader set of cytokines and is often used to label newly discovered cytokines. Cytokines have also been classified according to their various functions and can be found within the following groups. Interleukins are by far the broadest grouping of cytokines as the term interleukin is derived from the Latin "occurring between leukocytes." That definition can explain the actions of almost all cytokines and is likely why newly discovered cytokines are given that designation. Chemokines are small (8–10 kDa) and are so named due to their ability to attract other cells. Lymphokines include those cytokines produced by lymphocytes; thus, some cytokines such as IL-2, IL-6, and IL-10 can be both a lymphokine and an interleukin. Interferons (IFN) are a subset of cytokines that share a common ability to combat cancer and viral infection. They are glycoproteins and segregated into two groups, type I (IFNα and IFNβ) and type II (IFNγ). The original Tumor necrosis factor (TNF) family members include TNFα and TNFβ, the latter is now referred to as lymphotoxin alpha (LTα). Originally named for their ability to mediate tumor cytotoxicity, TNF family members (CD40L, FasL, etc.) are now identified based on sequence and structural similarities and mediate a variety of functions from the regulation of cell differentiation to cell survival. Table 1 summarizes common cytokines, their cellular source, function, and molecular weight.

The use of cytokines to combat cancer and autoimmunity has been heavily studied as many cytokines have well defined functions and can be synthetically engineered relatively easily. Cytokines approved for use in cancer therapy include IFNα, IL-2 and

Table 1
Common cytokines, their cellular source, function, and size [11]

Cytokine	Produced by	Actions	Size (kDa)[a]
IL-2	T cells	T cell proliferation	16, monomer
IL-4	T cell, mast cells	B cell activation and Th2 cell differentiation	16 and 18[b], monomer
IL-6	T cell, macrophages, endothelial cells	T and B cell growth and differentiation	24, monomer
IL-7	Thymic stromal cells, lymphatic endothelial cells	Homeostatic proliferation of naïve T cells and memory CD8 T cells	15 and 20[b], monomer and dimmers
IL-10	Monocytes, Th2 T cells, and Treg	Macrophage suppression and inhibition of Th1 cell differentiation	21, homodimer
IL-12	Macrophages and dendritic cells	NK cell activation Th1 cell differentiation	37 and 25[b], heterodimer
IL-15	Mononuclear phagocytes	Stimulation of T and NK cell growth and promotion of memory CD8 T cell survival	15 and 18[b], monomer
IL-17	T cell and macrophages	Induction of epithelial, endothelial and fibroblasts to produce proinflammatory cytokines	18, homodimer
IL-21	Th2 and Th17 T cells	Induction of T, B, and NK cell proliferation	17 and 18[b], monomer
CD40 L (CD 154)	T cells and mast cells	Activation of B cells and class switching	29, monomer, dimer and trimer
Lymphotoxin (LT, TNFβ)	Th1 and CD8 T cells	Activation of macrophages and neutrophils and inhibition of T cells and tumors	22, 33[c], trimer
Interferon-α	Leukocytes and dendritic cells	Antiviral	19, monomer
Interferon-β	Fibroblasts	Antiviral	19, monomer
Interferon-γ	Th1, CD8 T cells and NK cells	Activation of macrophages and NK cells. Inhibition of Th2 differentiation	19, 25[b], dimer and tetramer
C-CSF	Fibroblasts and monocytes	Promotion of neutrophil development	22, monomer
GM-CSF	T cells and macrophages	Promotion of dendritic cell differentiation. Activation of macrophages	16, 35[c] monomer
M-CSF	T cells, bone marrow stromal cells, osteoblasts	Promotion of macrophage development	α:29 β:60 Δ:60 Homodimer

(continued)

Table 1
(continued)

Cytokine	Produced by	Actions	Size (kDa)[a]
TGF-β	CD4 T cells and T regs	Inhibition of T cell growth but promotion of survival. Inhibition of macrophage activation	1:44 2:48 3:47 25[d], homo and heterodimer
TNF-α	T cells	Inhibition of tumorigenesis	26, 17[c], dimer and trimer

[a]Unless otherwise indicated sizes are for the unprocessed precursor protein
[b]Isoform
[c]Active, cleaved protein
[d]Glycosylated

GM-CSF [12]. And have been shown to support the role of both innate and adaptive immune responses to combat cancer. Cytokines used to treat autoimmune diseases are usually of an anti-inflammatory nature, and include IFNβ for Multiple Sclerosis [13], and IL-10 for psoriasis [14] and Crohn's disease [15]. Due to the potent nature of many cytokines, and reports of toxicity related with intravenous delivery, new approaches for cytokine therapy include the targeted delivery of cytokines to tissues using antibodies, or the slow release of cytokines encapsulated inside liposomes.

More recently, with advancing technology that increases detection sensitivity and throughput of biological samples, the measurement of cytokine concentration in serum and tissues have been considered as a diagnostic criterion. Clinical researchers have found unique cytokine profiles in disease states such as heart failure [16] and liver toxicity [17]. While the presence of a cytokine cannot point to a specific disease as can be done with a single antibody, the cytokine milieu can provide a fingerprint of a person's state of health that can be used to diagnose arguably any disease.

3 Inflammation and the Innate Response

In humans, the skin and mucosal membranes comprise the first line of defense barrier against a pathogen. When disrupted by infection or lesion the innate immune mechanisms in the skin begin to activate almost immediately.

One of the fastest acting mechanisms of the immune system is the complement reaction. When blood resident inactive complement proteins encounter a pathogen they will rapidly bind to it, either through direct interaction or in cooperation with a pathogen

specific antibody. Upon binding to a pathogen, a series of proteolytic cleavages will activate the complement proteins; this causes the formation of large multimeric complexes that in turn disrupt bacterial membranes, killing an invading pathogen directly.

Macrophages, DCs, and neutrophils form the next wave of the immune response to follow the acellular mechanisms such as complement. These innate immune cells are able to sense and home to danger signals associated with damage and/or infection. In addition, they will continuously sample their local microenvironments through phagocytosis. Pathogen associated molecular patterns (PAMPS) such as lipopolysaccharide (LPS), flagellin, single stranded RNA, unmethylated CpG, among others, bind to cell surface toll-like receptors (TLRs) on innate immune cells, causing the cells to become activated. Activated innate immune cells will begin production of proinflammatory cytokines and interferons, further amplifying the inflammatory response. Inflammatory cytokines may also have the effect of interrupting the spread of a pathogen; for example IFNα and IFNβ can make nearby cells much more resistant to infection with a viral pathogen.

Similar to PAMPs, a new class of immune activating molecules known as damage associated molecular patterns (DAMPs) has recently been defined. This group of immune stimulators is mostly composed of intracellular molecules which, once released into the extracellular environment, can activate innate immune cells and induce inflammation. The immune activating function of DAMPs may be particularly important in the case of cancer, as these altered cells lack any foreign molecular patterns which would classically be necessary to induce an immune response. The cellular damage incurred by unregulated growth of cancer cells can drive the recruitment of innate immune cells such as NK cells which can target and eliminate cancer.

In many cases the infecting pathogen is killed by the swift actions of the innate immune response described above. However in some cases more action is required. While some innate immune cells are involved in the direct pathogen attack, macrophages, and DCs will also transport pathogen and/or pathogen debris into a lymph node. Naïve T and B cells continuously circulate from the blood to the lymph nodes via specialized microvessels, the high endothelial venules (HEVs). HEV's express a variety of adhesion molecules such as selectins, integrins, members of the immunoglobulin superfamily and some mucin-like molecules that mediate the interactions with lymphocytes which allow them to migrate into the lymph node (Fig. 6). Antigen specific lymphocytes in lymphoid tissues are activated by APCs, which marks the beginning of a slower but much more targeted adaptive immune response.

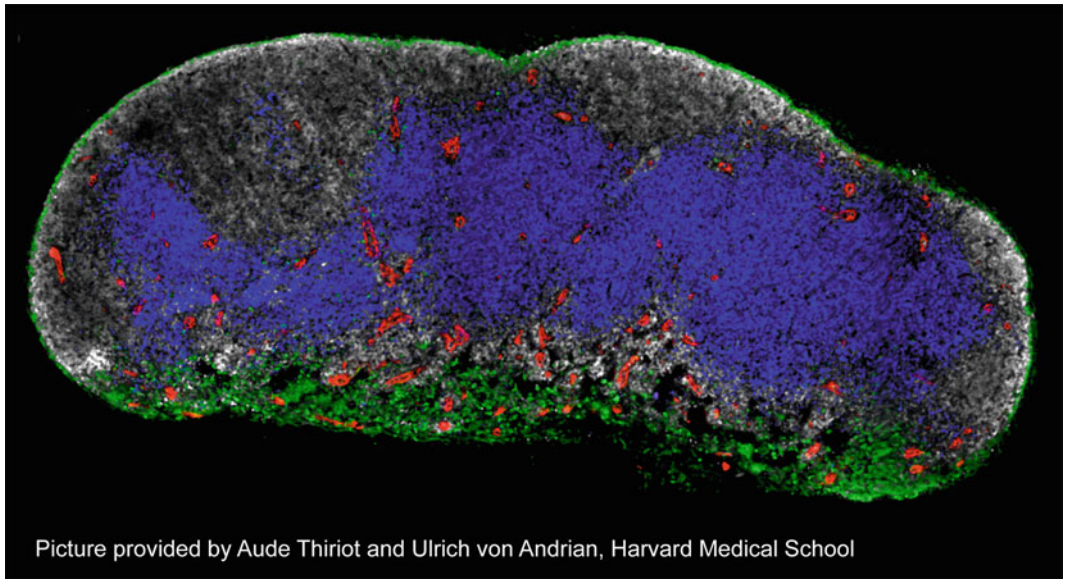


Fig. 6 The immune system relies on specialized organs, such as the lymph node shown above, where innate and adaptive immune cells come together. Macrophages (CD169-*green*) and dendritic cells phagocytose pathogens and digest them into small pieces. Dendritic cells then present small peptide fragments of the pathogens on their cell surface. T cells (TCRβ-*blue*) are in turn activated by the antigen presenting dendritic cells, and begin to proliferate. Additionally, B cells (CD19-*white*) can be activated by T cells and start to produce pathogen specific antibody. Venular endothelial cells are also shown (CD31-*red*)

4 Immune Activation and the Acquired Response

In contrast to the broad specificity of innate immune cells towards pathogens, adaptive immune cells are antigen specific which means that of the billion or so T and B cells in the human body only a handful (~10–100) are specific for any given pathogen protein. To streamline the interaction of an APC with its antigen specific T cell, T cells will circulate continuously throughout the lymphoid organs increasing their opportunity to eventually encounter their target.

As described earlier, APCs use a degradative pathway to break down phagocytosed debris into peptide pieces. These peptides become bound to MHC molecules and are then shuttled to the cell surface where they can encounter the TCR of adjacent T cells. There are two types of MHC receptors; MHC class I and MHC class II. Generally, proteins derived from an intracellular pathogen (viruses and bacteria) will bind to MHC-I, while those proteins derived from the extracellular environment (bacteria and parasites) will be bound to MHC-II. While all nucleated cells express MHC class I, only a handful of cells are capable of expressing MHC class II. Of all the cells capable of expressing MHC molecules, DCs are considered the most adept at activating and inducing the proliferation of naive T and B cells because they possess co-stimulatory

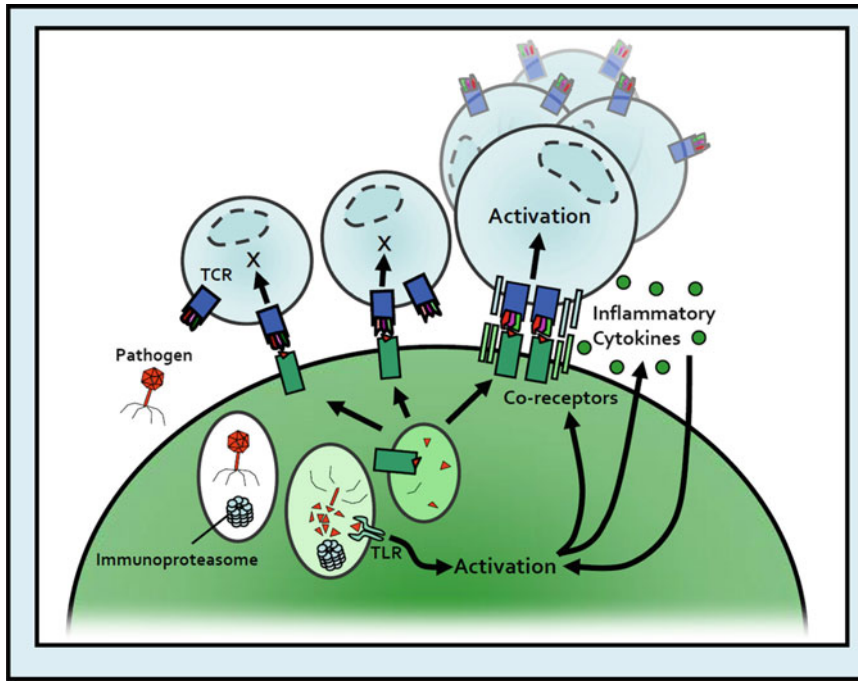


Fig. 7 An adaptive immune response relies on the sequential activation of different types of immune cells. Antigen presenting cells (*green*) become activated when they encounter a pathogen. Pathogen derived peptides bind to MHC molecules on antigen presenting cells and engage the T cell receptor on T cells. Only strong receptor ligand interactions will cause the activation of T cells

molecules and migrate more efficiently through lymphoid tissue to encounter a T cell.

The vast majority of TCR/MHC-peptide interactions will be of low affinity, last for a short period of time, and will not lead to T cell activation. However when a T cell encounters an APC and its T cell receptor has a high affinity for the MHC-peptide complex, the cell–cell interaction is reinforced with adhesion receptor/ligand interactions and the cells remain in contact for a longer period of time. It is important to note that the APC will express these co-receptors only if it has received adequate activation signals from inflammatory cytokines and/or directly through stimulation of its own TLRs. In addition to this, inflammatory cytokines released from the APC or other inflammatory cells can have a direct impact on the T-cells. Thus, a T cell can only become activated if it has the T cell receptor specific for the pathogen peptide in addition to the cytokine signals derived from inflammation associated with an infection (Fig. 7).

The point of interaction between the T cell and the antigen presenting cell is known as the immunological synapse. Upon T cell activation, many biochemical signaling complexes cause the rapid clonal proliferation of the T cell, as well as the secretion of IL-2 by

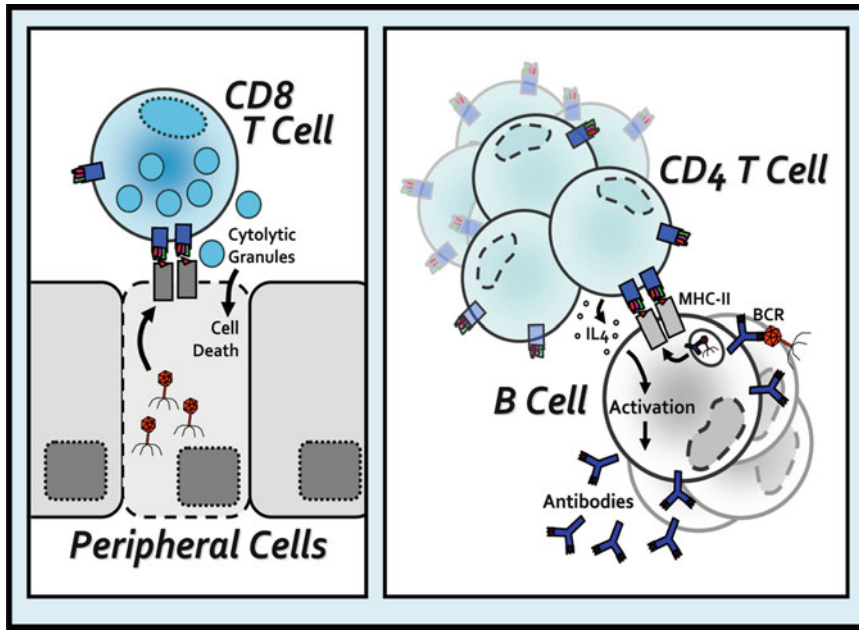


Fig. 8 T cell subsets have different functions. Activated CD8 T cells (*left*) move into peripheral organs and scan the MHC-I-peptide complexes presented by all cells. If the CD8 cell detects its target antigen then it will form a strong bond with the cell and transmit a death signal which kills the infected cell. In the case of activated Th2 CD4 T cells (*right*), these cells will bind to MHC-II-antigen complexes presented by B cells in the lymph node. If the CD4 T cell detects its target antigen then it will bind strongly and transmit an activation signal to the B cell. Activated B cells will then begin to produce antibody

T cells that supports their proliferation. During this rapid expansion, activated T-cells undergo differentiation to become "active" cells capable of exerting either direct cytolytic killing or rapid cytokine production within 7 days.

Depending on which subtype the T cells belong to, different "active" functions are performed. In the case of CD8 T cells, they move into the periphery where they scan for their cognate ligand, namely the MHC-I receptor bound to peptide. MHC-I receptor is present on all nucleated cells, therefore almost any cell type that becomes infected is subject to CD8 T cell scanning. If the activated CD8 T cell binds tightly to a target cell, it will release cytolytic granules into the cell synapse that penetrate the target cell causing its death. This allows for the clearance of intracellular pathogens by killing the infected cells. In the case of "helper" CD4 T cells, they can have divergent effects depending on the type of activation they received. If activated in the presence of IL-12 and IFNγ, CD4 T cells will aid in the clearance of a pathogen via the activation of CD8 T cells. This type of cell specific response is known as a type-I T-helper (Th1) response and is usually caused by intracellular infections caused by viruses, bacteria and protozoan. When CD4 T cells are activated in the presence of IL-4, they cause the activation of a B cell mediated antibody response. This type of CD4 T cell response is known as a type-II T-helper (Th2) response (Fig. 8).

Table 2
The various forms of antibodies and their functions

Antibodies		
Proteins, glycoproteins, and polysaccharides on the surfaces of pathogens, are all potential targets for antibodies. The specific area of the macromolecule which is targeted by the antibody is known as the epitope. B cells that have never encountered antigen initially express exclusively IgM and IgD isotypes in a membrane form. Once the cell has become activated through its interaction with a CD4 T cell (described earlier), it will initially release IgM antibodies into the serum in a pentameric form. Over time, the B cell may undergo additional genetic changes in a process known as *class switching* or *isotype switching*. Through switching, the variable region produced by VDJ recombination is matched with a new constant region, altering the function of the antibody. Depending on the specific mix of cytokines that the B cell receives during activation, it may produce *IgG*, *IgA*, or *IgE* antibody types		
Antibody	*Functions(s)*	*Location(s)*
IgM	Naïve BCR Early antibody production Activates the complement system	Membrane bound, released as a pentamer
IgD	Naïve BCR	Membrane bound
IgG	Directly neutralize target proteins Mark targets for phagocytosis Mark targets for neutrophil degranulation Aid in complement activation	Most common antibody in the bodily fluids Found in monomeric form Low levels at mucosal sites
IgA	Specialized for neutralization of targets—weakly induces phagocytosis or complement	The main antibody found at mucosal sites (e.g., intestine and lung) in a dimeric form
IgE	Specialized for activating mast cells which can induce rapid responses, such as in allergies	Bound to surface receptors upon mast cells

Naive B-cells in the lymph node will display a cell bound antibody on their cell surface, known as the BCR. By displaying their BCR to the extracellular environment the cells are constantly scanning, waiting for something to bind its receptor. Once the BCR is engaged, the receptor/ligand complex will then be internalized by the B cell. In some cases an entire pathogen is internalized with the receptor complex. The B cell then digests the pathogen and presents the antigens upon an MHC-II receptor. If the B cell then encounters an activated CD4 T cell with its cognate T cell receptor, it will form a synapse with the CD4 T cell. Through the release of several cytokines (such as IL-4) the B cell will become activated. Following activation, B cells will undergo massive expansion and differentiation, similar to what occurs in T cell activation. Unlike T cells however, B cells will change the way in which their specific receptor is expressed. Rather than displaying their BCR on the cell surface, the B cell will release its receptor as an antibody. At this point, the B cell is known as a plasma cell, and will produce a large quantity of antibody which will be released into the blood where it can specifically bind to and inactivate pathogens (Table 2).

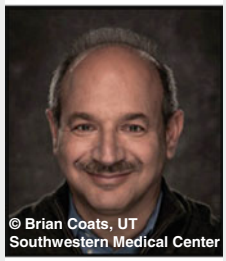
Unlike T cells, activated B cells can further improve the affinity of their immune receptors through mutation and selection. Those B cells bearing an enhanced receptor will be able to better bind their target pathogen, better recruit associated T cells, and will ultimately receive more activation signals. Thus through successive rounds of target binding, activation, and mutation, a B cell with significantly higher affinity can be selected.

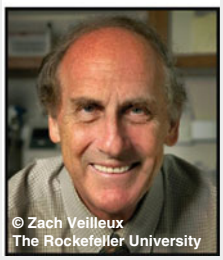
After lymphocytes become activated and the infecting pathogen is neutralized a majority of the lymphocytes will perish and about 5 % go on to form immunological memory. These memory lymphocytes take up residence within lymph node compartments and can survive there for decades. In the case of reinfection with the same or a similar pathogen, the memory cells react much more quickly, compared to naive lymphocytes, and can yield protective responses within 2–3 days instead of the 7 or more days it takes to generate a primary response.

The cells of the adaptive immune response possess receptors capable of binding to a nearly infinite number of pathogen molecules, but they can also lead to misdirected and harmful immune responses as seen in autoimmunity. In some cases, T cells can become activated to respond against antigens expressed by host cells. Naturally, an immune response directed against self antigens is highly undesirable, thus the immune system has developed mechanisms to stop this from happening. One example of this is the negative selection of self-interacting T cells which occurs in the thymus as T cells develop. Despite this and other protective mechanisms, the immune response can still sometimes become misdirected. In the case of rheumatic fever for example, antibodies developed against a strain of *Streptococcus* bacteria can cross-react with antigens in the heart, which can lead to serious heart damage.

5 Conclusion

Despite the abundance of basic scientific research in the field of immunology, it remains a science intimately bound to practical medical benefit. The development of vaccines is considered to be immunology's greatest success, as evidenced by the millions of lives saved. However, of the 500 known human infectious diseases, we are actively vaccinating against only 17 of them [18]. While immunologists have had great successes in the past, such as the elimination of smallpox and the near eradication of polio, many more complex diseases continue to challenge us (Fig. 9). To compete with rapidly evolving pathogens like influenza and drug-resistant tuberculosis, we need state of the art technologies to quickly identify the molecular changes in pathogens and immune responses. With modern techniques in proteomics, such as mass spectrometry, it





Dr.Jules A. Hoffmann (top), Dr. Bruce A. Beutler (middle) and Dr. Ralph M. Steinman (bottom) split the Nobel Prize in Physiology or Medicine in 2011.

The study of the interactions between the innate and adaptive immune systems has brought about important discoveries in immunology. An important example is the discovery of dendritic cells as the major source of T and B cell activation. With a clearer understanding of how dendritic cells operate, new therapies can be designed to target these cells making it possible to stimulate robust immune responses towards illnesses previously difficult to treat and prevent such as those caused by viruses, intracellular bacteria, and cancer.

In 2011, the Nobel prize for Physiology or Medicine was divided between Dr. Bruce A. Beutler, M.D and Dr. Jules A. Hoffmann, Ph.D., for their discovery of the toll like receptor responsible for innate cell activation and to Dr. Ralph M. Steinman, M.D., for his discovery of the cell that bridges innate and adaptive immunity, the dendritic cell. Dr. Steinman spent his life charting out a path for the use of dendritic cells in combating chronic infections such as tuberculosis, HIV and cancer. When Dr.'s Hoffmann and Beutler discovered that cell surface toll-like receptors were responsible for binding to pathogen debris and activating dendritic cells, driving the cytokine production that guided T and B cell activation, it became possible to create immune modulating designer vaccines. New vaccines can now incorporate toll-like receptor ligands as adjuvants (LPS or ssRNA) that activate dendritic cells in a controlled manner to stimulate T or B cell responses.

These discoveries have also led to a novel approach to cancer therapy, termed dendritic cell immunotherapy. When Dr. Steinman was diagnosed with pancreatic cancer in March of 2007 he teamed up with collaborators around the world to design a dendritic cell based therapy to target his own cancer. Dr. Steinman's own dendritic cells were cultured with cytokines *ex vivo*, and exposed to extracts of his tumour. The cytokine groomed dendritic cells, decorated with antigenic peptides from his tumour, were delivered back in the form of a customized vaccine.

Ideally this would have stimulated a robust anti-tumour T cell response. Unfortunately, Dr. Steinman passed away in October of 2011, just 3 days before it was announced he had won the Nobel prize. While it is not known whether the dendritic cell therapy lengthened Dr. Steinman's life, his contributions to science have most certainly advanced the field of cancer immunotherapy forward. In 2010, the first dendritic cell immunotherapy, Sipuleucel-T, was approved to treat prostate cancer. New research into cancer vaccines will continue, drawing on Dr. Steinman's contributions, at the Ralph Steinman center for cancer vaccines at Baylor in Dallas Texas.

Fig. 9 Nobel Prize for Physiology or Medicine winners of 2011

has become possible to quickly identify immune molecules in the lab and the clinic [19]. This technique generates whole proteome results quickly, providing scientists with the data to better understand the mechanisms behind successful immune response so that they can be translated into successful vaccines. It also provides a means to improve upon diagnostics in the clinic by generating a complete proteomic picture of a person's state of health. The faster we can refine our research methods and unveil protective immune mechanisms in humans, the better chance we stand at developing successful therapies and vaccines for the future.

References

1. Jenner E (1749–1823) The three original publications on vaccination against smallpox. The Harvard Classics. 1909–14. www.bartleby.com/38/4/
2. Cavaillon JM (2011) The historical milestones in the understanding of leukocyte biology initiated by Elie Metchnikoff. J Leukoc Biol 90(3):413–424
3. Kantha SS (1991) A centennial review; the 1890 tetanus antitoxin paper of von Behring and Kitasato and the related developments. Keio J Med 40(1):35–39
4. Turk JL (1994) Almroth Wright: phagocytosis and opsonization. J R Soc Med 87(10):576–577
5. Dreyer WJ, Bennett JC (1965) The molecular basis of antibody formation: a paradox. Proc Natl Acad Sci U S A 54(3):864–869
6. Feola M, Merklin R, Cho S, Brockman SK (1977) The terminal pathway of the lymphatic system of the human heart. Ann Thorac Surg 24(6):531–536
7. Medina E (2009) Neutrophil extracellular traps: a strategic tactic to defeat pathogens with potential consequences for the host. J Innate Immun 1(3):176–180
8. Paust S, von Andrian UH (2011) Natural killer cell memory. Nat Immunol 12(6):500–508
9. Berzins SP, Smyth MJ, Baxter AG (2011) Presumed guilty: natural killer T cell defects and human disease. Nat Rev Immunol 11(2):131–142
10. Bettelli E, Oukka M, Kuchroo VK (2007) TH-17 cells in the circle of immunity and autoimmunity. Nat Immunol 8(4):345–350
11. Murphy K, Travers P, Walport M (2008) Janeway's Immunobiology, 7th edn. Garland Science, New York
12. Dillman RO (2011) Cancer immunotherapy. Cancer Biother Radiopharm 26(1):1–64
13. Simon JH, Jacobs LD, Campion M et al (1998) Magnetic resonance studies of intramuscular interferon beta-1a for relapsing multiple sclerosis. The Multiple Sclerosis Collaborative Research Group. Ann Neurol 43(1):79–87
14. Asadullah K, Sterry W, Stephanek K, Jasulaitis D, Leupold M, Audring H, Volk HD, Döcke WD (1998) IL-10 is a key cytokine in psoriasis. Proof of principle by IL-10 therapy: a new therapeutic approach. J Clin Invest 101(4):783–794
15. Narula SK, Cutler D, Grint P (1998) Immunomodulation of Crohn's disease by interleukin-10. Agents Actions Suppl 49:57–65
16. Vistnes M, Christensen G, Omland T (2010) Multiple cytokine biomarkers in heart failure. Expert Rev Mol Diagn 10(2):147–157
17. Lacour S, Gautier JC, Pallardy M, Roberts R (2005) Cytokines as potential biomarkers of liver toxicity. Cancer Biomark 1(1):29–39
18. Osterholm MT, Kelly NS, Manske JM, Ballering KS, Leighton TR, Moore KA (2012) The compelling need for game-changing influenza vaccines. An analysis of the influenza vaccine enterprise and recommendations for the future. CIDRAP, University of Minnesota. 2
19. Vidal M, Chan DW, Gerstein M, Mann M, Omenn GS, Tagle D, Sechi S (2012) The human proteome: a scientific opportunity for transforming diagnostics, therapeutics, and healthcare. Clin Proteomics 9(1):6

Chapter 2

Immunoproteomics: Current Technology and Applications

Kelly M. Fulton and Susan M. Twine

Abstract

The varied landscape of the adaptive immune response is determined by the peptides presented by immune cells, derived from viral or microbial pathogens or cancerous cells. The study of immune biomarkers or antigens is not new and classical methods such as agglutination, enzyme-linked immunosorbent assay, or Western blotting have been used for many years to study the immune response to vaccination or disease. However, in many of these traditional techniques, protein or peptide identification has often been the bottleneck. Recent advances in genomics and proteomics, has led to many of the rapid advances in proteomics approaches. Immunoproteomics describes a rapidly growing collection of approaches that have the common goal of identifying and measuring antigenic peptides or proteins. This includes gel based, array based, mass spectrometry, DNA based, or *in silico* approaches. Immunoproteomics is yielding an understanding of disease and disease progression, vaccine candidates, and biomarkers. This review gives an overview of immunoproteomics and closely related technologies that are used to define the full set of antigens targeted by the immune system during disease.

Key words Immunoproteomics, Mass spectrometry, Antibody, Antigen, Cancer, Infectious disease, SERPA, SEREX, MHC, Epitope

1 Introduction

The landscape of the immune system is constantly changing and is determined by the peptides presented by immune cells, whether from viral or microbial pathogens or cancerous cells. Detection and identification of these immune-active proteins or peptides can therefore be investigated using many of the approaches that have been developed for proteomics studies. As an extension of the proteomics field, the term "immunoproteomics" was first used in 2001 [1]. The field is rapidly expanding and includes increasingly varied techniques that result in the identification of immune related proteins and peptides, derived from invading pathogens, host cells, or immune signalling molecules. The study of immune biomarkers or antigens is not new and classical methods such as agglutination, enzyme-linked immunosorbent assay, or Western blotting have

Kelly M. Fulton and Susan M. Twine (eds.), *Immunoproteomics: Methods and Protocols*, Methods in Molecular Biology, vol. 1061, DOI 10.1007/978-1-62703-589-7_2,

been used for many years to study the immune response to vaccination or disease. However, in many of these traditional techniques, protein or peptide identification has often been the bottleneck. Recent advances in genomics and proteomics, including mass spectrometry instrumentation, has led to many of the rapid advances in immunoproteomics approaches. Immunoproteomics is yielding an understanding of disease and disease progression, vaccine candidates, and biomarkers. Herein, we focus upon providing a broad overview of immunoproteomics and closely related techniques that are used to study the immune response and their role in further disease diagnostics and vaccine development.

2 Immunoproteomics for Characterization of Antibody Targets

One of the two major arms of the adaptive immune system, also classically referred to as the humoral immune response, relies on activated B-cells secreting large amounts of highly specific antibodies, which bind to microbial or cellular targets, either neutralizing them or tagging them for elimination. Antibodies can be generated against microbial invaders, cancer antigens and sometimes misdirected against self-antigens, resulting in autoimmune disease. For a more complete overview of the antibody based immune response, readers are directed to a recent review [2]. Many methods have been developed in order to study the antigen targets of the humoral immune response and in the following section we provide an overview of the most commonly used. Fig. 1 shows a summary overview of these methods.

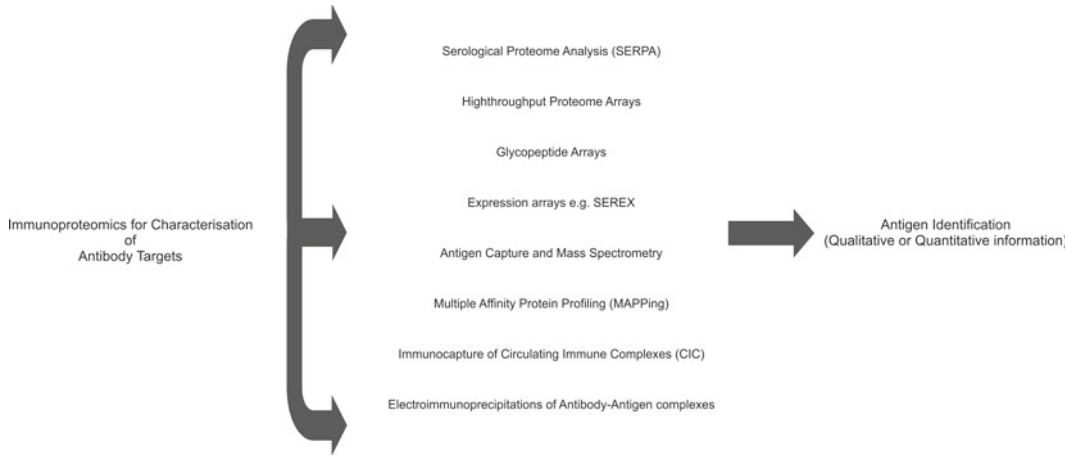


Fig. 1 Overview of methods commonly used to interrogate antigenic targets of the humoral immune response

2.1 Classical Immunoproteomics: Serological Proteome Analysis

One of the most commonly used immunoproteomics approaches relies upon 2D-PAGE, which separates proteins based upon orthogonal physical characteristics. When combined with Western blotting, the technique is commonly known as Serological Proteome Analysis (SERPA). The technique was originally developed in the 1970s and with some refinements popularized for use in biochemistry [3]. Early studies were hampered by challenges in protein identification, and instead used the gel maps to compare protein patterns under different cellular conditions. With many years of refinement, and rapid advances in mass spectrometry and genome sequencing, 2D-PAGE became the mainstay of comparative proteomics studies in the late 1990s and early 2000s. 2D-PAGE can be performed in most protein chemistry labs as a matter of routine, and advances in protein staining and image analysis software have made their use accessible to a broad scientific audience. With well-documented disadvantages, including difficulties in resolving very large, small, hydrophobic or basic proteins and the dynamic range of protein abundance, 2D-PAGE has been superseded by non-gel based proteomics approaches. However, 2D-PAGE has advantages and remains one of the few techniques that allow high quality analysis of intact proteins on a proteome wide scale, including detection of protein posttranslational modifications (PTMs). One of the most overlooked advantages is the ease and efficiency with which 2D-PAGE can interface with other biochemical techniques, such as Western blotting. When combined with Western blotting for detection of antigenic proteins, and mass spectrometry based identification of proteins from in-gel digests, 2D-PAGE provides a powerful approach for antigen identification. Combined, 2D-PAGE and Western blotting is commonly known as *ser*ologic *pr*oteome *a*nalysis (SERPA). The antigen used in these studies can be a whole cell proteome, or subproteome (e.g., membrane fraction). 2D-PAGE resolves the majority of proteins in a sample to a single protein spot, giving the potential to readily identify the antigenic proteins within the resolved proteome. Gels are then transferred to membranes and probed with sera from animal models or humans and developed as per any traditional Western blotting experiment. Many gels can be run in parallel to the blotting experiment, providing gels for reference maps and identification of immunoreactive proteins.

This now "classical" immunoproteomics approach is still widely used, and provides a robust way of screening the antibody reactivity profiles of serum in a variety of disease states, or post vaccination. Applications include discovery of antigenic proteins, biomarkers or correlates of protection, with many studies reporting bacterial diseases [4–29], cancers [30–40] and diseases such as multiple sclerosis [41]. Studies have included discovery of serodiagnostic markers for Q fever [42] and *Helicobacter pylori* [21, 25, 43, 44] as well as diagnostic markers of parasitic diseases, such as

Schistosomiasis [45]. Another report has used SERPA to discover proteins secreted *in vivo* by *Bacillus anthracis* [46]. SERPA has also been used to study the human serological response to vaccination with whole-cell pertussis vaccine [47], *Francisella tularensis* live vaccine strain [48] and human infection with *Francisella* [16, 48]. The latter studies focused upon discovering antibody based correlates of protection.

2.2 High Throughput Proteome Wide Screening of Antibody Targets: The Proteome Array

Some of the limitations of 2D-gel based immunoproteomics have been overcome with the development of proteome or protein arrays to study the humoral immune response. Here, each open reading frame of interest in the genome is amplified by PCR, followed by cloning, protein expression and microarray printing [49, 50]. Bacterial proteomes are sufficiently small that the entire complement of proteins from the genome can be printed on a single array.

The chips are then treated in a manner similar to traditional Western blotting, probed with sera and reactivity detected after incubation with a secondary antibody with fluorescent conjugate. The chip based technology has the advantage of screening closer to equal amounts of antigens, interrogation of the entire theoretical proteome of the organism, and reduced volume of serum required for screening (2 μL vs. ~50–100 μL for large 2D-Western blot). The reduced requirement for serum means that pooling of sera from multiple animals or humans in a study is not required, and individual differences can be readily detected. These benefits, combined with the high throughput capacity of proteome microarrays, make it an attractive method of rapidly screening hundreds of sera. The use of advanced data handling algorithms is a requirement, as with DNA based microarrays, for meaningful data interpretation [51].

The complexity of protein purification and high throughput gene expression systems means that it can be challenging to produce proteome arrays that represent the entire proteome of an organism. In addition, the expressed proteins lack native PTMs, processing and correct protein folding is not guaranteed. Investigation into the use of yeast based protein expression systems may help address the issue of PTMs, however many bacteria elaborate a unique repertoire of glycoconjugates and glycoproteins that cannot be replicated by yeast based systems. Lack of non-protein antigens can be addressed by addition of native molecules to arrays, in order to gain a broader perspective of the humoral immune response. To date, there have been reported advances in array technology that address challenging protein antigens, such as membrane proteins [52], and non-protein antigens, such as carbohydrates [53].

Proteome arrays have been used to study the humoral immune response of a wide range of pathogens, including smallpox vaccination [54, 55], Chlamydia infections [56, 57], Brucellosis [58, 59],

Mycobacterium tuberculosis infections [60, 61], salmonellosis [62], Herpes simplex virus [63, 64], *Plasmodium falciparum* [65–68], Q fever [42, 69, 70], toxoplasmosis [71], *Burkholderia pseudomallei* [72], *Borrelia burgdorferi* [73], *Francisella tularensis* [50, 74], and Epstein-Barr virus [75]. In the long term, this technology has the potential to aid development of improved serodiagnostic tests, vaccine development, epidemiological studies and shed light on the interaction of pathogens with the immune system.

2.3 Deciphering the Immune Response to Glycoprotein Antigens

Carbohydrate moieties and glycoconjugates, including glycoproteins, are increasingly being shown to have roles in various diseases, including cancers and bacterial infections. Protein glycosylation is a highly abundant PTM and aberrant glycosylation of proteins has been shown to be associated with cancers [76] and autoimmune diseases [77, 78]. Truncated glycan moieties on glycoproteins are recognized as nonself and result in the generation of autoantibodies to glycopeptide epitopes [79, 80]. For example, *O*-glycosylation of mucin (MUC1) is particularly important in cancers, with patients reported to have autoantibodies to distinct epitopes on MUC1 that harbor truncated sugar moieties [81]. Of note, these autoantibodies recognize cancer specific epitopes, composed of the combined peptide sequence and the carbohydrate moiety [76, 82, 83]. It is, therefore, likely that there are other glycopeptide antigens in cancers. Investigation of glycan associated autoantibodies has been carried out using variations of chip based screening technologies. These have included a microarray display platform that allows the large scale screening of *O*-glycopeptide libraries for the investigation of disease associated autoantibodies [80, 84–87].

Recently, a high throughput chemoenzymatic synthesis and microarray display platform has been described that enables the production and screening of large *O*-glycopeptide libraries for disease associated autoantibodies. A combined synthetic and enzymatic approach allowed immobilization and generation of a glycopeptide epitope library on a microarray chip. As outlined in Fig. 2, *O*-linked GlcNAc containing peptides were synthesized by standard solid-phase peptide synthesis (SPPS) [84]. These glycopeptides were then immobilized on microarray plates coated with amine-reactive NHS-ester groups. This was followed by on-slide glycosylation with different polypeptide GalNAc-transferases and other elongating glycosyltransferases. In this way, a diverse library of synthetic *O*-glycosylated MUC1-peptides was generated *in situ*. This was used for serological screening and the results showed that the array was able to detect autoantibodies in the sera of patients with a confirmed diagnosis of breast cancer [84]. Rapidly synthesized libraries which represent the potential diversity of glycopeptide or glycoprotein epitopes pave the way to broader screening of glycan-epitopes and the elucidation of glycan epitopes within existing immunodominant peptides.

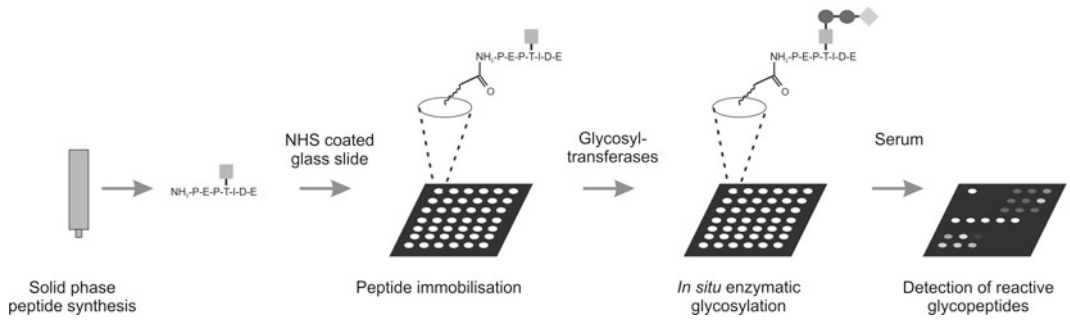


Fig. 2 Uncovering glycopeptides epitope. Peptides are synthesized using solid phase peptide synthesis, including amino acid harboring an *N*-acetyl glucosamine residue. Peptides are then immobilized on glass slides, coated with NHS esters. This serves as a partial purification step. Addition of glycosyltransferases allow *in situ* addition of carbohydrate moieties to generate a library of glycopeptide epitopes. This is then screened with sera and reactive epitopes identified

The diversity of carbohydrate moieties across the domain of Bacteria is substantially greater than that of eukaryotes. Many monosaccharides are found exclusively within bacteria and are genus, species, or strain specific. Consequently, these unique sugars are often readily identified by the host immune system as foreign entities during infection. Frequently these sugars are part of a pathogen associated molecular pattern (PAMP), such as lipopolysaccharide (LPS) or peptidoglycan, that is recognized by host pattern recognition receptors (PRR) such as toll-like receptors (TLR) [88, 89] or nucleotide oligomerization domains (Nod) [90–93] as part of an innate immune response. However, it is increasingly being reported that bacterial glycoproteins also play a role in stimulating innate [94] and adaptive [95–98] host immune responses. Several of these pathogen glycoproteins, including the flagellin of *Campylobacter coli* and *Campylobacter jejuni*, are responsible for serospecific antibody responses [95, 98]. Additionally, the antibody response to anthrose, a unique sugar decorating the *Bacillus anthracis* exosporium glycoproteins (BclA and BclB), is currently being exploited for its potential use in detection and diagnosis of anthrax [99–102]. Despite a growing recognition of the importance of bacterial glycoprotein antigens, immunproteomics methods directed specifically towards their identification are lacking. In fact, glycoprotein antigenicity is frequently discovered as a consequence of targeted glycoprotein characterization. Given the documented importance of bacterial glycoprotein antigens, methods designed for their global detection and identification would greatly benefit the field of immunoproteomics.

2.4 Antigen Discovery Using Expression Arrays

Expression arrays are composed of bacterial, yeast, mammalian, or cell free cDNA expression libraries that are used to identify novel antigens. Known as *se*rological analysis of *r*ecombinant cDNA

*ex*pression libraries (SEREX), these techniques have a large genetic component and have been termed by some as "reverse proteomics" [103]. SEREX was first developed for analysis of the humoral response to cancer in the 1990s [104], with the goal of identifying tumor specific antigens that elicit high titer immunoglobulin G (IgG) antibodies in patient sera. In this context, the technique permits the search for antibody responses and the molecular definition of immunogenic tumor proteins, based upon autologous patient sera (reviewed in ref. 103). Patient tumor mRNA is used to prepare prokaryotically expressed cDNA libraries which are then immunoscreened with absorbed and diluted patients' sera for the detection of tumor antigens that have elicited a high-titer IgG humoral response. This approach has the advantage of being able to identify antigens expressed *in vivo*, and is unbiased, based only upon the reactivity of clones with autologous patient sera. A second phase of screening is also carried out, using sera from normal patients in order to define antigens that show cancer-restricted immune recognition [105, 106]. SEREX has been applied to the study of many cancer types, including renal [105, 107, 108], colon [109–111] and breast [106, 112–123] cancers leading to the identification of cancer specific antigens. One antigen, NY-ESO-1, was identified in esophageal squamous cell carcinoma and the gene expressed in normal testis and ovary, with aberrant expression in various types of malignant tumors [124]. NY-ESO-1 shows restricted expression patterns, elicits both cell mediated and humoral immune responses [125] and has been under development as a cancer vaccine target (reviewed in ref. 126).

Despite many advantages, SEREX presents some challenges, in that it is time consuming to construct cDNA libraries for each tumor sample. In addition, false positives are possible, either due to reactivity with prokaryotic expression components or lack of expression of PTM in prokaryotic expression systems. In particular, protein glycosylation of eukaryotic proteins can represent important antigenic epitopes, including disease associated changes in glycosylation. A few autoantibodies to PTM-protein epitopes have been reported, including those found in cancers [79, 80] and autoimmune diseases [77]. The use of eukaryotic expression systems can ensure that expressed proteins are glycosylated [127–130]. Tumor associated antigens identified from SEREX screening are updated in the Cancer Immunome database (ref. 131; http://ludwig-sun5.unil.ch/CancerImmunomeDB/). Over 2000 autoantigens are listed in this online database. An excellent review that discusses the classes of SEREX defined antigens and the wider impact of this technique upon cancer vaccine and diagnostic development can be found here [103].

2.5 Antigen Capture and Mass Spectrometry

Immunocapture mass spectrometry aims to enrich antigen proteins from cell lysates, using mass spectrometry as the final means to identify captured proteins. There are many variants of

immunocapture and generally immunoglobulins from patient sera are immobilized on Protein A or Protein G, usually in column format. This is followed by the application of a cell or tissue lysate to the column, effectively enriching for antigenic proteins, i.e., those proteins to which there are antibodies in patient serum. Proteins are eluted from the column, enzymatically digested and subsequently identified by MS/MS [132].

2.5.1 Multiple Affinity Protein Profiling

*M*ultiple *A*ffinity *P*rotein *Profil**ing* (MAPPing) is an example of an immunocapture technique that has primarily been exploited to identify cancer related autoantigens [133, 134]. It is based upon two-dimensional immunoaffinity chromatography, whereby antigens from tumor lysates are separated based upon their affinity for immunoglobulins from healthy controls in the first dimension and immunoglobulins from cancer patients in the second. The first dimension removes autoantigens that are recognized by sera from healthy patients. Cancer restricted autoantigens then flow through to the second column, which then selectively binds them. The proteins eluted from the second chromatography step are therefore likely to be cancer specific and are identified by enzymatic digestion and MS/MS analyses [133, 134].

2.5.2 Capture and Identification of Circulating Immune Complexes

Another variation of immunocapture targets circulating immune complexes (CIC). Immune complexes are formed from the noncovalent interaction between antigens and antibodies and are usually removed by mononuclear phagocytes through complement receptors and Fc-receptors [135]. This process constantly occurs in healthy individuals and ensures the rapid clearance of denatured proteins, antigens of gut bacteria or dead cells. Studies have shown that these antigen–antibody complexes can play a role in disease progression of human autoimmune diseases [136], cancer [137], or infectious diseases [138]. There is some discrepancy in the literature regarding the utility of CIC in disease diagnosis, treatment or as an indicator of disease severity [139–144]. Some have argued that identification of antigens incorporated into CICs may be of greater relevance than information regarding free antigens [144], and that antigens in CICs could provide information useful to understanding disease progression, and in developing diagnostic and treatment strategies.

CICs can be isolated from serum, as described in a recent report [144]. Patient serum was immobilized on a Protein A or G column and cell lysates passed over the column. Proteins that were bound to the immobilized patient sera were eluted and identified using tandem mass spectrometry of their tryptic digests [143–145]. A recent study identified CICs containing the proteins thrombospondin-1 and platelet factor 4 in the serum of 81 and 52 % of a sampling of rheumatoid arthritis patients, respectively [143]. This method is applicable to many other diseases for inventory of antigens within CICs.

2.5.3 Electroimmunoprecipitation of Antigen–Antibody Complexes

Electroimmunoprecipitation can exploit differences in electrophoretic mobility between an antibody and its corresponding antigen, resulting antigen–antibody complexes embedded in an agarose gel. Staining of the gels permits visualization of precipitated complexes. Elution of these complexes, plus enzymatic digestion, and subsequent mass spectrometry analysis can identify the unknown antigenic proteins of interest [146]. In rocket immunoelectrophoresis (RIE), a monoclonal antibody is used. However, crossed immunoelectrophoresis (CIE) involves two dimensions of separation [147, 148] and can therefore be used to identify antigenic proteins reacting with mixtures of monoclonal antibodies, polyclonal antibodies, or serum. Therefore, electroimmunoprecipitation can be used to capture antigens relevant to various disease states or contribute to validation of antigenic proteins. Electroimmunoprecipitation has the added advantage of being quantitative [147–149] and can therefore also be used to monitor the level of serum antibody response to a known antigen.

2.6 Epitope Mapping

Discovery of antigenic proteins is the first step in profiling the humoral immune response to disease. There is often a need to then further dissect the immune response and determine the region of the antigenic protein, or epitope, that stimulates the immune response; particularly in antibody design or epitope based vaccine design [150]. This can be carried out using a wide variety of techniques, a full description of which is beyond the scope of this review and we direct the reader to recent reviews [151, 152].

3 Immunoproteomics in the Study of Major Histocompatibility Complex Peptides

The cell mediated immunity (CMI) arm of the adaptive immune response involves activation of cell populations such as phagocytes or T-cells and can include the release of communicator molecules, such as cytokines and chemokines in response to foreign invaders or antigens. T-cells recognize antigens that are displayed on the surface of host cells in complexes known as the major histocompatibility complex (MHC). The antigens found in complex with MHC molecules are short peptides that are derived from intracellular proteolysis of proteins. This antigen presentation and processing allows for the host recognition of foreign peptides from infected or transformed cells, by stimulating an immune response. In addition, there is constant surveillance of peptides derived from the host organism, and self-peptide presentation is involved in T-cell development in the thymus and regulation of self-tolerance.

There are two major subgroups of MHCs, denoted MHC I and MHC II, which are encoded by the human leukocyte antigen (HLA) gene clusters. These gene clusters are highly polymorphic, giving rise to hundreds of allelic forms, with only a subset present

in each individual. The polymorphism gives rise to differences in the MHC molecules, their binding pockets and affinity for particular peptide antigens, thereby influencing the repertoire of antigens presented to the immune system of an individual. The two major classes of MHC molecules (class I and II) are distinct in their three dimensional structure, pathways by which antigens are processed and the type of T-cell with which they interact. MHC class I gene cluster encodes the heterodimeric proteins that bind antigenic peptide from within cells, and are found on all nucleated cells types. MHC Class I molecules carrying peptide antigens complex with the CD8 co-receptor. This complex is primarily recognized by cytotoxic T-cells and leads to their activation and eventual death of the cell expressing the nonself antigen.

In comparison, MHC class II gene cluster encodes heterodimeric peptide-binding proteins and proteins that control peptides binding to the MHC heterodimers. Peptide loading onto MHC class II molecules occurs in the lysosomal pathway and MHC class II complexes are only found on specialized cell types, such as B-cells, neutrophils, and dendritic cells and can be induced on macrophages and human T cells. The CD4 T cell co-receptor recognizes MHC Class II antigen complexes, also resulting in T-cell activation. If the presented peptide is foreign, the T cells then proliferate, secrete cytokines, and differentiate into antigen-specific effector CD4 cells, which secrete cytokines and activate other cell types, such as B-cells. For both MHC class I and class II molecules, the antigens are peptide fragments which are recognized as nonself by T-cells, these antigens are known as T-cell epitopes. A detailed description of how these peptide fragments are generated is described in more detail [153–155].

3.1 MHC Peptide Enrichment

The identification and characterization of peptides displayed by MHC molecules and specific T-cell epitopes has become essential for modern immunological studies, in many aspects of basic and applied research. For example, the development of vaccines with enhanced T-cell immune response [156–158]. A broad array of functional and biochemical approaches have been developed to identify peptide epitopes, including forward and reverse immunoproteomics, and mass spectrometry centric approaches (for example refs. 158–168). A recent review describes T-cell epitope mapping based upon screening of peptide libraries and screening for T-cell activation [169]. In the following sections, we review the contributions of mass spectrometry based immunoproteomics to MHC peptide binding and T-cell epitope identification and how this knowledge is furthering vaccine and diagnostic development.

3.2 Mass Spectrometry in MHC Peptide Discovery

MHC class I and II proteins preferentially bind peptides of different lengths and general characteristics. Typically, MHC class I molecules have a binding cleft that accommodates peptides of

8–10 amino acids, whereas, MHC class II molecules bind peptides 8–30 amino acids in length. Peptides that bind the cleft of a MHC class II molecule are usually found to share a core sequence [170–175].

Over the past two decades, several methods of isolating MHC peptides have been developed. Early reports in the 1990s used acid treatment to elute peptides from the surface of cells [176]. Although simple to carry out, peptide elution was not specific to those bound to MHC complexes and difficulties arose when attempting to discriminate specific MHC peptides. Targeted immunoaffinity purification was also reported in the 1990s [177], in which monoclonal antibodies specific for an MHC class were used to enrich the MHC complexes. MHC bound peptides are then eluted by acid treatment and separated from proteins by size exclusion. Soluble MHC molecules, without a transmembrane domain, are secreted in transfected cells with MHC peptides bound. The secreted complexes are easily purified, for example with the use of immunoaffinity columns; this method is considered a facile method to isolate MHC peptides [178]. In all cases, it is assumed that peptides bound to MHC molecules are protected from proteolysis during sample preparation and that acid treatment is sufficient to dissociate peptides from their MHC binding partners. Immunoaffinity purification of MHC peptides has been applied in many areas, including the study of the central nervous system of multiple sclerosis patients [179, 180] and bronchoalveolar lavage cells isolated from patients with sarcoidosis [181]. Another study combines immunoaffinity enrichment with testing of subsequent fractions for biological reactivity, prior to peptide identification by mass spec [182] for the identification of tumor associated antigens. This approach has also been used for the successful identification of novel antigens in primary human breast cancer [183] and West Nile virus [184].

3.3 MHC Peptide Identification

Purified MHC peptides were largely analyzed using Edman degradation. In particular, the shorter length of the MHC class I peptide ligands made them amenable to amino acid sequencing by Edman degradation. The use of Edman chemistry on a pool of MHC class I peptides revealed an increased signal for a particular conserved amino acid, or amino acid position [185], allowing progress towards identifying conserved residues or sequence motifs. MHC class II peptides are less amenable to this approach, due to their longer length and greater heterogeneity. However, other early biochemical studies established consensus binding motifs for both MHC class I and II peptide ligands [170–175].

Due to the limitations in HPLC separation of peptides and Edman sequencing in early studies of MHC peptides, only short sequences of abundant peptides were determined. Pioneering studies in the early 1990s demonstrated the utility of the then

recently developed electrospray ionization mass spectrometry (ESI-MS), in combination with microcapillary HPLC to determine the length and sequence of peptides bound to HLA-A2.1 [159], one of the most widely distributed MHC class I molecules within the human population. Since this study, ESI-MS has been used extensively for the detection of peptides presented by major histocompatibility complex (MHC) molecules (for example refs. 159, 186, 187 and recently reviewed in ref. 188). Mass spectrometry affords the advantage of high resolution peptide mapping, allowing rapid identification of hundreds of MHC peptides in a single experiment.

Since the first report [189], rapid advances in mass spectrometry instrumentation, throughput and data handling mean that mass spectrometry is a widely used technique in the identification of T-cell epitopes. More recently, large scale proteomics methodologies have been used in comparative or quantitative studies of T-cell epitope identification. Studies have reported robust identification of epitopes, and refinements have been made to identify immunodominant epitopes and in distinguishing self and nonself MHC class I peptides. Precise splitting of the eluate from HPLC separation of MHC peptides, with a portion diverted to the mass spectrometer and the majority retained to assay T-cell activity, has allowed more precise correlation between MHC peptide identification and T-cell activation [190–192]. Other methods compared the LC-MS chromatograms of peptides eluted from MHC I complexes with those of reference cells. Mass spectrometry has been used to identify T-cell epitopes of *Plasmodium falciparum* [193], cancers [194, 195] and rheumatoid arthritis [196]. Others have employed novel approaches to hold antigen presenting cells in protein free medium, simplifying the repertoire of peptide antigens presented and reducing the background of peptides normally observed, allowing greater detection of exogenous MHC [197]. Fig. 3 gives an overview of the current workflow for MHC peptide isolation and identification.

3.4 Quantification of MHC Peptides Using Mass Spectrometry

Qualitative studies provide an inventory of detected MHC peptides, and with the development of advanced proteomics technologies comes the opportunity to carry out quantitative studies. Quantification of MHC peptides allows for comparison of peptide repertoire and abundance with time, between tissues, self and nonself, or test and control and between individuals. Quantification can be relative or comparative, achieved using peptide labeling strategies such as the commercially available ICAT system [198], isobaric tags such as iTRAQ [199] or chemical tags (mass coded abundance tagging, MCAT) [200]. A recent study, for example, reported robust identification of over 100 MHC II peptides, and their relative quantification using stable isotope labeling [187].

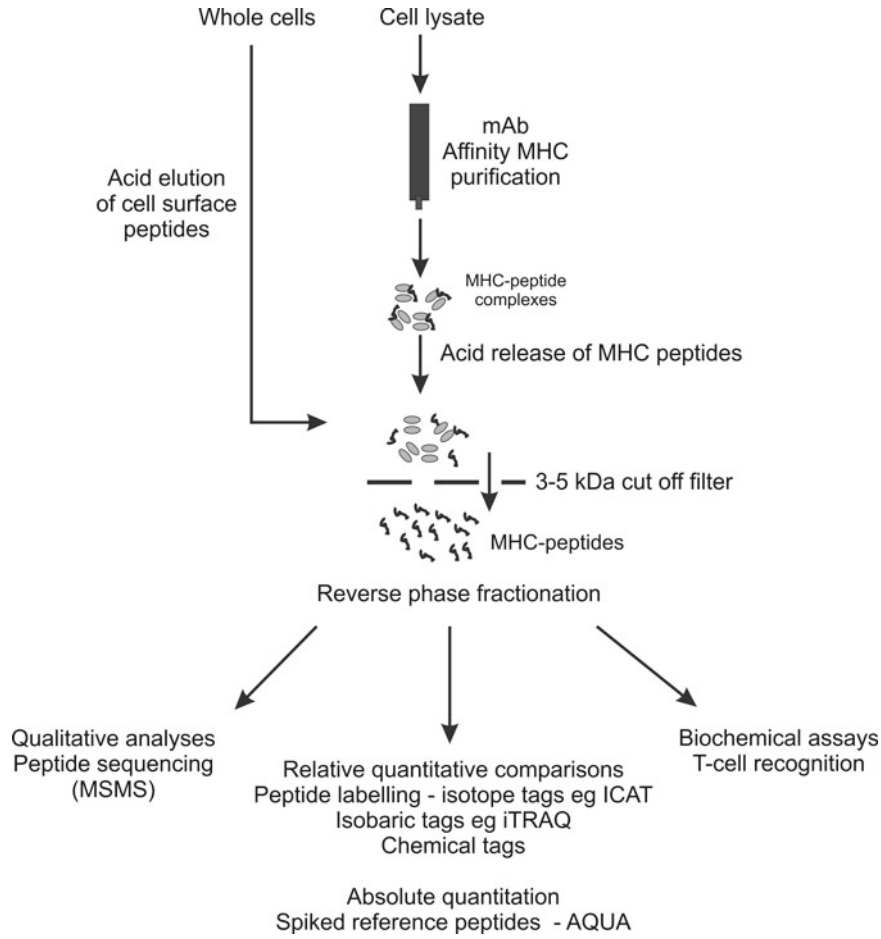


Fig. 3 Schematic overview of methods for MHC peptide purification and sequencing

Others have described the development of a selected reaction monitoring (SRM) method combined with absolute quantitation (AQUA) [201]. Selected reaction monitoring is a highly specific technique that targets specific peptides, with high sensitivity. When used in combination with a deuterated internal calibrant peptide, this permits the absolute quantification of target peptides [202, 203]. This approach was successful in quantifying the amount of a known ovalbumin peptide from the spleens of immunized mice after MHC affinity purification. Recently, the approach has been used to measure the presence and abundance of known MHC melanoma peptide antigens on the surface of several human melanoma cell lines [204]. SRM can be multiplexed for rapid and simultaneous identification and quantitation of hundreds of peptides, is robust and readily transferable between laboratories.

3.5 Characterization of Posttranslational Modifications of T-Cell Epitopes

Other studies have combined mass spectrometry and functional assays for T-cell epitope identification [195]. The genetic polymorphism of the HLA alleles results in variation in the MHC complexes across the population, with differing binding affinities. It can, therefore, prove challenging to identify antigenic MHC I peptides presented by MHC class I molecules that are less frequently found across a population. This is important in the development of peptide-based vaccines for the therapeutic treatment of melanoma and other cancers, which requires the identification of antigenic peptides that will allow the majority of the population, regardless of their MHC encoded phenotype, to stimulate a T-cell response.

Glycosylation is a common PTM of proteins in eukaryotes and increasingly discovered in bacteria. Although largely ignored until recently, carbohydrates, glycolipids, and glycopeptides [205] are now recognized to modulate T-cell recognition [206, 207] having been shown to be presented by MHC complexes [208]. This has important implications in the immune response to pathogens, tumor cells, and self-tolerance. Several studies in the late 1990s provided evidence that naturally modified *O*-GlcNAc peptides were ligands for MHC class I molecules [208–211], and a crystal structure showed the glycan moiety to be exposed for recognition by CD8 T-cells [212]. After affinity enrichment of MHC complexes and elution of bound peptides, many of the techniques developed for the study of glycoproteomes could be applied to target and identify glycopeptide MHC peptides. Some approaches such as those using lectin enrichment have already been successfully employed for the enrichment of MHC bound glycopeptides [208]. Other approaches, such as hydrazide capture [213] and chromatographic enrichments, combined with advanced mass spectrometry approaches, such as precursor ion scanning of signature glycan ions, could lead to rapid and specific identification of MHC glycopeptides.

Similarly, it has been proposed that phosphopeptides may also be T-cell antigens [214], presented by class I MHC molecules on malignant cells [215] or MHC class II [216] and be attractive targets for cancer immunotherapy [217, 218]. Phosphopeptides associated with class I MHC molecules on the surface of tumor cells can be enriched by immunoaffinity purification of the MHC complexes, followed by elution and enrichment of phosphopeptides with immobilized metal-affinity chromatography (IMAC) [214, 216–219].

4 Cytokines

Cytokines are low molecular weight secreted proteins, ranging from 8 to 40 kDa [220, 221], with diverse roles in controlling growth, survival, differentiation and the effector function of cells and tissues (recently reviewed in ref. 220, 221). They are critical to an immune

response, and the secreted profiles of certain cells determine the nature of the response—Th1 versus Th2 and dictate whether the immune response is cell mediated or antibody based. Production of cytokines is tightly regulated, with an uncontrolled response potentially leading to septic shock. Therefore, controlled production of cytokines is key to many aspects of inflammation and immunity, including a balanced immune response. Therefore the types and levels of cytokines can serve as markers of disease progression.

The number of cytokines and closely related growth factors that have been identified has increased dramatically in recent years [222]. Unlike hormones, cytokines are active over short distances at sites of inflammation and can act in combination with other cytokines to give a variety of biological responses. Cytokine profiles can potentially be indicative of a particular disease state, so in order to correlate this, methods are required that can simultaneously measure levels of multiple cytokines. Although some cytokines are produced at ng/mL concentrations in body fluids, most are expressed at pg/mL levels and therefore, the most widely used current methods are based upon immunoassays, RT-PCR or bead based bioassays. Other methods for detecting cytokines or cytokine secreting cells include radioimmunoassay (RIA), immunoradiometric assays [223], cellular enzyme-linked immunosorbent assay (CELISA), cytometric bead array (CBA), radioreceptor assay (RRA), reverse hemolytic plaque assay (RHPA), cell blot assay, and cytokine flow cytometry. Identifying and quantifying the cytokines secreted in response to a disease state or pathogen are of interest in diagnostics and as vaccine correlates of protection. The cytokine quantification assays that have gained popularity have become increasingly high throughput, allowing an increase in the amount of information that can be collected about the roles of cytokines during disease or post vaccination. The use of bead based assays has allowed the multiplex measurement of multiple cytokines simultaneously [224–229]. These assays are robust, but they are inherently biased towards a predetermined panel of cytokines and provide only quantitative information. In addition, these methods provide no information regarding PTM of cytokines, which can be of importance in some cases. For example, IL-24 activity is dependent upon formation of a disulfide bond and glycosylation [230].

Several different immunoproteomics approaches have been reported that are able to detect and quantify cytokines and provide information regarding PTMs. A recently reported technique, known as immunoaffinity capillary electrophoresis (IACE), captures cytokines by immunoaffinity using specific antibodies, then separates the captured proteins using capillary electrophoresis. The resulting protein or peptide fractions are then analyzed by tandem mass spectrometry, providing cytokine identification [231–233]. This two dimensional separation also allows for differentiation between protein isoforms and identification of PTMs. Another

cytokine detection method also exploited an immunoaffinity capture step coupled then directly to analysis by mass spectrometry for protein detection and quantitation [234]. With both methods, the immunoaffinity capture step limits cytokine detection to a predetermined panel. However, in the latter study, the authors' goal was to improve the speed of cytokine detection compared to current assay technologies (1–3 h) [234].

Other reports have focused upon unbiased detection of cytokines in serum, or *in vitro* secretion from immune cells, such as monocytes. Detection of cytokines in serum presents many challenges, characteristic of serum proteomics. Cytokines are typically a very small fraction of the low molecular mass proteome in serum. Although such proteins are amenable to detection using current mass spectrometry technologies, the challenge lies in their low abundance in relation to the high background of other serum proteins. In human serum, albumin and immunoglobulin G (IgG) make up 60–80 % of the total serum protein content [235], potentially masking the detection of low abundance proteins. The challenge of the dynamic range of proteins in serum is not new and there are many strategies for their depletion [236]. Additional concerns arise when albumin, known as the "tramp steamer" of the blood, interacts with many small molecules, fatty acids and proteins, acting as a transient carrier. Depletion of these transiently bound proteins, peptides and small molecules is possible and may distort the low abundance serum proteome. Methods have been reported for separation of low molecular weight serum proteins, using centrifugal ultrafiltration under solvent conditions that disrupt protein–protein interactions. Two dimensional liquid chromatography of tryptically digested proteins and identification using mass spectrometry facilitates the identification of the low MW proteome, including cytokines [237]. Others have also used ultracentrifugation, IEF [238] for identification of low MW serum proteome, while Groessl et al. [239] employed a label free MS based proteomics approach to characterize the human monocyte secretome, successfully identifying important proinflammatory proteins and cytokines. Advances in these mass spectrometry based methods pave the way for rapid, robust and unbiased serum cytokine detection, characterization and quantification during disease. This has the potential to contribute to understanding of disease progression, as well as revealing disease or post vaccination biomarkers.

5 Immunoinformatics

In silico prediction of T or B cell epitopes has become a mainstay of immune related research. This is part of a growing field of immunoinformatics, or computational immunology, which describes the application of informatics technologies to problems of the immune

system. Several studies have used the term "immunomics" to describe the study of the detailed map of immune reactions of a given host interacting with a foreign antigen (the immunome). *In silico* methods have been developed in order to predict the sequence, structure and affinity of various epitopes of the humoral and cell mediated immune systems. As with many rapidly growing fields, the overlap or complementarity between closely related areas means the boundaries are less easily defined. For example, immunoinformatic studies of peptide epitopes is important in immunoproteomics and many studies combine epitope prediction with epitope sequencing. The various algorithms and bioinformatics techniques complement proteomics identification of peptide epitopes, and combined *in silico* and *in vitro* approaches bring more power to peptide identification or mapping. In the following section, we provide a high level overview of the key areas.

5.1 Immunoinformatics and B-Cell Epitopes

B-cell epitopes are antigenic determinants from pathogens (or self) that interact with B-cell receptors [240]. The B-cell receptor contains a hydrophobic binding site composed of hypervariable loops that vary in length and amino acid composition. Epitopes that bind to the receptor are either continuous (linear) or discontinuous (conformational) [241]. According to accumulated knowledge, the majority of B-cell epitopes are discontinuous, with protein folding playing a large role in epitope formation. Prediction tools exist for prediction based upon amino acid sequence (for continuous epitopes) or structure based tools for discontinuous epitopes (for recent examples refs. 242–254). In the past, sequence based prediction tools have used amino acid hydrophobicity scales for epitope prediction. This approach is still used, for example BCIPEPT predicts continuous epitopes using propensity scale values, such as amino acid polarity, flexibility. The BCEPRED server [242] has been reported to predicted continuous B cell epitopes with an efficiency of 58.7 % [245]. Prediction of discontinuous epitopes is more challenging, with over 90 % of B-cell epitopes being discontinuous [255].

For both continuous and discontinuous epitopes, the current gold standard remains X-ray crystallography and observing the points of contact. From the accumulated structural data, several prediction methods have been developed, for example Discotope [243] and mapitope [246, 256]. Discotope combines amino acid statistics with protein spatial information and was trained on a dataset of X-ray crystal structures of antibody–antigen complexes. More detailed overviews of the current methods and databases are given [257, 258].

5.2 Immunoinformatic Prediction of T-Cell Epitopes

In order to accelerate experimental approaches to MHC epitope prediction, computational methods or algorithms have been developed that can predict MHC-binding peptides and their binding

affinity [259]. These approaches fall into two areas, and are either sequence based or structure based [260]. Numerous algorithms are now available to carry out sequence based peptide epitope predictions (reviewed in ref. 261–263). These have the advantage of being fast, potentially screening whole genomes, but require large amounts of experimental data regarding the peptide binding preferences of the MHC molecule of interest. In comparison, structure based epitope modeling is slower, requiring the X-ray crystal structures of the MHC molecules but can be applied to all MHC types, including those that are uncharacterized. Advanced approaches include matrix-driven methods, finding structural binding motifs, a quantitative structure activity relationship (QSAR) analysis, homology modeling, protein threading, docking techniques and design of several machine-learning algorithms. Structure based predications have the potential to discover non sequence based binders.

Both sequence and structure based computational approaches are based upon experimentally characterized peptides, but offer a more rapid indication of potential epitopes that could guide experimental studies. In both scenarios, experimental confirmation of peptide-MHC binding is still required.

In addition to sequence based or structure based predictions a number of computer algorithms have been developed that interrogate at the genome level for *in silico* prediction of T-cell antigens [264–267]. This has the potential to help in targeting low abundance T-cell epitopes in experimental studies. *In silico* methods, based upon various patterns in known MHC binding peptides, are cost effective and high throughput. They have the advantage of reducing the potential MHC binding peptide dataset, ruling out peptides that have no MHC binding potential. Even so, MHC binding is a prerequisite for T-cell activation, but does not guarantee it and experimental confirmation of T-cell activation is still required. There are also now epitope databases and web accessible tools for MHC binding prediction (for example http://www.iedb.org/). Other strategies have combined *in silico* prediction methods with mass spectrometry MHC peptide sequencing in order to increase the numbers of peptides identified [268–271]. This was exploited to target low abundance viral MHC peptides, synthesizing an *in silico* predicted MHC peptide as a calibrant, and using retention time and peptide mass–charge ratio in order to identify the corresponding experimental peptide [268, 269]. *In silico* prediction of MHC peptides is being demonstrated to be increasingly accurate when compared with experimental data [272, 273]. These approaches have the potential to increase the repertoire of detected MHC peptides. Moreover, sophisticated studies combining immunoproteomics and other approaches are beginning to decipher the origin and composition of the total repertoire of MHC peptides or "immunopeptidome" from a systems biology perspective [274].

The combined application of experimental studies and *in silico* based prediction will, in the long term impact upon vaccine development and personalized medicine. The information uncovers potential new antigens, which could be protein or peptide epitopes with the potential to stimulate protective immunity, i.e., to be part of a vaccine. The process is known as reverse vaccinology and has the potential to expedite the discovery and characterization of pathogen or disease epitopes. Reverse vaccinology identifies from whole genome sequences, antigenic extracellular proteins or peptides that are potential antigens. This approach has the potential to accelerate the sometimes slow and costly vaccine development pipeline. This was successfully pioneered for *Neisseria meningitidis*, causative agent of meningococcal meningitis and vaccines are now available for A, C, Y and W135 [275].

6 Emerging Technologies and Applications

In the previous sections, we have provided a high level summary of the current, most widely used techniques loosely grouped under "immunoproteomics". In the following subsections, we discuss emerging, or less widely used technologies that have potential to increase the breadth of immunoproteomics research.

6.1 Immuno-PCR

Immuno-PCR is a technique that was first reported in 1992 [276], and combines advantages of ELISA type assays, with the sensitivity of PCR and is aimed at detecting low abundance protein antigens. As outlined in Fig. 4, the antigen of interest is captured by a specific antibody and in a manner similar to traditional ELISA, a secondary antibody is used to detect binding. In this case, the secondary antibody is a chimeric antibody, with a DNA strand as the detection marker. The incorporation of a DNA tag allows amplification of the detection signal by PCR. This provides many of the advantages of PCR amplification, which are lacking in traditional ELISA assays. Immuno-PCR has been reported to have 10- to 1,000-fold increase in sensitivity compared to traditional antigen detection methods [276, 277], with high potential for the development of diagnostic assays. The technique has reported utility in detection of serological markers of ovarian cancer [278], CNS indicator proteins [279], detection and quantification of amyloid β-peptide in Alzheimer's disease [280], early diagnosis of infectious disease [281], cytokine detection [282], and toxin detection [283–287]. In addition, this method is not aimed at discovery immunoproteomics, and has been developed for speed and sensitivity for use as a clinical laboratory tool [288]. Development of real time quantitative immune-PCR has added the ability to measure the amounts of antigen in a sample [289–291]. An excellent review provides more details on this approach [292].

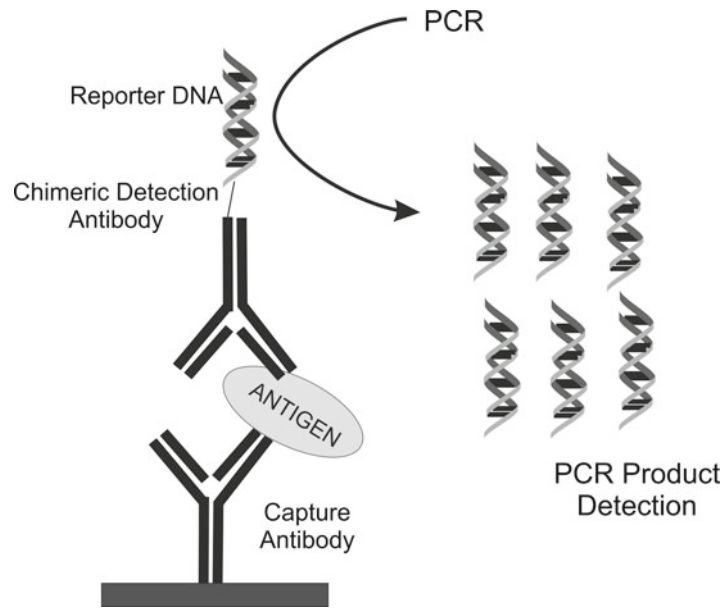


Fig. 4 Immuno-PCR. The setup of immune PCR is similar to that of traditional antigen detection ELISA. A capture antibody immobilizes the antigen, and detection antibody added. Instead of the antibody–enzyme conjugate used for colorimetric detection in ELISA, the chimeric antibody with reporter DNA is used. Addition of primers, nucleotides and polymerase allows amplification of the signal. The linear amplification of PCR means that the number of PCR amplicons generated is proportional to the initial amount of antigen detected. This shows a simplified scheme, and many variations have been developed

6.2 MALDI-TOF for Immune Cell Surface Discrimination

MALDI-TOF is seeing increasingly widespread use in clinical microbiology laboratories for the routine identification of bacterial species (for example ref. 293–296). The approach is based upon protein signatures (without protein identification), and exploits not only the differences in cell surface proteins between cells types, but the dynamic change in those proteins under certain conditions. This has been demonstrated to be a robust, reproducible, rapid and potentially cost saving approach in medical diagnostics [297]. Recently, this approach has been successful in discriminating intact immune cells, including lymphocytes, monocytes and polymorphonuclear cells for the generation of an immune cell database [298]. The same approach was also able distinguish between stimulated and unstimulated macrophages [298]. Further to this, distinct differences in the MALDI-TOF protein fingerprints of the surface of macrophages were detected with the addition of M1 agonists, IFN-γ, TNF, LPS, and LPS+IFN-γ, and the M2 agonists, IL-4, TGF-β1, and IL-10. The differences in macrophage surface fingerprints were specific and readily

identifiable [299]. The method is rapid and reproducible and opens the door to an alternative method of immune cell analysis.

6.3 In Vivo Microbial Antigen Discovery

In vivo microbial antigen discovery (InMAD) [300] was developed to identify circulating microbial antigens that are secreted or shed by bacteria, and detectable in sera. These circulating antigens can then be exploited for the development of rapid point of care immunoassays for bacterial diseases. The technique relies upon the humoral immune response to identify antigens that are circulating in sera. First carried out with the highly pathogenic bacteria, *Francisella tularensis* and *Burkholderia pseudomallei*, mice were infected with one or other organism and serum harvested [300]. The serum was filtered to remove whole bacteria, and termed InMAD serum. The filtered InMAD serum was then mixed with adjuvant and used to immunize mice. Bacterial proteins in the InMAD serum stimulate an immune response, which can then be monitored in order to determine the identity of the circulating bacterial proteins. Sera, collected from immunized mice was termed "InMAD immune serum" and was used in 2D Western blot or proteome array. In this way, the circulating bacterial proteins were identified [300] and have the potential to be rapidly translated into clinically relevant biomarkers for the disease diagnosis.

7 Applications

Immunoproteomics is still a relatively young field, with many academic reports, and a few being translated into clinical applications. However, there is huge potential for immunoproteomics-based assays to monitor or diagnose disease states or vaccine efficacy where antigens are involved. Bacterial and viral diseases are highly preventable through vaccination and an obvious application of immunoproteomics techniques is in antigen discovery for vaccine development. For example, efforts to develop a universal influenza vaccine with efficacy against all types of influenza need to be targeted against a conserved antibody or T-cell epitope. Mass spectrometry identification of influenza T-cell epitopes [301] is a step towards generating a vaccine that stimulates cross strain cell mediated immunity. A similar approach was used to identify conserved T-cell epitopes in dengue virus infected cells [302].

The remaining vaccine preventable diseases are challenging in terms of developing efficacious vaccines and discerning correlates of protection. Vaccinations against infectious disease are designed to stimulate a protective immune response. This immune response can be measured and correlated with the protection of the host against disease. In some cases, protective vaccination may only be established through detection of several immune parameters, such as immunodominant antibodies, cytokines etc. As immunoproteomics

studies advance in terms of sensitivity and throughput, this opens the door to rapid discovery of biomarkers of vaccine efficacy. Immunoproteomics approaches are being exploited to determine immune correlates of protection, which may then be used to monitor the protective status of the host. For example, proteome array studies have monitored the humoral immune response to smallpox and tularemia vaccines, and have noted a number of immunodominant proteins that have potential diagnostic applications [48, 54, 303–306]. These studies were extended further to investigate why smallpox vaccine fails to develop lesions in some individuals [307] and also comparing the antibody response to existing and next generation vaccinia virus vaccines [308].

Circulating antibodies represent important makers, reflecting the repertoire of nonself agents to which the immune system has been exposed. Antibodies amplify the signal of what may have been low abundance disease related proteins, have half lives of days to months and are stable to sample handling, so represent good biomarkers for diagnostic applications. As with all biomarker discoveries, validation and translation of immunoproteomic biomarkers to diagnostics is met with a number of challenges. Clinical diagnostic assays must be simple, robust, and sensitive, for example ELISA or antigen arrays.

Recombinant protein therapeutics are gaining popularity in a variety of applications. In addition to their desired therapeutic effects, they have the potential to stimulate an undesirable immune response against the recombinant protein. Protein therapeutics, such as recombinant IFNβ [309–311], IFNα [312, 313], and anti-TNFα antibodies [314, 315], are frequently observed to stimulate an undesirable immune response against the recombinant protein. The immune responses may be antibody or cell mediated and a combination of *in silico* prediction tools (reviewed in ref. 316) and *in vivo* validation by immunoproteomics methods could support prediction of immunogenicity for protein therapeutics, giving more rapid translation from discovery to clinic. Immunoproteomics approaches have the potential to have a high impact in this area, supporting the depletion of T-cell epitopes from protein therapeutics (reviewed in ref. 317).

8 Future Perspective

The breadth and sophistication of the techniques developed to study the immunoproteome have increased dramatically in the past decade. The field has benefited greatly from advances in proteomics and immunoinformatics and will continue to develop. Challenges remain, such as characterization of low abundance T-cell epitopes, and detection of low level serum cytokines. However, new avenues of investigation are emerging, including application of

interactomics to immunoproteomics studies, and comprehensive systems biology studies of the immune response to disease. As our depth of knowledge of the immune response to infection, cancer or self-antigens (misdirected autoimmunity) increases, so do the opportunities for discovery of robust disease biomarkers for early diagnosis. Combined *in silico* and experimental studies promise to yield efficacious vaccine candidates and correlates of vaccine protection. On a systems level, understanding the rapidly changing protein landscape of the immune system at various stages of life has the potential to provide immune markers of vaccine health, and predictive markers of the immune response, which may in the longer term, contribute to the development of personalized medicine.

References

1. Jungblut PR (2001) Proteome analysis of bacterial pathogens. Microbes Infect 3:831–840
2. McComb D, Thiriot A, Krishnan L, Stark F (2013) Introduction to the immune system. Immunoproteomics. Methods Mol Biol 1061:1–20
3. O'Farrell PH (1975) High resolution two-dimensional electrophoresis of proteins. J Biol Chem 250:4007–4021
4. Haas G, Karaali G, Ebermayer K, Metzger WG, Lamer S, Zimny-Arndt U, Diescher S, Goebel UB, Vogt K, Roznowski AB, Wiedenmann BJ, Meyer TF, Aebischer T, Jungblut PR (2002) Immunoproteomics of *Helicobacter pylori* infection and relation to gastric disease. Proteomics 2:313–324
5. Chen Z, Peng B, Wang S, Peng X (2004) Rapid screening of highly efficient vaccine candidates by immunoproteomics. Proteomics 4:3203–3213
6. Krah A, Miehlke S, Pleissner KP, Zimny-Arndt U, Kirsch C, Lehn N, Meyer TF, Jungblut PR, Aebischer T (2004) Identification of candidate antigens for serologic detection of *Helicobacter pylori*-infected patients with gastric carcinoma. Int J Cancer 108:456–463
7. Peng X, Ye X, Wang S (2004) Identification of novel immunogenic proteins of Shigella flexneri 2a by proteomic methodologies. Vaccine 22:2750–2756
8. Ying T, Wang H, Li M, Wang J, Shi Z, Feng E, Liu X, Su G, Wei K, Zhang X, Huang P, Huang L (2005) Immunoproteomics of outer membrane proteins and extracellular proteins of Shigella flexneri 2a 2457T. Proteomics 5:4777–4793
9. Ying TY, Wang JJ, Wang HL, Feng EL, Wei KH, Huang LY, Huang PT, Huang CF (2005) Immunoproteomics of membrane proteins of Shigella flexneri 2a 2457T. World J Gastroenterol 11:6880–6883
10. Connolly JP, Comerci D, Alefantis TG, Walz A, Quan M, Chafin R, Grewal P, Mujer CV, Ugalde RA, DelVecchio VG (2006) Proteomic analysis of Brucella abortus cell envelope and identification of immunogenic candidate proteins for vaccine development. Proteomics 6:3767–3780
11. Falisse-Poirrier N, Ruelle V, ElMoualij B, Zorzi D, Pierard O, Heinen E, De Pauw E, Zorzi W (2006) Advances in immunoproteomics for serological characterization of microbial antigens. J Microbiol Methods 67:593–596
12. Mini R, Bernardini G, Salzano AM, Renzone G, Scaloni A, Figura N, Santucci A (2006) Comparative proteomics and immunoproteomics of *Helicobacter pylori* related to different gastric pathologies. J Chromatogr B Analyt Technol Biomed Life Sci 833:63–79
13. Paul-Satyaseela M, Karched M, Bian Z, Ihalin R, Boren T, Arnqvist A, Chen C, Asikainen S (2006) Immunoproteomics of *Actinobacillus actinomycetemcomitans* outer-membrane proteins reveal a highly immunoreactive peptidoglycan-associated lipoprotein. J Med Microbiol 55:931–942
14. Bernardini G, Braconi D, Lusini P, Santucci A (2007) *Helicobacter pylori*, immunoproteomics related to different pathologies. Expert Rev Proteomics 4:679–689
15. Hagan EC, Mobley HL (2007) Uropathogenic *Escherichia coli* outer membrane antigens expressed during urinary tract infection. Infect Immun 75:3941–3949
16. Janovska S, Pavkova I, Reichelova M, Hubaleka M, Stulik J, Macela A (2007) Proteomic analysis of antibody response in a

case of laboratory-acquired infection with *Francisella tularensis* subsp. *tularensis*. Folia Microbiol (Praha) 52:194–198
17. Shin GW, Palaksha KJ, Kim YR, Nho SW, Kim S, Heo GJ, Park SC, Jung TS (2007) Application of immunoproteomics in developing a *Streptococcus iniae* vaccine for olive flounder (*Paralichthys olivaceus*). J Chromatogr B Analyt Technol Biomed Life Sci 849:315–322
18. Zhang W, Lu CP (2007) Immunoproteomics of extracellular proteins of Chinese virulent strains of *Streptococcus suis* type 2. Proteomics 7:4468–4476
19. Li H, Xiong XP, Peng B, Xu CX, Ye MZ, Yang TC, Wang SY, Peng XX (2009) Identification of broad cross-protective immunogens using heterogeneous antiserum-based immunoproteomic approach. J Proteome Res 8:4342–4349
20. Park SH, Kwon SJ, Lee SJ, Kim YC, Hwang KY, Kang YH, Lee KJ (2009) Identification of immunogenic antigen candidate for *Chlamydophila pneumoniae* diagnosis. J Proteome Res 8:2933–2943
21. Lahner E, Bernardini G, Santucci A, Annibale B (2010) *Helicobacter pylori* immunoproteomics in gastric cancer and gastritis of the carcinoma phenotype. Expert Rev Proteomics 7:239–248
22. Li H, Ye MZ, Peng B, Wu HK, Xu CX, Xiong XP, Wang C, Wang SY, Peng XX (2010) Immunoproteomic identification of polyvalent vaccine candidates from *Vibrio parahaemolyticus* outer membrane proteins. J Proteome Res 9:2573–2583
23. Roy K, Bartels S, Qadri F, Fleckenstein JM (2010) Enterotoxigenic *Escherichia coli* elicits immune responses to multiple surface proteins. Infect Immun 78:3027–3035
24. Santi L, Silva WO, Pinto AF, Schrank A, Vainstein MH (2010) Metarhizium anisopliae host-pathogen interaction, differential immunoproteomics reveals proteins involved in the infection process of arthropods. Fungal Biol 114:312–319
25. Lahner E, Bernardini G, Possenti S, Renzone G, Scaloni A, Santucci A, Annibale B (2011) Immunoproteomics of *Helicobacter pylori* infection in patients with atrophic body gastritis, a predisposing condition for gastric cancer. Int J Med Microbiol 301:125–132
26. Wang Y, Dang Y, Wang X, Lu H, Lang X, Li X, Feng S, Zhang F, Ren L (2011) Comparative proteomic analyses of *Streptococcus suis* serotype 2 cell wall-associated proteins. Curr Microbiol 62:578–588
27. Yang Y, Wang L, Yin J, Wang X, Cheng S, Lang X, Qu H, Sun C, Wang J, Zhang R (2011) Immunoproteomic analysis of *Brucella melitensis* and identification of a new immunogenic candidate protein for the development of brucellosis subunit vaccine. Mol Immunol 49:175–184
28. Hu Q, Ding C, Tu J, Wang X, Han X, Duan Y, Yu S (2012) Immunoproteomics analysis of whole cell bacterial proteins of *Riemerella anatipestifer*. Vet Microbiol 157:428–438
29. Pajuaba AC, Silva DA, Almeida KC, Cunha-Junior JP, Pirovani CP, Camillo LR, Mineo JR (2012) Immunoproteomics of *Brucella abortus* reveals differential antibody profiles between S19-vaccinated and naturally infected cattle. Proteomics 12:820–831
30. Cha HJ, Yoon HG, Kim YW, Lee HS, Kim JW, Kweon KS, Oh BH, Park KH (1998) Molecular and enzymatic characterization of a maltogenic amylase that hydrolyzes and transglycosylates acarbose. Eur J Biochem 253:251–262
31. Hamrita B, Chahed K, Kabbage M, Guillier CL, Trimeche M, Chaieb A, Chouchane L (2008) Identification of tumor antigens that elicit a humoral immune response in breast cancer patients' sera by serological proteome analysis (SERPA). Clin Chim Acta 393: 95–102
32. Li L, Chen SH, Yu CH, Li YM, Wang SQ (2008) Identification of hepatocellular-carcinoma-associated antigens and autoantibodies by serological proteome analysis combined with protein microarray. J Proteome Res 7:611–620
33. Forgber M, Gellrich S, Sharav T, Sterry W, Walden P (2009) Proteome-based analysis of serologically defined tumor-associated antigens in cutaneous lymphoma. PLoS One 4:e8376
34. Forgber M, Trefzer U, Sterry W, Walden P (2009) Proteome serological determination of tumor-associated antigens in melanoma. PLoS One 4:e5199
35. Mou Z, He Y, Wu Y (2009) Immunoproteomics to identify tumor-associated antigens eliciting humoral response. Cancer Lett 278:123–129
36. Shukla S, Pranay A, D'Cruz AK, Chaturvedi P, Kane SV, Zingde SM (2009) Immunoproteomics reveals that cancer of the tongue and the gingivobuccal complex exhibit differential autoantibody response. Cancer Biomark 5:127–135
37. Liu R, Wang K, Yuan K, Wei Y, Huang C (2010) Integrative oncoproteomics strategies for anticancer drug discovery. Expert Rev Proteomics 7:411–429
38. Suzuki A, Iizuka A, Komiyama M, Takikawa M, Kume A, Tai S, Ohshita C, Kurusu A,

Nakamura Y, Yamamoto A, Yamazaki N, Yoshikawa S, Kiyohara Y, Akiyama Y (2010) Identification of melanoma antigens using a Serological Proteome Approach (SERPA). Cancer Genomics Proteomics 7:17–23
39. Mojtahedi Z, Safaei A, Yousefi Z, Ghaderi A (2011) Immunoproteomics of HER2-positive and HER2-negative breast cancer patients with positive lymph nodes. OMICS 15:409–418
40. O'Meara MM, Disis ML (2011) Therapeutic cancer vaccines and translating vaccinomics science to the global health clinic, emerging applications toward proof of concept. OMICS 15:579–588
41. Almeras L, Lefranc D, Drobecq H, de Seze J, Dubucquoi S, Vermersch P, Prin L (2004) New antigenic candidates in multiple sclerosis, identification by serological proteome analysis. Proteomics 4:2184–2194
42. Xiong X, Wang X, Wen B, Graves S, Stenos J (2012) Potential serodiagnostic markers for Q fever identified in *Coxiella burnetii* by immunoproteomic and protein microarray approaches. BMC Microbiol 12:35
43. Kornilovs'ka I, Nilsson I, Utt M, Ljungh A, Wadstrom T (2002) Immunogenic proteins of *Helicobacter pullorum*, *Helicobacter bilis* and *Helicobacter hepaticus* identified by two-dimensional gel electrophoresis and immunoblotting. Proteomics 2:775–783
44. Utt M, Nilsson I, Ljungh A, Wadstrom T (2002) Identification of novel immunogenic proteins of *Helicobacter pylori* by proteome technology. J Immunol Methods 259:1–10
45. Zhong ZR, Zhou HB, Li XY, Luo QL, Song XR, Wang W, Wen HQ, Yu L, Wei W, Shen JL (2010) Serological proteome-oriented screening and application of antigens for the diagnosis of *Schistosomiasis japonica*. Acta Trop 116:1–8
46. Chitlaru T, Gat O, Grosfeld H, Inbar I, Gozlan Y, Shafferman A (2007) Identification of in vivo-expressed immunogenic proteins by serological proteome analysis of the *Bacillus anthracis* secretome. Infect Immun 75:2841–2852
47. Zhu YZ, Cai CS, Zhang W, Guo HX, Zhang JP, Ji YY, Ma GY, Wu JL, Li QT, Lu CP, Guo XK (2010) Immunoproteomic analysis of human serological antibody responses to vaccination with whole-cell pertussis vaccine (WCV). PLoS One 5:e13915
48. Fulton KM, Zhao X, Petit MD, Kilmury SL, Wolfraim LA, House RV, Sjostedt A, Twine SM (2011) Immunoproteomic analysis of the human antibody response to natural tularemia infection with Type A or Type B strains or LVS vaccination. Int J Med Microbiol 301:591–601
49. Davies DH, Liang X, Hernandez JE, Randall A, Hirst S, Mu Y, Romero KM, Nguyen TT, Kalantari-Dehaghi M, Crotty S, Baldi P, Villarreal LP, Felgner PL (2005) Profiling the humoral immune response to infection by using proteome microarrays, high-throughput vaccine and diagnostic antigen discovery. Proc Natl Acad Sci U S A 102:547–552
50. Eyles JE, Unal B, Hartley MG, Newstead SL, Flick-Smith H, Prior JL, Oyston PC, Randall A, Mu Y, Hirst S, Molina DM, Davies DH, Milne T, Griffin KF, Baldi P, Titball RW, Felgner PL (2007) Immunodominant *Francisella tularensis* antigens identified using proteome microarray. Proteomics 7: 2172–2183
51. Sundaresh S, Doolan DL, Hirst S, Mu Y, Unal B, Davies DH, Felgner PL, Baldi P (2006) Identification of humoral immune responses in protein microarrays using DNA microarray data analysis techniques. Bioinformatics 22:1760–1766
52. Fang Y, Frutos AG, Lahiri J (2002) Membrane protein microarrays. J Am Chem Soc 124:2394–2395
53. Shin I, Cho JW, Boo DW (2004) Carbohydrate arrays for functional studies of carbohydrates. Comb Chem High Throughput Screen 7:565–574
54. Davies DH, Molina DM, Wrammert J, Miller J, Hirst S, Mu Y, Pablo J, Unal B, Nakajima-Sasaki R, Liang X, Crotty S, Karem KL, Damon IK, Ahmed R, Villarreal L, Felgner PL (2007) Proteome-wide analysis of the serological response to vaccinia and smallpox. Proteomics 7:1678–1686
55. Benhnia MR, Maybeno M, Blum D, Aguilar-Sino R, Matho M, Meng X, Head S, Felgner PL, Zajonc DM, Koriazova L, Kato S, Burton DR, Xiang Y, Crowe JE Jr, Peters B, Crotty S (2013) Unusual features of vaccinia virus extracellular virion form (EV) neutralization resistance revealed in human antibody responses to the smallpox vaccine. J Virol 87:1569–1585
56. Cruz-Fisher MI, Cheng C, Sun G, Pal S, Teng A, Molina DM, Kayala MA, Vigil A, Baldi P, Felgner PL, Liang X, de la Maza LM (2011) Identification of immunodominant antigens by probing a whole Chlamydia trachomatis open reading frame proteome microarray using sera from immunized mice. Infect Immun 79:246–257
57. Teng A, Cruz-Fisher MI, Cheng C, Pal S, Sun G, Ralli-Jain P, Molina DM, Felgner PL, Liang X, de la Maza LM (2012) Proteomic identification of immunodominant chlamydial antigens in a mouse model. J Proteomics 77:176–186
58. Conejero L, Patel N, de Reynal M, Oberdorf S, Prior J, Felgner PL, Titball RW, Salguero

FJ, Bancroft GJ (2011) Low-dose exposure of C57BL/6 mice to *burkholderia pseudomallei* mimics chronic human melioidosis. Am J Pathol 179:270–280
59. Cannella AP, Tsolis RM, Liang L, Felgner PL, Saito M, Sette A, Gotuzzo E, Vinetz JM (2012) Antigen-specific acquired immunity in human brucellosis, implications for diagnosis, prognosis, and vaccine development. Front Cell Infect Microbiol 2:1
60. Kunnath-Velayudhan S, Salamon H, Wang HY, Davidow AL, Molina DM, Huynh VT, Cirillo DM, Michel G, Talbot EA, Perkins MD, Felgner PL, Liang X, Gennaro ML (2010) Dynamic antibody responses to the *Mycobacterium tuberculosis* proteome. Proc Natl Acad Sci U S A 107:14703–14708
61. Kunnath-Velayudhan S, Davidow AL, Wang HY, Molina DM, Huynh VT, Salamon H, Pine R, Michel G, Perkins MD, Xiaowu L, Felgner PL, Flynn JL, Catanzaro A, Gennaro ML (2012) Proteome-scale antibody responses and outcome of *Mycobacterium tuberculosis* infection in nonhuman primates and in tuberculosis patients. J Infect Dis 206:697–705
62. Lee SJ, Liang L, Juarez S, Nanton MR, Gondwe EN, Msefula CL, Kayala MA, Necchi F, Heath JN, Hart P, Tsolis RM, Heyderman RS, MacLennan CA, Felgner PL, Davies DH, McSorley SJ (2012) Identification of a common immune signature in murine and human systemic Salmonellosis. Proc Natl Acad Sci U S A 109:4998–5003
63. Dasgupta G, Chentoufi AA, Kalantari M, Falatoonzadeh P, Chun S, Lim CH, Felgner PL, Davies DH, BenMohamed L (2012) Immunodominant "asymptomatic" herpes simplex virus 1 and 2 protein antigens identified by probing whole-ORFome microarrays with serum antibodies from seropositive asymptomatic versus symptomatic individuals. J Virol 86:4358–4369
64. Kalantari-Dehaghi M, Chun S, Chentoufi AA, Pablo J, Liang L, Dasgupta G, Molina DM, Jasinskas A, Nakajima-Sasaki R, Felgner J, Hermanson G, BenMohamed L, Felgner PL, Davies DH (2012) Discovery of potential diagnostic and vaccine antigens in herpes simplex virus 1 and 2 by proteome-wide antibody profiling. J Virol 86:4328–4339
65. Doolan DL, Mu Y, Unal B, Sundaresh S, Hirst S, Valdez C, Randall A, Molina D, Liang X, Freilich DA, Oloo JA, Blair PL, Aguiar JC, Baldi P, Davies DH, Felgner PL (2008) Profiling humoral immune responses to P. falciparum infection with protein microarrays. Proteomics 8:4680–4694
66. Barry AE, Trieu A, Fowkes FJ, Pablo J, Kalantari-Dehaghi M, Jasinskas A, Tan X, Kayala MA, Tavul L, Siba PM, Day KP, Baldi P, Felgner PL, Doolan DL (2011) The stability and complexity of antibody responses to the major surface antigen of *Plasmodium falciparum* are associated with age in a malaria endemic area. Mol Cell Proteomics 10:M111 008326
67. Nnedu ON, O'Leary MP, Mutua D, Mutai B, Kalantari-Dehaghi M, Jasinskas A, Nakajima-Sasaki R, John-Stewart G, Otieno P, Liang X, Waitumbi J, Kimani F, Camerini D, Felgner PL, Walson JL, Vigil A (2011) Humoral immune responses to *Plasmodium falciparum* among HIV-1-infected Kenyan adults. Proteomics Clin Appl 5:613–623
68. Fan YT, Wang Y, Ju C, Zhang T, Xu B, Hu W, Chen JH (2013) Systematic analysis of natural antibody responses to *P. falciparum* merozoite antigens by protein arrays. J Proteomics 78:148–158
69. Beare PA, Chen C, Bouman T, Pablo J, Unal B, Cockrell DC, Brown WC, Barbian KD, Porcella SF, Samuel JE, Felgner PL, Heinzen RA (2008) Candidate antigens for Q fever serodiagnosis revealed by immunoscreening of a Coxiella burnetii protein microarray. Clin Vaccine Immunol 15:1771–1779
70. Vigil A, Chen C, Jain A, Nakajima-Sasaki R, Jasinskas A, Pablo J, Hendrix LR, Samuel JE, Felgner PL (2011) Profiling the humoral immune response of acute and chronic Q fever by protein microarray. Mol Cell Proteomics 10:M110 006304
71. Liang L, Doskaya M, Juarez S, Caner A, Jasinskas A, Tan X, Hajagos BE, Bradley PJ, Korkmaz M, Guruz Y, Felgner PL, Davies DH (2011) Identification of potential serodiagnostic and subunit vaccine antigens by antibody profiling of toxoplasmosis cases in Turkey. Mol Cell Proteomics 10:M110 006916
72. Suwannasaen D, Mahawantung J, Chaowagul W, Limmathurotsakul D, Felgner PL, Davies H, Bancroft GJ, Titball RW, Lertmemongkolchai G (2011) Human immune responses to Burkholderia pseudomallei characterized by protein microarray analysis. J Infect Dis 203:1002–1011
73. Barbour AG, Jasinskas A, Kayala MA, Davies DH, Steere AC, Baldi P, Felgner PL (2008) A genome-wide proteome array reveals a limited set of immunogens in natural infections of humans and white-footed mice with Borrelia burgdorferi. Infect Immun 76:3374–3389
74. Sundaresh S, Randall A, Unal B, Petersen JM, Belisle JT, Hartley MG, Duffield M, Titball RW, Davies DH, Felgner PL, Baldi P (2007) From protein microarrays to diagnostic antigen discovery, a study of the pathogen *Francisella tularensis*. Bioinformatics 23: i508–i518

75. Zheng D, Wan J, Cho YG, Wang L, Chiou CJ, Pai S, Woodard C, Zhu J, Liao G, Martinez-Maza O, Qian J, Zhu H, Hayward GS, Ambinder RF, Hayward SD (2011) Comparison of humoral immune responses to Epstein-Barr virus and Kaposi's sarcoma-associated herpesvirus using a viral proteome microarray. J Infect Dis 204:1683–1691
76. Hakomori S (2002) Glycosylation defining cancer malignancy, new wine in an old bottle. Proc Natl Acad Sci U S A 99:10231–10233
77. Coenen D, Verschueren P, Westhovens R, Bossuyt X (2007) Technical and diagnostic performance of 6 assays for the measurement of citrullinated protein/peptide antibodies in the diagnosis of rheumatoid arthritis. Clin Chem 53:498–504
78. Hiki Y (2009) O-linked oligosaccharides of the IgA1 hinge region, roles of its aberrant structure in the occurrence and/or progression of IgA nephropathy. Clin Exp Nephrol 13:415–423
79. Schietinger A, Philip M, Yoshida BA, Azadi P, Liu H, Meredith SC, Schreiber H (2006) A mutant chaperone converts a wild-type protein into a tumor-specific antigen. Science 314:304–308
80. Wandall HH, Blixt O, Tarp MA, Pedersen JW, Bennett EP, Mandel U, Ragupathi G, Livingston PO, Hollingsworth MA, Taylor-Papadimitriou J, Burchell J, Clausen H (2010) Cancer biomarkers defined by autoantibody signatures to aberrant O-glycopeptide epitopes. Cancer Res 70:1306–1313
81. Smorodin EP, Kurtenkov OA, Shevchuk IN, Tanner RH (2005) The isolation and characterization of human natural alphaGal-specific IgG antibodies applicable to the detection of alphaGal-glycosphingolipids. J Immunoassay Immunochem 26:145–156
82. Reis CA, David L, Seixas M, Burchell J, Sobrinho-Simoes M (1998) Expression of fully and under-glycosylated forms of MUC1 mucin in gastric carcinoma. Int J Cancer 79:402–410
83. Brockhausen I (1999) Pathways of O-glycan biosynthesis in cancer cells. Biochim Biophys Acta 1473:67–95
84. Blixt O, Clo E, Nudelman AS, Sorensen KK, Clausen T, Wandall HH, Livingston PO, Clausen H, Jensen KJ (2010) A high-throughput O-glycopeptide discovery platform for seromic profiling. J Proteome Res 9:5250–5261
85. Kracun SK, Clo E, Clausen H, Levery SB, Jensen KJ, Blixt O (2010) Random glycopeptide bead libraries for seromic biomarker discovery. J Proteome Res 9:6705–6714
86. Pedersen JW, Blixt O, Bennett EP, Tarp MA, Dar I, Mandel U, Poulsen SS, Pedersen AE, Rasmussen S, Jess P, Clausen H, Wandall HH (2011) Seromic profiling of colorectal cancer patients with novel glycopeptide microarray. Int J Cancer 128:1860–1871
87. Reis CA, Campos D, Osorio H, Santos LL (2011) Glycopeptide microarray for autoantibody detection in cancer. Expert Rev Proteomics 8:435–437
88. Chow JC, Young DW, Golenbock DT, Christ WJ, Gusovsky F (1999) Toll-like receptor-4 mediates lipopolysaccharide-induced signal transduction. J Biol Chem 274:10689–10692
89. Poltorak A, Ricciardi-Castagnoli P, Citterio S, Beutler B (2000) Physical contact between lipopolysaccharide and toll-like receptor 4 revealed by genetic complementation. Proc Natl Acad Sci U S A 97:2163–2167
90. Girardin SE, Boneca IG, Carneiro LA, Antignac A, Jehanno M, Viala J, Tedin K, Taha MK, Labigne A, Zahringer U, Coyle AJ, DiStefano PS, Bertin J, Sansonetti PJ, Philpott DJ (2003) Nod1 detects a unique muropeptide from gram-negative bacterial peptidoglycan. Science 300:1584–1587
91. Girardin SE, Boneca IG, Viala J, Chamaillard M, Labigne A, Thomas G, Philpott DJ, Sansonetti PJ (2003) Nod2 is a general sensor of peptidoglycan through muramyl dipeptide (MDP) detection. J Biol Chem 278: 8869–8872
92. Girardin SE, Hugot JP, Sansonetti PJ (2003) Lessons from Nod2 studies, towards a link between Crohn's disease and bacterial sensing. Trends Immunol 24:652–658
93. Girardin SE, Travassos LH, Herve M, Blanot D, Boneca IG, Philpott DJ, Sansonetti PJ, Mengin-Lecreulx D (2003) Peptidoglycan molecular requirements allowing detection by Nod1 and Nod2. J Biol Chem 278: 41702–41708
94. Verma A, Arora SK, Kuravi SK, Ramphal R (2005) Roles of specific amino acids in the N terminus of *Pseudomonas aeruginosa* flagellin and of flagellin glycosylation in the innate immune response. Infect Immun 73: 8237–8246
95. Doig P, Kinsella N, Guerry P, Trust TJ (1996) Characterization of a post-translational modification of Campylobacter flagellin, identification of a sero-specific glycosyl moiety. Mol Microbiol 19:379–387
96. Horn C, Namane A, Pescher P, Riviere M, Romain F, Puzo G, Barzu O, Marchal G (1999) Decreased capacity of recombinant 45/47-kDa molecules (Apa) of *Mycobacterium tuberculosis* to stimulate T lymphocyte responses related to changes in their mannosylation pattern. J Biol Chem 274:32023–32030
97. Romain F, Horn C, Pescher P, Namane A, Riviere M, Puzo G, Barzu O, Marchal G (1999) Deglycosylation of the 45/47-kilodalton antigen complex of

Mycobacterium tuberculosis decreases its capacity to elicit in vivo or in vitro cellular immune responses. Infect Immun 67:5567–5572
98. Logan SM, Kelly JF, Thibault P, Ewing CP, Guerry P (2002) Structural heterogeneity of carbohydrate modifications affects serospecificity of Campylobacter flagellins. Mol Microbiol 46:587–597
99. Mehta AS, Saile E, Zhong W, Buskas T, Carlson R, Kannenberg E, Reed Y, Quinn CP, Boons GJ (2006) Synthesis and antigenic analysis of the BclA glycoprotein oligosaccharide from the *Bacillus anthracis* exosporium. Chemistry 12:9136–9149
100. Wang D, Carroll GT, Turro NJ, Koberstein JT, Kovac P, Saksena R, Adamo R, Herzenberg LA, Steinman L (2007) Photogenerated glycan arrays identify immunogenic sugar moieties of *Bacillus anthracis* exosporium. Proteomics 7:180–184
101. Dhenin SG, Moreau V, Morel N, Nevers MC, Volland H, Creminon C, Djedaini-Pilard F (2008) Synthesis of an anthrose derivative and production of polyclonal antibodies for the detection of anthrax spores. Carbohydr Res 343:2101–2110
102. Dhenin SG, Moreau V, Nevers MC, Creminon C, Djedaini-Pilard F (2009) Sensitive and specific enzyme immunoassays for antigenic trisaccharide from *Bacillus anthracis* spores. Org Biomol Chem 7:5184–5199
103. Lee SY, Jeoung D (2007) The reverse proteomics for identification of tumor antigens. J Microbiol Biotechnol 17:879–890
104. Sahin U, Tureci O, Pfreundschuh M (1997) Serological identification of human tumor antigens. Curr Opin Immunol 9:709–716
105. Scanlan MJ, Gordan JD, Williamson B, Stockert E, Bander NH, Jongeneel V, Gure AO, Jager D, Jager E, Knuth A, Chen YT, Old LJ (1999) Antigens recognized by autologous antibody in patients with renal-cell carcinoma. Int J Cancer 83:456–464
106. Scanlan MJ, Gout I, Gordon CM, Williamson B, Stockert E, Gure AO, Jager D, Chen YT, Mackay A, O'Hare MJ, Old LJ (2001) Humoral immunity to human breast cancer, antigen definition and quantitative analysis of mRNA expression. Cancer Immun 1:4
107. Koroleva EP, Lagarkova MA, Mesheryakov AA, Scanlan MJ, Old LJ, Nedospasov SA, Kuprash DV (2002) Serological identification of antigens associated with renal cell carcinoma. Russ J Immunol 7:229–238
108. Devitt G, Meyer C, Wiedemann N, Eichmuller S, Kopp-Schneider A, Haferkamp A, Hautmann R, Zoller M (2006) Serological analysis of human renal cell carcinoma. Int J Cancer 118:2210–2219
109. Scanlan MJ, Chen YT, Williamson B, Gure AO, Stockert E, Gordan JD, Tureci O, Sahin U, Pfreundschuh M, Old LJ (1998) Characterization of human colon cancer antigens recognized by autologous antibodies. Int J Cancer 76:652–658
110. Line A, Slucka Z, Stengrevics A, Silina K, Li G, Rees RC (2002) Characterisation of tumour-associated antigens in colon cancer. Cancer Immunol Immunother 51:574–582
111. Ishikawa T, Fujita T, Suzuki Y, Okabe S, Yuasa Y, Iwai T, Kawakami Y (2003) Tumor-specific immunological recognition of frameshift-mutated peptides in colon cancer with microsatellite instability. Cancer Res 63:5564–5572
112. Obata Y, Akahashi T, Tamaki H, Tominaga S, Murai H, Iwase T, Iwata H, Mizutani M, Chen YT, Old LJ, Miura S (1999) Identification of cancer antigens in breast cancer by the SEREX expression cloning method. Breast Cancer 6:305–311
113. Forti S, Scanlan MJ, Invernizzi A, Castiglioni F, Pupa S, Agresti R, Fontanelli R, Morelli D, Old LJ, Pupa SM, Menard S (2002) Identification of breast cancer-restricted antigens by antibody screening of SKBR3 cDNA library using a preselected patient's serum. Breast Cancer Res Treat 73:245–256
114. Jager D, Unkelbach M, Frei C, Bert F, Scanlan MJ, Jager E, Old LJ, Chen YT, Knuth A (2002) Identification of tumor-restricted antigens NY-BR-1, SCP-1, and a new cancer/testis-like antigen NW-BR-3 by serological screening of a testicular library with breast cancer serum. Cancer Immun 2:5
115. Minenkova O, Pucci A, Pavoni E, De Tomassi A, Fortugno P, Gargano N, Cianfriglia M, Barca S, De Placido S, Martignetti A, Felici F, Cortese R, Monaci P (2003) Identification of tumor-associated antigens by screening phage-displayed human cDNA libraries with sera from tumor patients. Int J Cancer 106:534–544
116. Jager D, Taverna C, Zippelius A, Knuth A (2004) Identification of tumor antigens as potential target antigens for immunotherapy by serological expression cloning. Cancer Immunol Immunother 53:144–147
117. Qian F, Odunsi K, Blatt LM, Scanlan MJ, Mannan M, Shah N, Montgomery J, Haddad F, Taylor M (2005) Tumor associated antigen recognition by autologous serum in patients with breast cancer. Int J Mol Med 15: 137–144
118. Shimada H, Nakashima K, Ochiai T, Nabeya Y, Takiguchi M, Nomura F, Hiwasa T (2005) Serological identification of tumor antigens of esophageal squamous cell carcinoma. Int J Oncol 26:77–86

119. Jager D (2007) Potential target antigens for immunotherapy identified by serological expression cloning (SEREX). Methods Mol Biol 360:319–326
120. Kiyamova R, Kostianets O, Malyuchik S, Filonenko V, Usenko V, Gurtovyy V, Khozayenko Y, Antonuk S, Old L, Gout I (2010) Identification of tumor-associated antigens from medullary breast carcinoma by a modified SEREX approach. Mol Biotechnol 46:105–112
121. Song MH, Ha JC, Lee SM, Park YM, Lee SY (2011) Identification of BCP-20 (FBXO39) as a cancer/testis antigen from colon cancer patients by SEREX. Biochem Biophys Res Commun 408:195–201
122. Kostianets O, Shyian M, Sergiy D, Antoniuk S, Gout I, Filonenko V, Kiyamova R (2012) Serological analysis of SEREX-defined medullary breast carcinoma-associated antigens. Cancer Invest 30:519–527
123. Song MH, Choi KU, Shin DH, Lee CH, Lee SY (2012) Identification of the cancer/testis antigens AKAP3 and CTp11 by SEREX in hepatocellular carcinoma. Oncol Rep 28:1792–1798
124. Chen YT, Scanlan MJ, Sahin U, Tureci O, Gure AO, Tsang S, Williamson B, Stockert E, Pfreundschuh M, Old LJ (1997) A testicular antigen aberrantly expressed in human cancers detected by autologous antibody screening. Proc Natl Acad Sci U S A 94: 1914–1918
125. Jager E, Jager D, Knuth A (1999) CTL-defined cancer vaccines, perspectives for active immunotherapeutic interventions in minimal residual disease. Cancer Metastasis Rev 18:143–150
126. Old LJ (2008) Cancer vaccines, an overview. Cancer Immun 8(Suppl 1):1
127. Boder ET, Wittrup KD (1997) Yeast surface display for screening combinatorial polypeptide libraries. Nat Biotechnol 15:553–557
128. Mischo A, Wadle A, Watzig K, Jager D, Stockert E, Santiago D, Ritter G, Regitz E, Jager E, Knuth A, Old L, Pfreundschuh M, Renner C (2003) Recombinant antigen expression on yeast surface (RAYS) for the detection of serological immune responses in cancer patients. Cancer Immun 3:5
129. Wadle A, Mischo A, Imig J, Wullner B, Hensel D, Watzig K, Neumann F, Kubuschok B, Schmidt W, Old LJ, Pfreundschuh M, Renner C (2005) Serological identification of breast cancer-related antigens from a *Saccharomyces cerevisiae* surface display library. Int J Cancer 117:104–113
130. Kim MS, Choi HY, Choi YS, Kim J, Kim YS (2007) Optimized serological isolation of lung-cancer-associated antigens from a yeast surface-expressed cDNA library. J Microbiol Biotechnol 17:993–1001
131. Jongeneel V (2001) Towards a cancer immunome database. Cancer Immun 1:3
132. Caron M, Choquet-Kastylevsky G, Joubert-Caron R (2007) Cancer immunomics using autoantibody signatures for biomarker discovery. Mol Cell Proteomics 6:1115–1122
133. Hardouin J, Lasserre JP, Canelle L, Duchateau M, Vlieghe C, Choquet-Kastylevsky G, Joubert-Caron R, Caron M (2007) Usefulness of autoantigens depletion to detect autoantibody signatures by multiple affinity protein profiling. J Sep Sci 30:352–358
134. Hardouin J, Lasserre JP, Sylvius L, Joubert-Caron R, Caron M (2007) Cancer immunomics, from serological proteome analysis to multiple affinity protein profiling. Ann N Y Acad Sci 1107:223–230
135. Roozendaal R, Carroll MC (2007) Complement receptors CD21 and CD35 in humoral immunity. Immunol Rev 219: 157–166
136. Solomon S, Kassahn D, Illges H (2005) The role of the complement and the Fc gamma R system in the pathogenesis of arthritis. Arthritis Res Ther 7:129–135
137. Croce MV, Fejes M, Riera N, Minoldo DA, Segal-Eiras A (1985) Clinical importance of circulating immune complexes in human acute lymphoblastic leukemia. Cancer Immunol Immunother 20:91–95
138. Liu P, Overman RG, Yates NL, Alam SM, Vandergrift N, Chen Y, Graw F, Freel SA, Kappes JC, Ochsenbauer C, Montefiori DC, Gao F, Perelson AS, Cohen MS, Haynes BF, Tomaras GD (2011) Dynamic antibody specificities and virion concentrations in circulating immune complexes in acute to chronic HIV-1 infection. J Virol 85:11196–11207
139. Coppo R, Bosticardo GM, Basolo B, Messina M, Mazzucco G, Stratta P, Quarello F, Alloatti S, Piccoli G (1982) Clinical significance of the detection of circulating immune complexes in lupus nephritis. Nephron 32: 320–328
140. Soltis RD, Hasz DE (1983) The effect of serum immunoglobulin concentration on immune complex detection by polyethylene glycol. J Immunol Methods 57:275–282
141. Valentijn RM, van Overhagen H, Hazevoet HM, Hermans J, Cats A, Daha MR, van EL (1985) The value of complement and immune complex determinations in monitoring disease activity in patients with systemic lupus erythematosus. Arthritis Rheum 28:904–913
142. Lock RJ, Unsworth DJ (2000) Measurement of immune complexes is not useful in routine clinical practice. Ann Clin Biochem 37(Pt 3): 253–261

143. Ohyama K, Ueki Y, Kawakami A, Kishikawa N, Tamai M, Osaki M, Kamihira S, Nakashima K, Kuroda N (2011) Immune complexome analysis of serum and its application in screening for immune complex antigens in rheumatoid arthritis. Clin Chem 57:905–909
144. Ohyama K, Kuroda N (2012) Proteomic approaches to profiling the humoral immune response and identifying disease-associated antigens. Biol Pharm Bull 35:1409–1412
145. Ohyama K, Kawakami A, Tamai M, Baba M, Kishikawa N, Kuroda N (2012) Serum immune complex containing thrombospondin-1, a novel biomarker for early rheumatoid arthritis. Ann Rheum Dis 71:1916–1917
146. Beyer NH, Schou C, Houen G, Heegaard NH (2008) Extraction and identification of electroimmunoprecipitated proteins from agarose gels. J Immunol Methods 330: 24–33
147. Grubb AO (1974) Crossed immunoelectrophoresis and electroimmunoassay of IgG. J Immunol 113:343–347
148. Grubb AO (1974) Crossed immunoelectrophoresis and electroimmunoassay of IgM. J Immunol 112:1420–1425
149. Laurell CB (1966) Quantitative estimation of proteins by electrophoresis in agarose gel containing antibodies. Anal Biochem 15:45–52
150. Gershoni JM, Roitburd-Berman A, Siman-Tov DD, Tarnovitski Freund N, Weiss Y (2007) Epitope mapping, the first step in developing epitope-based vaccines. BioDrugs 21:145–156
151. Carter JM, Loomis-Price L (2004) B cell epitope mapping using synthetic peptides. Curr Protoc Immunol Chapter 9, Unit 9 4
152. Ladner RC (2007) Mapping the epitopes of antibodies. Biotechnol Genet Eng Rev 24:1–30
153. Hugues S, Malherbe L, Filippi C, Glaichenhaus N (2002) Generation and use of alternative multimers of peptide/MHC complexes. J Immunol Methods 268:83–92
154. Watts C, Moss CX, Mazzeo D, West MA, Matthews SP, Li DN, Manoury B (2003) Creation versus destruction of T cell epitopes in the class II MHC pathway. Ann N Y Acad Sci 987:9–14
155. Saunders PM, van Endert P (2011) Running the gauntlet, from peptide generation to antigen presentation by MHC class I. Tissue Antigens 78:161–170
156. Pogue RR, Eron J, Frelinger JA, Matsui M (1995) Amino-terminal alteration of the HLA-A*0201-restricted human immunodeficiency virus pol peptide increases complex stability and in vitro immunogenicity. Proc Natl Acad Sci U S A 92:8166–8170
157. Abdel-Motal UM, Friedline R, Poligone B, Pogue-Caley RR, Frelinger JA, Tisch R (2001) Dendritic cell vaccination induces cross-reactive cytotoxic T lymphocytes specific for wild-type and natural variant human immunodeficiency virus type 1 epitopes in HLA-A*0201/Kb transgenic mice. Clin Immunol 101:51–58
158. Kessler JH, Melief CJ (2007) Identification of T-cell epitopes for cancer immunotherapy. Leukemia 21:1859–1874
159. Hunt DF, Michel H, Dickinson TA, Shabanowitz J, Cox AL, Sakaguchi K, Appella E, Grey HM, Sette A (1992) Peptides presented to the immune system by the murine class II major histocompatibility complex molecule I-Ad. Science 256:1817–1820
160. Anthony DD, Lehmann PV (2003) T-cell epitope mapping using the ELISPOT approach. Methods 29:260–269
161. Koelle DM (2003) Expression cloning for the discovery of viral antigens and epitopes recognized by T cells. Methods 29:213–226
162. Lemmel C, Stevanovic S (2003) The use of HPLC-MS in T-cell epitope identification. Methods 29:248–259
163. Wang RF (2003) Identification of MHC class II-restricted tumor antigens recognized by CD4+ T cells. Methods 29:227–235
164. Buus S, Claesson MH (2004) Identifying multiple tumor-specific epitopes from large-scale screening for overexpressed mRNA. Curr Opin Immunol 16:137–142
165. Wang M, Bai F, Pries M, Buus S, Prause JU, Nissen MH (2006) Identification of MHC class I H-2 Kb/Db-restricted immunogenic peptides derived from retinal proteins. Invest Ophthalmol Vis Sci 47:3939–3945
166. Gunawardana CG, Diamandis EP (2007) High throughput proteomic strategies for identifying tumour-associated antigens. Cancer Lett 249:110–119
167. Thorn M, Wang M, Kloverpris H, Schmidt EG, Fomsgaard A, Wenandy L, Berntsen A, Brunak S, Buus S, Claesson MH (2007) Identification of a new hTERT-derived HLA-A*0201 restricted, naturally processed CTL epitope. Cancer Immunol Immunother 56:1755–1763
168. Richards KA, Topham D, Chaves FA, Sant AJ (2010) Cutting edge, CD4 T cells generated from encounter with seasonal influenza viruses and vaccines have broad protein specificity and can directly recognize naturally generated epitopes derived from the live pandemic H1N1 virus. J Immunol 185:4998–5002
169. Provenzano M, Panelli MC, Mocellin S, Bracci L, Sais G, Stroncek DF, Spagnoli GC, Marincola FM (2006) MHC-peptide specificity and T-cell epitope mapping, where

immunotherapy starts. Trends Mol Med 12:465–472
170. Rudensky A, Preston-Hurlburt P, Hong SC, Barlow A, Janeway CA Jr (1991) Sequence analysis of peptides bound to MHC class II molecules. Nature 353:622–627
171. Rudensky A, Preston-Hurlburt P, al-Ramadi BK, Rothbard J, Janeway CA Jr (1992) Truncation variants of peptides isolated from MHC class II molecules suggest sequence motifs. Nature 359:429–431
172. Engelhard VH (1994) Structure of peptides associated with MHC class I molecules. Curr Opin Immunol 6:13–23
173. Engelhard VH (1994) Structure of peptides associated with class I and class II MHC molecules. Annu Rev Immunol 12:181–207
174. Alfonso C, Karlsson L (2000) Nonclassical MHC class II molecules. Annu Rev Immunol 18:113–142
175. Santambrogio L, Strominger JL (2006) The ins and outs of MHC class II proteins in dendritic cells. Immunity 25:857–859
176. Storkus WJ, Zeh HJ 3rd, Salter RD, Lotze MT (1993) Identification of T-cell epitopes, rapid isolation of class I-presented peptides from viable cells by mild acid elution. J Immunother Emphasis Tumor Immunol 14:94–103
177. Van Bleek GM, Nathenson SG (1990) Isolation of an endogenously processed immunodominant viral peptide from the class I H-2Kb molecule. Nature 348:213–216
178. Prilliman K, Lindsey M, Zuo Y, Jackson KW, Zhang Y, Hildebrand W (1997) Large-scale production of class I bound peptides, assigning a signature to HLA-B*1501. Immunogenetics 45:379–385
179. Duyar H, Dengjel J, de Graaf KL, Wiesmuller KH, Stevanovic S, Weissert R (2005) Peptide motif for the rat MHC class II molecule RT1. Da, similarities to the multiple sclerosis-associated HLA-DRB1*1501 molecule. Immunogenetics 57:69–76
180. Fissolo N, Haag S, de Graaf KL, Drews O, Stevanovic S, Rammensee HG, Weissert R (2009) Naturally presented peptides on major histocompatibility complex I and II molecules eluted from central nervous system of multiple sclerosis patients. Mol Cell Proteomics 8:2090–2101
181. Wahlstrom J, Dengjel J, Persson B, Duyar H, Rammensee HG, Stevanovic S, Eklund A, Weissert R, Grunewald J (2007) Identification of HLA-DR-bound peptides presented by human bronchoalveolar lavage cells in sarcoidosis. J Clin Invest 117:3576–3582
182. Demine R, Sherev T, Walden P (2003) Biochemical determination of natural tumor-associated T-cell epitopes. Mol Biotechnol 25:71–78
183. Verma B, Hawkins OE, Neethling FA, Caseltine SL, Largo SR, Hildebrand WH, Weidanz JA (2010) Direct discovery and validation of a peptide/MHC epitope expressed in primary human breast cancer cells using a TCRm monoclonal antibody with profound antitumor properties. Cancer Immunol Immunother 59:563–573
184. McMurtrey CP, Lelic A, Piazza P, Chakrabarti AK, Yablonsky EJ, Wahl A, Bardet W, Eckerd A, Cook RL, Hess R, Buchli R, Loeb M, Rinaldo CR, Bramson J, Hildebrand WH (2008) Epitope discovery in West Nile virus infection, Identification and immune recognition of viral epitopes. Proc Natl Acad Sci U S A 105:2981–2986
185. Falk K, Rotzschke O (1993) Consensus motifs and peptide ligands of MHC class I molecules. Semin Immunol 5:81–94
186. Lemmel C, Weik S, Eberle U, Dengjel J, Kratt T, Becker HD, Rammensee HG, Stevanovic S (2004) Differential quantitative analysis of MHC ligands by mass spectrometry using stable isotope labeling. Nat Biotechnol 22:450–454
187. Bozzacco L, Yu H, Zebroski HA, Dengjel J, Deng H, Mojsov S, Steinman RM (2011) Mass spectrometry analysis and quantitation of peptides presented on the MHC II molecules of mouse spleen dendritic cells. J Proteome Res 10:5016–5030
188. Mester G, Hoffmann V, Stevanovic S (2011) Insights into MHC class I antigen processing gained from large-scale analysis of class I ligands. Cell Mol Life Sci 68: 1521–1532
189. Hunt DF, Henderson RA, Shabanowitz J, Sakaguchi K, Michel H, Sevilir N, Cox AL, Appella E, Engelhard VH (1992) Characterization of peptides bound to the class I MHC molecule HLA-A2.1 by mass spectrometry. Science 255:1261–1263
190. Henderson RA, Cox AL, Sakaguchi K, Appella E, Shabanowitz J, Hunt DF, Engelhard VH (1993) Direct identification of an endogenous peptide recognized by multiple HLA-A2.1-specific cytotoxic T cells. Proc Natl Acad Sci U S A 90:10275–10279
191. Cox AL, Skipper J, Chen Y, Henderson RA, Darrow TL, Shabanowitz J, Engelhard VH, Hunt DF, Slingluff CL Jr (1994) Identification of a peptide recognized by five melanoma-specific human cytotoxic T cell lines. Science 264:716–719
192. den Haan JM, Sherman NE, Blokland E, Huczko E, Koning F, Drijfhout JW, Skipper J, Shabanowitz J, Hunt DF, Engelhard VH

et al (1995) Identification of a graft versus host disease-associated human minor histocompatibility antigen. Science 268:1476–1480
193. Carralot JP, Lemmel C, Stevanovic S, Pascolo S (2008) Mass spectrometric identification of an HLA-A*0201 epitope from *Plasmodium falciparum* MSP-1. Int Immunol 20:1451–1456
194. Hawkins OE, Vangundy RS, Eckerd AM, Bardet W, Buchli R, Weidanz JA, Hildebrand WH (2008) Identification of breast cancer peptide epitopes presented by HLA-A*0201. J Proteome Res 7:1445–1457
195. Weinzierl AO, Maurer D, Altenberend F, Schneiderhan-Marra N, Klingel K, Schoor O, Wernet D, Joos T, Rammensee HG, Stevanovic S (2008) A cryptic vascular endothelial growth factor T-cell epitope, identification and characterization by mass spectrometry and T-cell assays. Cancer Res 68:2447–2454
196. Seward RJ, Drouin EE, Steere AC, Costello CE (2011) Peptides presented by HLA-DR molecules in synovia of patients with rheumatoid arthritis or antibiotic-refractory Lyme arthritis. Mol Cell Proteomics 10(M110):002477
197. Park S, Lim Y, Lee D, Cho B, Bang YJ, Sung S, Kim HY, Kim DK, Lee YS, Song Y, Jeoung DI (2003) Identification and characterization of a novel cancer/testis antigen gene CAGE-1. Biochim Biophys Acta 1625:173–182
198. Gygi SP, Rist B, Griffin TJ, Eng J, Aebersold R (2002) Proteome analysis of low-abundance proteins using multidimensional chromatography and isotope-coded affinity tags. J Proteome Res 1:47–54
199. Ross PL, Huang YN, Marchese JN, Williamson B, Parker K, Hattan S, Khainovski N, Pillai S, Dey S, Daniels S, Purkayastha S, Juhasz P, Martin S, Bartlet-Jones M, He F, Jacobson A, Pappin DJ (2004) Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. Mol Cell Proteomics 3:1154–1169
200. Cagney G, Emili A (2002) De novo peptide sequencing and quantitative profiling of complex protein mixtures using mass-coded abundance tagging. Nat Biotechnol 20:163–170
201. Tan CT, Croft NP, Dudek NL, Williamson NA, Purcell AW (2011) Direct quantitation of MHC-bound peptide epitopes by selected reaction monitoring. Proteomics 11:2336–2340
202. Lange V, Picotti P, Domon B, Aebersold R (2008) Selected reaction monitoring for quantitative proteomics, a tutorial. Mol Syst Biol 4:222
203. Picotti P, Rinner O, Stallmach R, Dautel F, Farrah T, Domon B, Wenschuh H, Aebersold R (2010) High-throughput generation of selected reaction-monitoring assays for proteins and proteomes. Nat Methods 7:43–46
204. Hogan KT, Sutton JN, Chu KU, Busby JA, Shabanowitz J, Hunt DF, Slingluff CL Jr (2005) Use of selected reaction monitoring mass spectrometry for the detection of specific MHC class I peptide antigens on A3 supertype family members. Cancer Immunol Immunother 54:359–371
205. Ishioka GY, Lamont AG, Thomson D, Bulbow N, Gaeta FC, Sette A, Grey HM (1992) MHC interaction and T cell recognition of carbohydrates and glycopeptides. J Immunol 148:2446–2451
206. Carbone FR, Gleeson PA (1997) Carbohydrates and antigen recognition by T cells. Glycobiology 7:725–730
207. Kastrup IB, Andersen MH, Elliott T, Haurum JS (2001) MHC-restricted T cell responses against posttranslationally modified peptide antigens. Adv Immunol 78:267–289
208. Kastrup IB, Stevanovic S, Arsequell G, Valencia G, Zeuthen J, Rammensee HG, Elliott T, Haurum JS (2000) Lectin purified human class I MHC-derived peptides, evidence for presentation of glycopeptides in vivo. Tissue Antigens 56:129–135
209. Haurum JS, Arsequell G, Lellouch AC, Wong SY, Dwek RA, McMichael AJ, Elliott T (1994) Recognition of carbohydrate by major histocompatibility complex class I-restricted, glycopeptide-specific cytotoxic T lymphocytes. J Exp Med 180:739–744
210. Haurum JS, Tan L, Arsequell G, Frodsham P, Lellouch AC, Moss PA, Dwek RA, McMichael AJ, Elliott T (1995) Peptide anchor residue glycosylation, effect on class I major histocompatibility complex binding and cytotoxic T lymphocyte recognition. Eur J Immunol 25:3270–3276
211. Haurum JS, Hoier IB, Arsequell G, Neisig A, Valencia G, Zeuthen J, Neefjes J, Elliott T (1999) Presentation of cytosolic glycosylated peptides by human class I major histocompatibility complex molecules in vivo. J Exp Med 190:145–150
212. Glithero A, Tormo J, Haurum JS, Arsequell G, Valencia G, Edwards J, Springer S, Townsend A, Pao YL, Wormald M, Dwek RA, Jones EY, Elliott T (1999) Crystal structures of two H-2Db/glycopeptide complexes suggest a molecular basis for CTL cross-reactivity. Immunity 10:63–74
213. Zhang H, Li XJ, Martin DB, Aebersold R (2003) Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. Nat Biotechnol 21:660–666
214. Andersen MH, Bonfill JE, Neisig A, Arsequell G, Sondergaard I, Valencia G, Neefjes J,

Zeuthen J, Elliott T, Haurum JS (1999) Phosphorylated peptides can be transported by TAP molecules, presented by class I MHC molecules, and recognized by phosphopeptide-specific CTL. J Immunol 163:3812–3818
215. Zarling AL, Ficarro SB, White FM, Shabanowitz J, Hunt DF, Engelhard VH (2000) Phosphorylated peptides are naturally processed and presented by major histocompatibility complex class I molecules in vivo. J Exp Med 192:1755–1762
216. Meyer VS, Drews O, Gunder M, Hennenlotter J, Rammensee HG, Stevanovic S (2009) Identification of natural MHC class II presented phosphopeptides and tumor-derived MHC class I phospholigands. J Proteome Res 8:3666–3674
217. Zarling AL, Polefrone JM, Evans AM, Mikesh LM, Shabanowitz J, Lewis ST, Engelhard VH, Hunt DF (2006) Identification of class I MHC-associated phosphopeptides as targets for cancer immunotherapy. Proc Natl Acad Sci U S A 103:14889–14894
218. Depontieu FR, Qian J, Zarling AL, McMiller TL, Salay TM, Norris A, English AM, Shabanowitz J, Engelhard VH, Hunt DF, Topalian SL (2009) Identification of tumor-associated, MHC class II-restricted phosphopeptides as targets for immunotherapy. Proc Natl Acad Sci U S A 106:12073–12078
219. Li Y, Depontieu FR, Sidney J, Salay TM, Engelhard VH, Hunt DF, Sette A, Topalian SL, Mariuzza RA (2010) Structural basis for the presentation of tumor-associated MHC class II-restricted phosphopeptides to CD4+ T cells. J Mol Biol 399:596–603
220. Dinarello CA (2007) Historical insights into cytokines. Eur J Immunol 37(Suppl 1): S34–S45
221. Dinarello CA (2010) Anti-inflammatory agents, present and future. Cell 140:935–950
222. Bruserud O, Kittang AO (2010) The chemokine system in experimental and clinical hematology. Curr Top Microbiol Immunol 341:3–12
223. Pult I, Netter HJ, Bruns M, Prassolov A, Sirma H, Hohenberg H, Chang SF, Frolich K, Krone O, Kaleta EF, Will H (2001) Identification and analysis of a new hepadnavirus in white storks. Virology 289:114–128
224. Leng SX, McElhaney JE, Walston JD, Xie D, Fedarko NS, Kuchel GA (2008) ELISA and multiplex technologies for cytokine measurement in inflammation and aging research. J Gerontol A Biol Sci Med Sci 63:879–884
225. Yu X, Schneiderhan-Marra N, Hsu HY, Bachmann J, Joos TO (2009) Protein microarrays, effective tools for the study of inflammatory diseases. Methods Mol Biol 577:199–214
226. Han KC, Ahn DR, Yang EG (2010) An approach to multiplexing an immunosorbent assay with antibody-oligonucleotide conjugates. Bioconjug Chem 21:2190–2196
227. Szarka A, Rigo J Jr, Lazar L, Beko G, Molvarec A (2010) Circulating cytokines, chemokines and adhesion molecules in normal pregnancy and preeclampsia determined by multiplex suspension array. BMC Immunol 11:59
228. Yu X, Hartmann M, Wang Q, Poetz O, Schneiderhan-Marra N, Stoll D, Kazmaier C, Joos TO (2010) microFBI, a microfluidic bead-based immunoassay for multiplexed detection of proteins from a microL sample volume. PLoS One 1:e3125
229. Han KC, Yang EG, Ahn DR (2012) A highly sensitive, multiplex immunoassay using gold nanoparticle-enhanced signal amplification. Chem Commun (Camb) 48:5895–5897
230. Fuson KL, Zheng M, Craxton M, Pataer A, Ramesh R, Chada S, Sutton RB (2009) Structural mapping of post-translational modifications in human interleukin-24, role of N-linked glycosylation and disulfide bonds in secretion and activity. J Biol Chem 284:30526–30533
231. Guzman NA, Phillips TM (2005) Immunoaffinity CE for proteomics studies. Anal Chem 77:61A–67A
232. Guzman NA, Blanc T, Phillips TM (2008) Immunoaffinity capillary electrophoresis as a powerful strategy for the quantification of low-abundance biomarkers, drugs, and metabolites in biological matrices. Electrophoresis 29:3259–3278
233. Guzman NA, Phillips TM (2011) Immunoaffinity capillary electrophoresis, a new versatile tool for determining protein biomarkers in inflammatory processes. Electrophoresis 32:1565–1578
234. Boyle MD, Hess JL, Nuara AA, Buller RM (2006) Application of immunoproteomics to rapid cytokine detection. Methods 38: 342–350
235. Hortin GL, Sviridov D, Anderson NL (2008) High-abundance polypeptides of the human plasma proteome comprising the top 4 logs of polypeptide abundance. Clin Chem 54:1608–1616
236. Bjorhall K, Miliotis T, Davidsson P (2005) Comparison of different depletion strategies for improved resolution in proteomic analysis of human serum samples. Proteomics 5:307–317
237. Tirumalai RS, Chan KC, Prieto DA, Issaq HJ, Conrads TP, Veenstra TD (2003)

Characterization of the low molecular weight human serum proteome. Mol Cell Proteomics 2:1096–1103
238. Harper RG, Workman SR, Schuetzner S, Timperman AT, Sutton JN (2004) Low-molecular-weight human serum proteome using ultrafiltration, isoelectric focusing, and mass spectrometry. Electrophoresis 25: 1299–1306
239. Groessl M, Luksch H, Rosen-Wolff A, Shevchenko A, Gentzel M (2012) Profiling of the human monocytic cell secretome by quantitative label-free mass spectrometry identifies stimulus-specific cytokines and pro-inflammatory proteins. Proteomics 12: 2833–2842
240. Van Regenmortel MH (2009) What is a B-cell epitope? Methods Mol Biol 524:3–20
241. Greenbaum JA, Andersen PH, Blythe M, Bui HH, Cachau RE, Crowe J, Davies M, Kolaskar AS, Lund O, Morrison S, Mumey B, Ofran Y, Pellequer JL, Pinilla C, Ponomarenko JV, Raghava GP, van Regenmortel MH, Roggen EL, Sette A, Schlessinger A, Sollner J, Zand M, Peters B (2007) Towards a consensus on datasets and evaluation metrics for developing B-cell epitope prediction tools. J Mol Recognit 20:75–82
242. Saha S, Bhasin M, Raghava GP (2005) Bcipep, a database of B-cell epitopes. BMC Genomics 6:79
243. Haste Andersen P, Nielsen M, Lund O (2006) Prediction of residues in discontinuous B-cell epitopes using protein 3D structures. Protein Sci 15:2558–2567
244. Moreau V, Granier C, Villard S, Laune D, Molina F (2006) Discontinuous epitope prediction based on mimotope analysis. Bioinformatics 22:1088–1095
245. Saha S, Raghava GP (2006) Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. Proteins 65:40–48
246. Bublil EM, Freund NT, Mayrose I, Penn O, Roitburd-Berman A, Rubinstein ND, Pupko T, Gershoni JM (2007) Stepwise prediction of conformational discontinuous B-cell epitopes using the Mapitope algorithm. Proteins 68:294–304
247. Saha S, Raghava GP (2007) Prediction methods for B-cell epitopes. Methods Mol Biol 409:387–394
248. Saha S, Raghava GP (2007) Searching and mapping of B-cell epitopes in Bcipep database. Methods Mol Biol 409:113–124
249. Moreau V, Fleury C, Piquer D, Nguyen C, Novali N, Villard S, Laune D, Granier C, Molina F (2008) PEPOP, computational design of immunogenic peptides. BMC Bioinformatics 9:71
250. Ponomarenko J, Bui HH, Li W, Fusseder N, Bourne PE, Sette A, Peters B (2008) ElliPro, a new structure-based tool for the prediction of antibody epitopes. BMC Bioinformatics 9:514
251. Sweredoski MJ, Baldi P (2008) PEPITO, improved discontinuous B-cell epitope prediction using multiple distance thresholds and half sphere exposure. Bioinformatics 24:1459–1460
252. Sun J, Wu D, Xu T, Wang X, Xu X, Tao L, Li YX, Cao ZW (2009) SEPPA, a computational server for spatial epitope prediction of protein antigens. Nucleic Acids Res 37:W612–W616
253. Liang S, Zheng D, Standley DM, Yao B, Zacharias M, Zhang C (2010) EPSVR and EPMeta, prediction of antigenic epitopes using support vector regression and multiple server results. BMC Bioinformatics 11:381
254. Zhang W, Xiong Y, Zhao M, Zou H, Ye X, Liu J (2011) Prediction of conformational B-cell epitopes from 3D structures by random forests with a distance-based feature. BMC Bioinformatics 12:341
255. Evans MC (2008) Recent advances in immunoinformatics, application of in silico tools to drug development. Curr Opin Drug Discov Devel 11:233–241
256. Bublil EM, Yeger-Azuz S, Gershoni JM (2006) Computational prediction of the cross-reactive neutralizing epitope corresponding to the [corrected] monclonal [corrected] antibody b12 specific for HIV-1 gp120. FASEB J 20:1762–1774
257. Tong JC, Ren EC (2009) Immunoinformatics, current trends and future directions. Drug Discov Today 14:684–689
258. Tomar N, De RK (2010) Immunoinformatics, an integrated scenario. Immunology 131: 153–168
259. Bordner AJ, Mittelmann HD (2010) MultiRTA, a simple yet reliable method for predicting peptide binding affinities for multiple class II MHC allotypes. BMC Bioinformatics 11:482
260. Bordner AJ, Abagyan R (2006) Ab initio prediction of peptide-MHC binding geometry for diverse class I MHC allotypes. Proteins 63:512–526
261. Tsurui H, Takahashi T (2007) Prediction of T-cell epitope. J Pharmacol Sci 105:299–316
262. Lundegaard C, Lund O, Buus S, Nielsen M (2010) Major histocompatibility complex class I binding predictions as a tool in epitope discovery. Immunology 130:309–318
263. Lundegaard C, Lund O, Nielsen M (2011) Prediction of epitopes using neural network based methods. J Immunol Methods 374: 26–34

264. Doolan DL, Aguiar JC, Weiss WR, Sette A, Felgner PL, Regis DP, Quinones-Casas P, Yates JR 3rd, Blair PL, Richie TL, Hoffman SL, Carucci DJ (2003) Utilization of genomic sequence information to develop malaria vaccines. J Exp Biol 206:3789–3802
265. Moise L, McMurry JA, Buus S, Frey S, Martin WD, De Groot AS (2009) In silico-accelerated identification of conserved and immunogenic variola/vaccinia T-cell epitopes. Vaccine 27:6471–6479
266. Walshe VA, Hattotuwagama CK, Doytchinova IA, Wong M, Macdonald IK, Mulder A, Claas FH, Pellegrino P, Turner J, Williams I, Turnbull EL, Borrow P, Flower DR (2009) Integrating in silico and in vitro analysis of peptide binding affinity to HLA-Cw*0102, a bioinformatic approach to the prediction of new epitopes. PLoS One 4:e8095
267. Driguez P, Doolan DL, Loukas A, Felgner PL, McManus DP (2010) Schistosomiasis vaccine discovery using immunomics. Parasit Vectors 3:4
268. Schirle M, Keilholz W, Weber B, Gouttefangeas C, Dumrese T, Becker HD, Stevanovic S, Rammensee HG (2000) Identification of tumor-associated MHC class I ligands by a novel T cell-independent approach. Eur J Immunol 30:2216–2225
269. Pascolo S, Schirle M, Guckel B, Dumrese T, Stumm S, Kayser S, Moris A, Wallwiener D, Rammensee HG, Stevanovic S (2001) A MAGE-A1 HLA-A A*0201 epitope identified by mass spectrometry. Cancer Res 61:4072–4077
270. Schirle M, Weinschenk T, Stevanovic S (2001) Combining computer algorithms with experimental approaches permits the rapid and accurate identification of T cell epitopes from defined antigens. J Immunol Methods 257:1–16
271. Bozzacco L, Yu H, Dengjel J, Trumpfheller C, Zebroski HA 3rd, Zhang N, Kuttner V, Ueberheide BM, Deng H, Chait BT, Steinman RM, Mojsov S, Fenyo D (2012) Strategy for identifying dendritic cell-processed CD4+ T cell epitopes from the HIV gag p24 protein. PLoS One 7:e41897
272. Messitt TJ, Terry F, Moise L, Martin W, De Groot AS (2011) A comparison of two methods for T cell epitope mapping, "cell free" in vitro versus immunoinformatics. Immunome Res 7:6
273. Nakayasu ES, Sobreira TJ, Torres R Jr, Ganiko L, Oliveira PS, Marques AF, Almeida IC (2012) Improved proteomic approach for the discovery of potential vaccine targets in Trypanosoma cruzi. J Proteome Res 11:237–246
274. Caron E, Vincent K, Fortier MH, Laverdure JP, Bramoulle A, Hardy MP, Voisin G, Roux PP, Lemieux S, Thibault P, Perreault C (2011) The MHC I immunopeptidome conveys to the cell surface an integrative view of cellular regulation. Mol Syst Biol 7:533
275. Pizza M, Scarlato V, Masignani V, Giuliani MM, Arico B, Comanducci M, Jennings GT, Baldi L, Bartolini E, Capecchi B, Galeotti CL, Luzzi E, Manetti R, Marchetti E, Mora M, Nuti S, Ratti G, Santini L, Savino S, Scarselli M, Storni E, Zuo P, Broeker M, Hundt E, Knapp B, Blair E, Mason T, Tettelin H, Hood DW, Jeffries AC, Saunders NJ, Granoff DM, Venter JC, Moxon ER, Grandi G, Rappuoli R (2000) Identification of vaccine candidates against serogroup B meningococcus by whole-genome sequencing. Science 287:1816–1820
276. Sano T, Smith CL, Cantor CR (1992) Immuno-PCR, very sensitive antigen detection by means of specific antibody-DNA conjugates. Science 258:120–122
277. Niemeyer CM, Adler M, Wacker R (2007) Detecting antigens by quantitative immuno-PCR. Nat Protoc 2:1918–1930
278. Kasai N, Kobayashi K, Shioya S, Yoshikawa Y, Yotsumoto F, Miyamoto S, Mekada E, Enokizono J (2012) Soluble heparin-binding EGF-like growth factor (HB-EGF) detected by newly developed immuno-PCR method is a clear-cut serological biomarker for ovarian cancer. Am J Transl Res 4:415–421
279. Kuczius T, Becker K, Fischer A, Zhang W (2012) Simultaneous detection of three CNS indicator proteins in complex suspensions using a single immuno-PCR protocol. Anal Biochem 431:4–10
280. Hashimoto M, Aoki M, Winblad B, Tjernberg LO (2012) A novel approach for Abeta(1)(−)(4)(0) quantification using immuno-PCR. J Neurosci Methods 205:364–367
281. Malou N, Renvoise A, Nappez C, Raoult D (2012) Immuno-PCR for the early serological diagnosis of acute infectious diseases, the Q fever paradigm. Eur J Clin Microbiol Infect Dis 31:1951–1960
282. Potuckova L, Franko F, Bambouskova M, Draber P (2011) Rapid and sensitive detection of cytokines using functionalized gold nanoparticle-based immuno-PCR, comparison with immuno-PCR and ELISA. J Immunol Methods 371:38–47
283. Chao HY, Wang YC, Tang SS, Liu HW (2004) A highly sensitive immuno-polymerase chain reaction assay for *Clostridium botulinum* neurotoxin type A. Toxicon 43:27–34
284. Allen RC, Rogelj S, Cordova SE, Kieft TL (2006) An immuno-PCR method for detecting Bacillus thuringiensis Cry1Ac toxin. J Immunol Methods 308:109–115
285. Fischer A, von Eiff C, Kuczius T, Omoe K, Peters G, Becker K (2007) A quantitative

real-time immuno-PCR approach for detection of staphylococcal enterotoxins. J Mol Med (Berl) 85:461–469
286. Zhang W, Bielaszewska M, Pulz M, Becker K, Friedrich AW, Karch H, Kuczius T (2008) New immuno-PCR assay for detection of low concentrations of shiga toxin 2 and its variants. J Clin Microbiol 46:1292–1297
287. He X, Qi W, Quinones B, McMahon S, Cooley M, Mandrell RE (2011) Sensitive detection of Shiga Toxin 2 and some of its variants in environmental samples by a novel immuno-PCR assay. Appl Environ Microbiol 77:3558–3564
288. Adler M (2005) Immuno-PCR as a clinical laboratory tool. Adv Clin Chem 39:239–292
289. Adler M, Wacker R, Niemeyer CM (2003) A real-time immuno-PCR assay for routine ultrasensitive quantification of proteins. Biochem Biophys Res Commun 308: 240–250
290. Lind K, Kubista M (2005) Development and evaluation of three real-time immuno-PCR assemblages for quantification of PSA. J Immunol Methods 304:107–116
291. Niemeyer CM, Adler M, Wacker R (2005) Immuno-PCR, high sensitivity detection of proteins by nucleic acid amplification. Trends Biotechnol 23:208–216
292. Malou N, Raoult D (2011) Immuno-PCR, a promising ultrasensitive diagnostic method to detect antigens and antibodies. Trends Microbiol 19:295–302
293. Ferreira L, Sanchez-Juanes F, Munoz-Bellido JL, Gonzalez-Buitrago JM (2011) Rapid method for direct identification of bacteria in urine and blood culture samples by matrix-assisted laser desorption ionization time-of-flight mass spectrometry, intact cell vs. extraction method. Clin Microbiol Infect 17:1007–1012
294. Welker M (2011) Proteomics for routine identification of microorganisms. Proteomics 11:3143–3153
295. Welker M, Moore ER (2011) Applications of whole-cell matrix-assisted laser-desorption/ionization time-of-flight mass spectrometry in systematic microbiology. Syst Appl Microbiol 34:2–11
296. Sogawa K, Watanabe M, Sato K, Segawa S, Miyabe A, Murata S, Saito T, Nomura F (2012) Rapid identification of microorganisms by mass spectrometry, improved performance by incorporation of in-house spectral data into a commercial database. Anal Bioanal Chem 403:1811–1822
297. El-Bouri K, Johnston S, Rees E, Thomas I, Bome-Mannathoko N, Jones C, Reid M, Ben-Ismaeil B, Davies AR, Harris LG, Mack D (2012) Comparison of bacterial identification by MALDI-TOF mass spectrometry and conventional diagnostic microbiology methods, agreement, speed and cost implications. Br J Biomed Sci 69:47–55
298. Ouedraogo R, Flaudrops C, Ben Amara A, Capo C, Raoult D, Mege JL (2010) Global analysis of circulating immune cells by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. PLoS One 5:e13691
299. Ouedraogo R, Daumas A, Ghigo E, Capo C, Mege JL, Textoris J (2012) Whole-cell MALDI-TOF MS, a new tool to assess the multifaceted activation of macrophages. J Proteomics 75:5523–5532
300. Nuti DE, Crump RB, Dwi Handayani F, Chantratita N, Peacock SJ, Bowen R, Felgner PL, Davies DH, Wu T, Lyons CR, Brett PJ, Burtnick MN, Kozel TR, AuCoin DP (2011) Identification of circulating bacterial antigens by in vivo microbial antigen discovery. MBio 16:e00136
301. Testa JS, Shetty V, Hafner J, Nickens Z, Kamal S, Sinnathamby G, Philip R (2012) MHC class I-presented T cell epitopes identified by immunoproteomics analysis are targets for a cross reactive influenza-specific T cell response. PLoS One 7:e48484
302. Testa JS, Shetty V, Sinnathamby G, Nickens Z, Hafner J, Kamal S, Zhang X, Jett M, Philip R (2012) Conserved MHC class I-presented dengue virus epitopes identified by immunoproteomics analysis are targets for cross-serotype reactive T-cell response. J Infect Dis 205:647–655
303. Macela A, Stulik J, Hernychova L, Kroca M, Krocova Z, Kovarova H (1996) The immune response against *Francisella tularensis* live vaccine strain in Lps(n) and Lps(d) mice. FEMS Immunol Med Microbiol 13:235–238
304. Havlasova J, Hernychova L, Halada P, Pellantova V, Krejsek J, Stulik J, Macela A, Jungblut PR, Larsson P, Forsman M (2002) Mapping of immunoreactive antigens of *Francisella tularensis* live vaccine strain. Proteomics 2:857–867
305. Janovska S, Pavkova I, Hubalek M, Lenco J, Macela A, Stulik J (2007) Identification of immunoreactive antigens in membrane proteins enriched fraction from *Francisella tularensis* LVS. Immunol Lett 108:151–159
306. Twine SM, Petit MD, Fulton KM, House RV, Conlan JW (2010) Immunoproteomics analysis of the murine antibody response to vaccination with an improved *Francisella tularensis* live vaccine strain (LVS). PLoS One 5:e10000
307. Tan X, Chun S, Pablo J, Felgner P, Liang X, Davies DH (2012) Failure of the smallpox

vaccine to develop a skin lesion in vaccinia virus-naive individuals is related to differences in antibody profiles before vaccination, not after. Clin Vaccine Immunol 19:418–428
308. Hermanson G, Chun S, Felgner J, Tan X, Pablo J, Nakajima-Sasaki R, Molina DM, Felgner PL, Liang X, Davies DH (2012) Measurement of antibody responses to Modified Vaccinia virus Ankara (MVA) and Dryvax((R)) using proteome microarrays and development of recombinant protein ELISAs. Vaccine 30:614–625
309. Cook SD, Quinless JR, Jotkowitz A, Beaton P (2001) Serum IFN neutralizing antibodies and neopterin levels in a cross-section of MS patients. Neurology 57:1080–1084
310. Bertolotto A, Malucchi S, Sala A, Orefice G, Carrieri PB, Capobianco M, Milano E, Melis F, Giordana MT (2002) Differential effects of three interferon betas on neutralising antibodies in patients with multiple sclerosis, a follow up study in an independent laboratory. J Neurol Neurosurg Psychiatry 73:148–153
311. Bertolotto A, Sala A, Malucchi S, Marnetto F, Caldano M, Di Sapio A, Capobianco M, Gilli F (2004) Biological activity of interferon betas in patients with multiple sclerosis is affected by treatment regimen and neutralising antibodies. J Neurol Neurosurg Psychiatry 75:1294–1299
312. Von Wussow P, Hehlmann R, Hochhaus T, Jakschies D, Nolte KU, Prummer O, Ansari H, Hasford J, Heimpel H, Deicher H (1994) Roferon (rIFN-alpha 2a) is more immunogenic than intron A (rIFN-alpha 2b) in patients with chronic myelogenous leukemia. J Interferon Res 14:217–219
313. Ryff JC (1997) Clinical investigation of the immunogenicity of interferon-alpha 2a. J Interferon Cytokine Res 17(Suppl 1):S29–S33
314. Baert F, Noman M, Vermeire S, Van Assche G, D' Haens G, Carbonez A, Rutgeerts P (2003) Influence of immunogenicity on the long-term efficacy of infliximab in Crohn's disease. N Engl J Med 348:601–608
315. Radstake TR, Svenson M, Eijsbouts AM, van den Hoogen FH, Enevold C, van Riel PL, Bendtzen K (2009) Formation of antibodies against infliximab and adalimumab strongly correlates with functional drug levels and clinical responses in rheumatoid arthritis. Ann Rheum Dis 68:1739–1745
316. Baker MP, Jones TD (2007) Identification and removal of immunogenicity in therapeutic proteins. Curr Opin Drug Discov Devel 10:219–227
317. Baker MP, Reynolds HM, Lumicisi B, Bryson CJ (2010) Immunogenicity of protein therapeutics, The key causes, consequences and challenges. Self Nonself 1:314–322

Chapter 3

Antigen Identification Using SEREX

Ugur Sahin and Özlem Türeci

Abstract

Serological analysis of recombinant cDNA expression libraries (SEREX) allows systematic identification of antigens recognized by the spontaneous autoantibody repertoire of patients with cancer and autoimmune disease. SEREX is based on screening of lambda phage expression libraries constructed from diseased tissue with the autologous patient's serum. This chapter provides the key protocols for SEREX immunoscreening and for downstream validation of autoantigens.

Key words SEREX, Autoantibodies, Serology, Autoantigens, Expression cloning, Immunoproteomics, Biomarker

1 Introduction

Autoantibody-based screening methods are potent tools for the identification of disease-associated antigens and analysis of complex immune responses.

In the past, a few proteins were examined at a time using techniques such as one dimensional (1D) sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS PAGE) and sandwich enzyme-linked immunosorbent assays (ELISA). Nowadays, with the development of high-throughput strategies, multiple potential antigens are unraveled in a single experiment. Serological analysis of recombinant cDNA expression libraries (SEREX) [1], the first such approach to be established, is of outstanding sensitivity and coverage [2] and is being used both in cancer and in autoimmunity [3, 4].

From the methodological point of view, characteristics of the SEREX procedure can be summarized as follows (*see* Fig. 1).

cDNA expression libraries are constructed from fresh-frozen disease tissue and packaged into lambda phage vectors. The quality of the library is critical to cast a wide and unbiased net to identify any and all autoantigens prevailing in the patient.

The library is expressed by lytic infection of *Escherichia coli*. Recombinant proteins released in lytic plaques are transferred onto

Kelly M. Fulton and Susan M. Twine (eds.), *Immunoproteomics: Methods and Protocols*, Methods in Molecular Biology, vol. 1061, DOI 10.1007/978-1-62703-589-7_3,

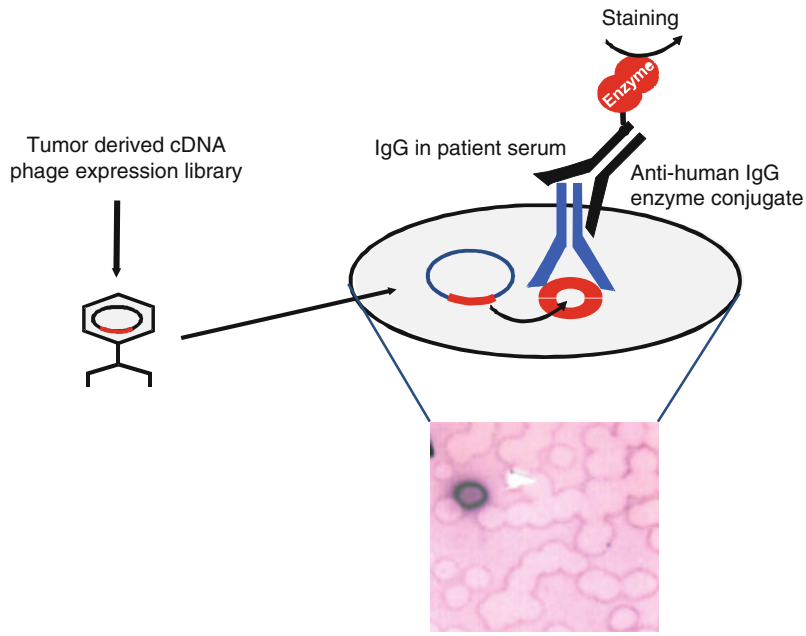


Fig. 1 Serological analysis of recombinant cDNA expression libraries (SEREX)

nitrocellulose membranes (plaque lift). Advantages of this lytic bacterial expression system are its high processivity and broad coverage of potential autoantigens including toxic gene products and linear as well as conformational epitopes. A downside is the lack of autoantigenic epitopes derived from eukaryotic posttranslational modification (e.g., glycosylation).

Nitrocellulose filters are screened with highly diluted autologous patient serum for gene products reactive with high-titered IgG, thus more likely to be immunologically relevant autoantigens with prevalence of T-cell responses (particularly $CD4^+$ T cells). SEREX-defined clones can be directly sequenced for immediate molecular definition of the antigenic targets and subsequent investigation of their expression profiles.

As the majority of cognate autoantibodies in human sera are patient-specific rather than disease-associated, any practical approach for identifying cancer-related autoantigens must include an integral strategy for demonstrating disease relevance early on.

This chapter provides protocols for the integrated screening and validation process.

2 Materials

2.1 Disposables and Equipment

1. Petri dishes (Ø 8 cm, 13.5 cm) (Greiner Bio-One).
2. 50 mL conical Falcon tubes (BD Biosciences).

3. 14 mL Falcon round-bottom white cap tubes (BD Biosciences).
4. Safe lock tubes (0.5 mL, 1.5 mL) (Eppendorf).
5. Nitrocellulose Blotting Membranes, 45 μm (Sartorius AG).
6. Square Petri dish 24.5 cm × 24.5 cm.
7. Self-made plastic spacers, size 23 cm × 2 cm × 0.1 cm to fit into a square Petri dish forming 24 sections (23 cm × 1 cm in size).
8. Self-made incubator dish to allow parallel processing of 24 nitrocellulose membrane slips each in a different serum. Floor space of 25 cm × 56 cm subdivided into incubation segments each 2 cm × 24 cm in size.
9. Branson Sonic Power Sonifier (Co. A. Smithkline).

2.2 Kits

1. Messenger RNA Isolation Kit (Stratagene): Denaturing solution, β-mercaptoethanol, oligo(dT) cellulose, high-salt buffer, low-salt buffer, elution buffer, 5 M NaCl, 3 M sodium acetate, 20 mg/mL glycogen, transfer pipets (RNase-free), push columns (RNase-free), syringes (10 mL) (RNase-free).
2. Uni-ZAP cDNA Synthesis Kit (Stratagene):
   (a) *First-strand reagents*: AccuScript reverse transcriptase (AccuScript-RT), RNase Block ribonuclease inhibitor (40 U/μL), first-strand methyl nucleotide mixture (10 mM dATP, dGTP, and dTTP plus 5 mM 5-methyl dCTP), 10× first-strand buffer, 1.4 μg/μL Linker-primer, diethylpyrocarbonate (DEPC)-treated water.
   (b) *Second-strand reagents*: 10× Second-strand buffer, second-strand dNTP mixture (10 mM dATP, dGTP, and dTTP plus 26 mM dCTP), 1.5 U/μL *Escherichia coli* RNase H, 9.0 U/μL *Escherichia coli* DNA polymerase I.
   (c) *Blunting*: Blunting dNTP mixture (2.5 mM dATP, dGTP, dTTP, and dCTP), 2.5 U/μL cloned *Pfu* DNA polymerase, 3 M sodium acetate.
   (d) *Ligation*, *phosphorylation*, *and digestion reagents*: 0.4 μg/μL *Eco*RI adapters, 10× Ligase buffer, 10 mM rATP, 4 U/μL T4 DNA ligase, 5 U/μL T4 polynucleotide kinase, *Xho*I digestion reagents, 40 U/μL *Xho*I 600 U, *Xho*I buffer supplement.
3. Sizesep 400 Spun columns (Amersham-Bioscience): Spun columns prepacked with Sepharose CL-4B and pre-equilibrated in distilled water containing 0.15 % Kathon CG/IcP Biocide.
4. Uni-Zap XR Vector Kit (Stratagene): Uni-ZAP XR vector predigested with *Eco*RI and *Xho*I, XL1-Blue MRF' strain, SOLR strain, ExAssist interference-resistant helper phage.
5. ZAP-cDNA Gigapack III Gold cloning Kit (Stratagene): Gigapack III Gold packaging extracts.

Only those components of the kits are listed, which are required for protocols described here. Instructions of manufacturers have been modified as outlined below in Subheading 3.

2.3 Reagents

1. DEPC-treated water (RNAse-free).
2. Phenol pre-equilibrated with Tris–HCl to pH 8 (Sigma).
3. Chloroform.
4. 3 M sodium acetate (NaAc) in water.
5. 1 M glycine.
6. 96 % and 70 % (v/v) ethanol with water.
7. 1 M isopropyl β-D-1-thiogalactopyranoside (IPTG) in water, store stock at –20 °C.
8. 50 mg/mL 5-bromo-4-chloro-3-indol-phosphate-toluidine (BCIP) in 100 % *N*-*N* dimethylformamide (DMF).
9. 75 mg/mL nitroblue-tetrazoliumchloride (NBT) in 70 % DMF.
10. 10 % (w/v) NaN_3 in water, store at 4 °C.
11. 10 % (w/v) Thimerosal in water, store at 4 °C.
12. Goat anti-Human IgG labeled with alkaline phosphatase (Dianova).
13. 1 M $MgCl_2$ in water.
14. 1 M $MgSO_4$ in water.
15. Glutaraldehyde-activated kieselgel affinity adsorbent.
16. 10 % (w/v) Maltose in water.
17. 250 mg/mL X-Gal in DMF.
18. 5 mg/mL Tetracycline in 96 % ethanol, store at –20 °C.
19. 100 µg/mL Ampicillin in water, store at –20 °C.
20. 50 µg/mL Kanamycin in water, store at –20 °C.

2.4 Buffers and Growth Media

1. Lysogeny broth (LB): 10 g NaCl, 10 g tryptone, 5 g yeast extract, add deionized H_2O to a final volume of 1 L.
2. LB Broth with 10 mM $MgSO_4$ and 0.2 % (w/v) maltose.
3. LB Agar: 10 g NaCl, 10 g tryptone, 5 g yeast extract, 20 g agar, add deionized H_2O to a final volume of 1 L. Cool 1 L of autoclaved LB agar to 55 °C and add tetracycline (final concentration 12.5 µg/mL), ampicillin (final concentration 150 µg/mL), or kanamycin (final concentration 50 µg/mL). Pour in Petri dishes and store at 4 °C.
4. LB Top Agar: Prepare 1 L LB broth, add 0.4 % (w/v) agarose (Applichem) and 0.4 % agar. Melt and cool to 40 °C before use.
5. Sodium–Tris–EDTA buffer (STE): 0.1 M NaCl, 10 mM Tris–HCl, pH 8.0, 1 mM EDTA, pH 8.0

6. 20× Phosphate-buffered saline (PBS): 2.8 M NaCl, 54 mM KCl, 162 mM Na_2HPO_4, 30 mM KH_2PO_4 adjust pH to 7.4–7.6.
7. SM Buffer: 5.8 g NaCl, 2.0 g $MgSO_4$, 50 mL of 1 M Tris–HCl (pH 7.5), 5 mL of 2 % (w/v) gelatin, add deionized H_2O to a final volume of 1 L.
8. Dilution buffer for patient serum: 0.01 % NaN_3, 0.01 % Thimerosal and 0.5 % low-fat milk powder in 1× Tris-buffered saline (TBS).
9. 10× TBS: 0.5 M Tris-Base, 1.5 M NaCl, adjust pH 7.5 with concentrated hydrochloric acid.
10. 10× Tris-buffered saline and Tween (TBST): add 0.5 % Tween20 to 10× TBS.
11. Blocking buffer: 1× TBST with 5 % low-fat milk powder.
12. Tris–HCl, pH 3.0: 0.05 M Tris-Base, adjust pH to 3 with HCl.
13. 10× 3-(*N*-morpholino)propanesulfonic acid (MOPS) buffer: 200 mM MOPS, 50 mM NaAc, 10 mM ethylenediaminetetraacetic acid (EDTA) in 1 L, adjust to pH 7.0 with NaOH.
14. 10× Color development solution (CDS): 0.1 M Tris–HCl, pH 9.5, 5 mM $MgCl_2$, 0.1 M NaCl, pH 9.5.

3 Methods

3.1 Generation of Lambda Phage Expression Library

3.1.1 Synthesis of Double-Stranded cDNA

1. Place the nonenzymatic reagents of the Uni-ZAP XR cDNA library synthesis kit (*see* **Note 1**) on ice. Only the reverse transcriptase should be kept at –20 °C until use.
2. Pipette 5 µL of 10× first-strand buffer, 3 µL of first-strand methyl nucleotide mixture, 2 µL of linker-primer (1.4 µg/µL), (36–*X*) µL of DEPC-treated water, 1 µL of RNase Block Ribonuclease Inhibitor (40 U/µL). Mix and add the required volume (X µL) of poly(A) RNA (5 µg) (*see* **Note 2**). Incubate for 10 min at room temperature.
3. Add 3 µL of Accuscript-RT (50 U/µL). Mix the first-strand synthesis reaction, spin down and incubate at 37 °C for 1 h. To stop the reaction, place tube on ice.
4. Place all nonenzymatic second-strand components on ice.
5. Add on ice to the first-strand synthesis reaction: 20 µL of 10× second-strand buffer, 6 µL of second-strand dNTP mixture, 114 µL of sterile distilled water, 2 µL of RNase H (1.5 U/µL), 11 µL of DNA polymerase I (9.0 U/µL). Incubate for 2.5 h at exactly 16 °C.
6. Place the reaction tube on ice and add 23 µL of blunting dNTP mixture, 2 µL of cloned pfu DNA polymerase (2.5 U). Incubate at 72 °C for 30 min.

7. Prepare a 1:1 (v/v) emulsion of phenol (pH 7–8) and chloroform and add 200 µL of it to the reaction. Vortex and spin the reaction briefly at maximum speed in a bench top microcentrifuge, at room temperature. Transfer the upper aqueous layer to a new tube without removing any interface.
8. Add an equal volume of chloroform. Vortex, spin, and transfer again the upper layer. Precipitate the cDNA overnight at –20 °C after adding 20 µL of 3 M sodium acetate and 400 µL of 100 % ethanol.
9. Spin at maximum speed for 60 min at 4 °C. Remove supernatant without disturbing the pellet. Wash the pellet with 70 % (v/v) ethanol. Do not mix or vortex. Spin again at maximum speed for 2 min at room temperature with the microfuge tube in the same orientation. Dry the pellet, which contains blunted cDNA.
10. Resuspend the pellet in 9 µL of *Eco*RI adapters and incubate at 4 °C for at least 30 min.
11. Add to the tube containing the blunted cDNA and the *Eco*RI adapters: 1 µL of 10× ligase buffer, 1 µL of 10 mM rATP, 1 µL of T4 DNA ligase (4 U/µL). Incubate overnight at 8 °C.
12. Heat-inactivate at 70 °C for 30 min. Spin and cool to room temperature for 5 min.
13. Add 1 µL of 10× ligase buffer, 2 µL of 10 mM rATP, 6 µL of sterile water, 1 µL of T4 polynucleotide kinase (10 U/µL). Incubate for 30 min at 37 °C.
14. Heat-inactivate at 70 °C for 30 min. Spin and cool to room temperature for 5 min.
15. Add 28 µL of *Xho*I buffer supplement and 3 µL of *Xho*I (40 U/µL). Incubate for 1.5 h at 37 °C.
16. Add 50 µL of a 1:1 (v/v) emulsion of phenol (equilibrated to pH 7–8) and chloroform to the reaction. Vortex and spin briefly at maximum speed at room temperature and transfer the upper aqueous layer to a new tube without removing any interface.
17. Add an equal volume of chloroform and vortex. Spin and transfer again the upper layer. This contains the double-stranded cDNA.

*3.1.2 cDNA Size Fractionation (See **Note 3**)*

1. Place the SizeSep™ 400 spun column in 14 mL Falcon tubes and overlay the column with 1× STE buffer to wash the column.
2. Centrifuge (700 × *g*) for 2 min, discard the flow through. Repeat this washing step. To dry the column, centrifuge for 3 min at 700 × *g*.

3. Place a 1.5 mL tube into the 14 mL Falcon tube beneath the column. Load the cDNA reaction mixture on the middle of the column without touching it.
4. Spin 3 min at 700 × *g* and collect the flow through (1st fraction). Pipet 50 µL 1× STE buffer on the column, spin (3 min at 700 × *g*). Collect the flow through (2nd fraction). Repeat this twice to obtain the 3rd and 4th fraction.
5. Add 5 µL of 3 M sodium acetate and 55 µL 96 % ethanol to each fraction, shake gently and precipitate the cDNA overnight at –20 °C.
6. Spin at maximum speed in a benchtop microfuge tube for 1 h at 4 °C. Wash the pellet by adding 250 µL of 70 % (v/v) ethanol without disturbing it. Centrifuge at maximum speed for 5 min at room temperature in the same orientation as before.
7. Air-dry the pellets and resuspend in 8 µL of sterile H_2O.

3.1.3 Ligation of cDNA into the Lambda Phage Vector and Packaging

1. Add 1 µL of the Uni-ZAP XR vector (1 µg), *X*µL of resuspended cDNA (*see* **Note 4**), 0.5 µL of 10× ligase buffer, 0.5 µL of 10 mM rATP (pH 7.5), and 0.5 µL of T4 DNA ligase (4 U/µL). The three latter reagents may be taken from the Uni-ZAP cDNA Synthesis Kit. Adjust with water for a final volume of 5 µL.
2. Incubate for 2 days at 4 °C or overnight at 12–14 °C.
3. To test packaging into phages 1 day after initiation of ligation, remove one tube of MaxPlax lambda packaging extract (*see* **Note 5**) from the freezer, quickly thaw it between your fingers.
4. Add up to 0.5 µL of the ligated DNA to the packaging extract. Mix well and incubate at room temperature for a maximum of 2 h.
5. Add 500 µL of SM buffer and 20 µL of chloroform and mix. Make serial dilutions of the packaged phage in SM buffer (10^{-2} to 10^{-6}). Prepare the XL1-Blue MRF' host bacteria as in Subheading 3.1.4 (*see* **Note 6**). Infect them with plaques, plate and titer as described in paragraph of Subheading 3.1.5.
6. Count the plaques and determine the titer (pfu/mL):

$$\frac{(\text{Number of plaques}) \times (\text{dilution factor}) \times (\text{total reaction volume})}{(\text{volume of dilution plated}) \times (\text{amount of DNA packaged})}$$

7. Package the remaining ligation reaction at day 2 of ligation repeating **steps 3–6**. Calculate the number of packaging extracts based on the efficiency of the test ligation.
8. After packaging, spin tubes briefly to sediment debris. The supernatants containing the phage are ready for titering and can be stored at 4 °C.

3.1.4 Preparation of Host Bacteria

1. Streak the host strain XLl-Blue MRF' onto an LB agar plate containing the tetracycline and incubate overnight at 37 °C.
2. Inoculate LB Broth with 10 mM $MgSO_4$ and 0.2 % (w/v) maltose, with a single colony and grow at 37 °C on a shaker to an OD_{600} of 0.7–0.9.
3. Spin the cells at 500×*g* for 10 min, discard the supernatant and resuspend in sterile 10 mM $MgSO_4$ to an OD_{600} of 0.5.

3.1.5 Plating and Titering the Library

1. To plate the packaged ligation product, mix 1 µL of the final packaged reaction and 1 µL of its 1:10 dilution each with 200 µL of XL1-Blue MRF' cells to an OD_{600} of 0.5.
2. Allow the phage to attach to the cells by incubating them at 37 °C for 15 min. Add 4 mL LB top agar 15 µL of 0.5 M IPTG, 50 µL of X-gal.
3. Plate onto the LB-tetracycline agar plates in small Petri dishes (Ø 8 cm). Invert the plates and incubate at 37 °C overnight.
4. Count the plaques and calculate the efficiency (*see* Subheading 3.1.3, **step 6**). Background plaques are blue, recombinant plaques will be white and should be at least 50 times more abundant.

*3.1.6 Amplification of the Library (See **Note 7**)*

1. Plate the entire library on large Petri dishes (10^5 pfu/plate) to obtain subconfluent phage plaques by proceeding as described in **steps 1–3** of Subheading 3.1.5 and incubate overnight.
2. Overlay the plates with 8 mL of SM buffer and incubate at 4 °C for 3 h with gentle rocking. Recover the bacteriophage suspension from each plate and pool into a sterile polypropylene container. No glass ware, as phages may stick to surface!
3. Rinse the plates with additional 2 mL of SM buffer and recover.
4. Add chloroform to 5 % (v/v) final concentration. Mix and incubate for 15 min at room temperature.
5. Remove the cell debris by centrifugation for 10 min at 500×*g*. Recover and transfer the supernatant to a new polypropylene container. Add chloroform to a final concentration of 0.3 % (v/v) and store at 4 °C. If storage for a longer period of time is planned, store at –80 °C with 7 % DMSO. Determine quality of library (*see* **Note 8**) and titer before use.

3.2 Preparation of Patient Sera for Immunoscreening

To ensure specific and sensitive detection of autoantigenic epitopes displayed by recombinant proteins expressed in plaques, sera of patients have to be cleared from reactivities against components of the expression system (*E. coli* bacteria, lambda phage proteins, etc.). To this aim, sera are extensively preabsorbed before use in the three-step process described below (*see* **Note 9**).

3.2.1 Preabsorption Against Mechanically Disrupted Bacteria

1. Inoculate LB Broth supplemented with 10 mM $MgSO_4$ and 0.2 % (w/v) maltose, with a single colony of XL1-Blue MRF' and grow overnight at 37 °C with gentle rocking.
2. Harvest the bacteria by centrifugation (3,500 × *g*, 15 min) and resuspend pellet in 5 mL 1× MOPS buffer.
3. Place cells on ice and disrupt by sonication for 5× 20 s.
4. For each individual patient serum fill 2 g of affinity adsorbent in a 50 mL Falcon tube, add 3 mL 1× PBS/0.01 % NaN_3, and incubate for 10 min. Spin and discard supernatant.
5. Add sonicated bacterial lysate to the affinity adsorbent and rotate on an overhead shaker for 4 h at room temperature or overnight at 4 °C.
6. Spin down affinity adsorbent (100 × *g*, 1 min), discard supernatant, and wash column 2× 10 min with 30 mL 1× PBS/0.01 % NaN_3 on an overhead shaker.
7. Incubate affinity adsorbent for 2 h at 4 °C with 1 M glycine. Wash 2× 10 min with 1× PBS/0.01 % NaN_3.
8. Dilute human serum 1:10 in serum dilution buffer. Transfer not more than 40 mL of the diluted serum to each 50 mL Falcon tube with the prepared resin. Shake overhead for 4 h at room temperature or overnight at 4 °C.
9. Spin down the resin (100 × *g*, 1 min) (*see* **Note 10**) and harvest the patient sera.

3.2.2 Preabsorption Against Bacteriophage Proteins and Lytically Infected Bacteria

1. Harvest XL1-Blue MRF' and grow overnight at 37 °C in LB Broth supplemented with 10 mM $MgSO_4$ and 0.2 % (w/v) maltose by centrifugation (3,500 × *g*, 15 min). Discard supernatant and resuspend the pellet in 2 mL 10 mM $MgSO_4$. Transfer 200 μL of it to 4 mL fresh LB Broth supplemented with 10 mM $MgSO_4$ and 0.2 % (w/v) maltose. Remaining bacteria are kept at 4 °C until use in **step 4**.
2. Infect freshly inoculated LB Broth with 10^4 pfu of a wild type lambda ZAPII phage ("blue phages" obtained in Subheading 3.1.5). Shake at 37 °C for 4 h.
3. Add remaining bacterial suspension from **step 2** and shake 2 h at 37 °C.
4. Place cells on ice and disrupt by sonication for 5× 20 s.
5. Fill 2 g of affinity adsorbent in a 50 mL Falcon tube, add 3 mL 1× PBS/0.0 1 % NaN_3, rotate for 10 min. Spin the tube and discard supernatant.
6. Add phage lysed bacteria to the affinity adsorbent and rotate on an overhead shaker for 4 h at room temperature or overnight at 4 °C.

7. Spin down the matrix (100×*g*, 1 min) and discard supernatant. Wash column first 2× 10 min with 30 mL 1× PBS/0.01 % NaN_3 on an overhead shaker. Discard washing buffer.
8. Incubate affinity adsorbent for 2 h at 4 °C with 1 M glycine and wash subsequently 2× 10 min with 1× PBS/0.01 % NaN_3.
9. Transfer the sera obtained by Subheading 3.2.1 to the 50 mL Falcon tubes with the prepared resin. Incubate for 4 h at room temperature or overnight 4 °C on an overhead shaker.
10. Centrifuge to spin down the resin (100×*g*, 1 min) (*see* **Note 10**) and harvest the patient sera.

3.2.3 Preabsorption Against Lytic Filters

1. Prepare *E. coli* XL1-Blue MRF' as described in Subheading 3.1.4.
2. For each large Petri dish agar plate (Ø 13.5 cm) infect 600 µL bacteria (OD_{600} 0.5) with 10^5 pfu of wild type lambda ZAPII phage ("blue phage" obtained in Subheading 3.1.6). Incubate at 37 °C for 15 min. Add 40 µL 1 M IPTG and 8 mL LB top Agar.
3. Pour onto LB-tetracycline agar plates, allow the top agar to solidify and incubate upside down overnight at 37 °C. The high MOI (multiplicity of infection) ensures that plates are lytic the next morning without spared bacterial lawn.
4. Place nitrocellulose filters on plates and incubate for 4 h at 37 °C.
5. Lift filters and replace each by a second one, repeat incubation for 6 h at 37 °C.
6. Immerse filters in a large container with 1× TBST and incubate under strong agitation.
7. Wash 3× 10 min in 1× TBST. Agar residues have to be removed completely.
8. Place each filter into an individual Petri dish and incubate in blocking buffer for 1 h at room temperature with gentle agitation. Wash filters 3× 10 min in a large container with 1× TBS.
9. Place each filter blotted side facing upward into a Petri dish. Transfer each individual human serum processed as described in Subheadings 3.2.1 and 3.2.2 to a filter, incubate for 4 h at room temperature on a horizontal shaker.
10. Remove the membrane, recover the serum completely, and incubate it for a second time with a fresh lytic membrane.
11. Transfer serum into a 500 mL bottle or falcon tubes and store at 4 °C. The serum is now ready to use for immunoscreening.

3.3 Immunoscreening of Lambda Phage Expression Library

If the library has been generated from tissue containing B-lymphocytes (e.g., diseased tissue of primary lymphatic organs, organ tissue with disease-associated brisk B-cell infiltrate) (*see* **Note 11**).

1. Inoculate a single colony from the freshly plated XL1-Blue MRF' cells in 20 mL LB medium containing 0.2 % maltose and 10 mM $MgSO_4$ and grow overnight at 30 °C to an OD_{600} 0.7.
2. Spin the cells gently in Falcon tubes, remove the supernatant, and resuspend cells in 10 mM $MgSO_4$ to OD_{600} 0.5.
3. Add 5×10^4 pfu of the library to 0.6 mL cells. Mix well and incubate at 37 °C for 15 min without shaking.
4. Add 15 μL of 1 M IPTG and 10 mL LB top agar, mix well, and pour the mixture onto a prewarmed LB-tetracycline agar plate (Ø 13.5 cm) (*see* **Note 12**). Invert the plates and incubate at 37 °C overnight.
5. Number nitrocellulose filters. Soak them in TBS until they are completely wet. Do not use nitrocellulose that does not wet properly (*see* **Note 13**). Place filters on blotting paper to drain residual fluid. Take plates out of the incubator and place filters carefully onto the top agar starting at an edge. Incubate plates at 37 °C over night.
6. Chill plates at 4 °C for at least 1 h. Mark the orientation of the filters relative to the agar plate before removing them (e.g., by piercing a needle through filter and plate and generating holes marking corners of a non-equilateral triangle).
7. Remove filters carefully from plates, immerse them in a large container with 1× TBST (*see* **Note 13**). Shake them vigorously on a platform shaker for 30–60 min.
8. Remove any remaining top agar by lightly rubbing filters submerged in 1× TBST with gloved fingers.
9. Wash filters for further 15 min under vigorous agitation in a large container with fresh 1× TBST buffer.
10. Transfer each filter into an individual Petri dish. Immerse them in blocking buffer for 1 h at room temperature with gentle agitation.
11. Wash filters 3× 10 min in a large container with 1× TBS. Make sure that precipitates of milk powder are rubbed off with gloved fingers.
12. Transfer each filter again into an individual Petri dish. Add preabsorbed serum prepared as described in Subheading 3.2. Incubate on the shaker at room temperature overnight.
13. Collect the preabsorbed patient serum from all plates, pool it in a glass bottle and store it at 4 °C (*see* **Note 14**).
14. Wash filters 3× 10 min in a large container with 1× TBS.
15. Transfer each filter again into an individual Petri dish. Add goat anti-Human IgG labeled with alkaline phosphatase (1: 2,500 dilution) in 1× TBS with 1 % dry milk. Incubate for 1.5 h on the shaker at room temperature.

16. Wash filters 3× 10 min in a large container with 1× TBS.
17. Prepare staining solution by adding nitroblue-tetrazolium-chloride (NBT) to a final concentration of 0.3 mg/mL and 5-bromo-4-chloro-3-indol-phosphate-toluidine (BCIP) to a final concentration of 0.15 mg/mL to color development solution under continuous stirring. Filter this staining solution to avoid precipitates.
18. Transfer each filter into an individual Petri dish and immerse them in staining solution. Place in the dark until positive reactions are clearly visible. Positive clones can be identified based on the staining of the circumference of phage plaques.
19. Identify positive signals and determine the respective plaque on the agar plate using needle holes for orientation.
20. Pick positive plaques with 1–2 adjacent negative plaques (*see* **Note 15**) using a pipette tip and place each agar plug in a separate tube containing 1 mL of SM and 20 µL of chloroform. Vortex vigorously and incubate at 4 °C overnight.

3.4 Monoclonalization of Phages

1. Divide the LB-tetracycline agar plate in a large Petri dish (Ø 13.5 cm) in four quarters by cutting out a cross shape with a scalpel.
2. Dilute the primary stock of oligoclonal phages at 1:500 and 1:100 in SM buffer, add 2 µL from these dilutions and IPTG to 2–3 mL Top agar. Pour each of these dilutions on one quarter. Thus two different titers of two different oligoclonal phage stocks can be assayed on one 13.5 cm Petri plate (*see* **Note 16**).
3. Repeat **steps 5–19** as described in Subheading 3.3.
4. Pick two monoclonal phages for each clone confirmed as positive. Transfer them into 0.5 mL SM with 20 µL chloroform for *in vivo* excision.

3.5 In Vivo Excision of the Phagemid from the Lambda Phage Vector

The pBluescript phagemid containing the cloned insert can be easily excised from the Uni-ZAP XR vector by simultaneously infecting *E. coli* with both the lambda vector and the fl bacteriophage.

1. Streak the host strain *E. coli* XL1-Blue MRF' and *E. coli* XL1-Blue SOLR onto an LB agar plate containing tetracycline/kanamycin and incubate the plate overnight at 37 °C.
2. Grow overnight cultures of XL1-Blue MRF' and SOLR cells in LB broth, supplemented with 0.2 % (w/v) maltose and 10 mM $MgSO_4$, at 30 °C to an OD_{600} of 0.7–0.9. Spin the cells down and resuspend in 10 mM $MgSO_4$ for an OD_{600} of 1.0.
3. Pipette 200 µL of XL1-Blue MRF' cells at an OD_{600} of 1.0, 250 µL of monoclonal secondary positive phage stock (containing $>1 \times 10^5$ phage particles), 1 µL of the ExAssist helper phage ($>1 \times 10^6$ pfu/µL) into a Falcon 2059 polypropylene tube. Incubate at 37 °C for 15 min.

3. Add 3 mL of LB broth and incubate for 2.5–3 h at 37 °C.
4. Heat the tube at 65–70 °C for 20 min to inactivate the ExAssist helper phage, spin the tube at 1,000 × *g* for 15 min.
5. Transfer supernatant containing the excised pBluescript phagemid packaged as filamentous phage particles to a sterile Falcon tube. Store at 4 °C.
6. Add 10 and 100 µL of the phage supernatant each to 200 µL of freshly grown SOLR cells from **step 2** (OD_{600} 1.0) to 1.5 mL microcentrifuge tube. Incubate the tube at 37 °C for 15 min.
7. Streak 10–20 µL of the cell mixture from microcentrifuge tube on LB-ampicillin agar plates (50 µg/mL) and incubate the plates overnight at 37 °C. Colonies appearing on the plate contain the pBluescript double-stranded phagemid with the cloned DNA insert and can be subjected to plasmid purification as a starting point for further analysis, for example, DNA sequence analysis (*see* **Note 17**).

3.6 Differential Serology

To determine which of the SEREX-defined clones is disease-associated seroassays with panels of allogeneic sera from patients with the respective disease as well as appropriate control groups have to be conducted (*see* **Note 18**).

3.6.1 Conventional Plaque Lift Assay

1. Carefully divide agar-tetracycline plates in 15 cm Petri dishes into quarters, without detaching them from the dish.
2. Mix monoclonal phages representing individual antigens of interest 5:1 with wild type lambda ZAP phage as a control. Infect 0.6 mL *E. coli* XL1-Blue MRF' host strain ($OD_{600} = 0.5$) in the presence of 0.8 mM IPTG with 2×10^3 pfu.
3. Add 2–3 mL LB top agar and pour each individual phage mixture on one quarter of the prewarmed LB agar plates for overnight incubation. Thus four different phage clones are ready to blot on one nitrocellulose filter and can be processed in parallel.
4. Proceed as described in Subheading 3.3, **steps 5–20**. Each nitrocellulose filter representing four different autoantigens is incubated with preabsorbed serum from a different donor. Typically up to 20 sera from patients plus the same number of sera from healthy controls can be easily handled.

3.6.2 SeroGRID

To allow for higher throughput multiplex analysis of a panel of antigens with a larger panel of sera SeroGRID may be applied (13).

1. Prepare a LB-tetracycline plate in a 24.5 cm × 24.5 cm × 2.5 cm square Petri dish.
2. Insert 22 plastic spacers into the agar creating 23 strip-type separated spaces 1 cm × 24 cm in size.

3. Mix each individual phage clone with wild type phage as internal negative control (ratio 4:1). Infect *E. coli* XL1-Blue MRF with 6×10^2 pfu of this mixture. Add 3 mL of top agar and IPTG and pour into the prepared slots on the plate. Thus, you may feature 21 different phage clones you want to test, a wild type phage as a negative control and as a positive control for quality of patients' sera a phage containing for example PINCH (GenBank accession no. U09284), an adapter protein of LIM family for signal transduction in the integrin and growth factor pathway, which reacts with most human sera.
4. Incubate overnight, remove spacers, and place a 24 cm × 23 cm filters on the plate.
5. Wash and block as described in **steps 7–11** of Subheading 3.3.
6. Cut the filters in 1 cm × 23 cm sized strips. Direction of cuts is in a right angle as compared to the position of spacers. Thus, each strip features 1 cm × 1 cm squares each with a different phage clone derived autoantigen.
7. Each strip is placed in an individual segment of the incubation chamber and incubated with a different patient serum overnight.
8. Prepare stained filters by following **steps 13–20** of Subheading 3.3.

4 Notes

1. Protocols related to the construction of the lambda phage expression library refer to the lambda ZAPII-based system (Uni-ZAP XR cDNA library synthesis kit, Stratagene), which we prefer, because lytic plaques are large and signals clearly distinguishable. We and others have also used the following expression systems for SEREX: Lambda Uni-ZAP (Stratagene), Lambda ZAP Express (Stratagene), and Lambda TRIPLX vector (Clontech).
2. High-quality mRNA as starting material is of utmost importance. We propose to use the guanidium isothiocyanate–phenol–chloroform extraction method [5, 6]. For subsequent purification of poly(A) RNA, we use the messenger RNA Isolation Kit (Stratagene) exactly as suggested by the manufacturer. All precautions should be taken to avoid contamination with RNAses.
3. It is crucial for efficient ligation of the cDNA-fragments with the phage vector arms to completely eliminate residual adaptors present in great molar excess. We recommend the use of Sizesep separation columns. In addition to removal of adaptors, these columns allow size selection of long cDNA-fragments.

4. Quantification of yield: Take 1 µL from each tube and add it to 4 µL H_2O. Transfer 0.5 µL of these into new tubes with 4.5 µL H_2O (1:50 dilution). Dot 1 µL from both dilutions of each fraction onto an agarose/ethidium bromide gel. Dot increasing dilutions of a DNA marker of known concentration. Compare on a UV transilluminator. The amount of cDNA to be used for ligation should be such that insert and vector are in equal molar ratios to prevent multiple inserts. Maximum volume of resuspended cDNA to be introduced into the ligation reaction is 2.5 µL.
5. For packaging of methylated DNA commercial kits of comparable convenience and efficacy are also available from different manufacturers (Gigapack Gold, Stratagene; MaxPlax, Invitrogen).
6. Messenger RNA is primed in the first-strand synthesis with an *Xho*I site containing linker-primer and is reverse transcribed using 5-methyl dCTP. Hemimethylated cDNA is generated to protect from digestion of the cDNA with *Xho*I used for adaptor digestion. Hemimethylated DNA introduced in a bacterial strain would be efficiently digested by the *mcrA* and *mcrB* restriction systems. It is of importance to have a first pass of the library through a $McrA^-$ $McrB^-$ host (e.g., XL1-Blue MRF'), otherwise you will lose your library. After that, the library is no longer hemimethylated and can be grown on $McrA^+$ $McrB^+$ strains (e.g., XL1-Blue).
7. The titer of a primary lambda phage library may drop with time resulting in the loss of rare cDNA species. One round of library amplification is recommended to obtain a large and stable high-titer stock of the library. We prefer amplification on agar plates rather than in liquid medium to reduce the risk of major skewing in the representation of individual clones. You may want to conduct an initial screening round directly after the library has been packaged prior to amplification to gather the diversity of transcript species. If primary phage count is $>4 \times 10^6$, immunoscreening may be conducted without library amplification. In that case make sure that XL1-Blue MRF' cells are used as host strain in each screening round (*see* **Note 6**).
8. Number of clones, rate of recombinants, and median and maximal insert size of inserts describe the quality of an expression library. The number of primary plaques indicates the representation of individual transcripts in the library. cDNA clones in lambda ZAP libraries are expressed as fusion proteins containing N-terminal beta-galactosidase fragments. As only one-third of directionally cloned cDNA are in the appropriate frame and some cDNA-fragments are cloned with a leading 5′-UTR possibly introducing premature stop codons, libraries to be used for SEREX should contain $>2 \times 10^6$ primary plaques.

The median size of human mRNA derived inserts should be >1 kb and the library should also contain larger inserts with sizes >4–5 kb. To estimate the average and maximum fragment size, amplify $>10^5$ phages as template by PCR using oligonucleotides binding to insert flanking regions of the vector backbone and analyze size distribution on a cDNA agarose gel. The rate of recombinants (white plaque number/total plaque number) should be >95 %. Lower rates may indicate inefficient ligation of the cDNA-fragments (e.g., mismatched amounts of vector and insert, defective cDNA-fragment or vector ends, contamination with adaptors). We recommend to perform sequencing of randomly picked clones to ensure the correctness of inserts (correct direction, presence of adaptor sequences at 5′ end, presence of poly(A) tail at the 3′ end) if libraries with lower rates of recombinants will be used for screening.

9. The quality of the preabsorbed patient serum is critical. Our protocol uses not only mechanically disrupted bacteria which deplete antibodies against intracellular components and cell wall elements of bacteria. We also preabsorb against bacteria lysed by phages, to address reactivities against phage and bacterial proteins induced secondary to infection (e.g., shock proteins). Many human sera can be used after having been subjected to one round of the described tripartite preabsorption process. If this does not reduce unspecific cross reactivity sufficiently, repeat parts of the process. If mainly the bacterial lawn in between plaques is contributing to the background, repeat steps of Subheading 3.2.1. If the plaque circumferences are stained unspecifically, repeat steps of Subheadings 3.2.2 or 3.2.3.
10. Lysate-loaded resins can be regenerated and reused. For regeneration wash 3× 10 min with 30 mL 50 mM Tris–HCl pH 3. Subsequently, wash 2× 10 min with 30 mL TBS and 10 min with 30 mL TBS and 0.01 % NaN_3 and Thimerosal. Store the resin in TBS and 0.01 % NaN_3 and Thimerosal at 4 °C. Maximal shelf life of regenerated resin is 3 months. The binding capacity is sufficient for processing of 6 mL undiluted serum in total (equivalent to 60 mL 1:10 pre-diluted serum).
11. If libraries are constructed from tissue with a considerable fraction of B-lymphocytes, fragments of human immunoglobulins will be expressed as well. Immunoscreening of such libraries will yield false-positive signals as the secondary anti-human-IgG antibody will not only detect patient serum derived IgG bound to plaque lifted autoantigens but also to IgG fragments recombinantly expressed as part of the library. Depending on composition of the tissue sample, up to 90 % of all identified signals may be such false positives. Funneling all these clones through the entire process and late disclosure of their

IgG origin on the level of sequencing to eventually discard them is highly inefficient. One option is to deplete immunoglobulin sequences on the level of library construction using subtractive cDNA library approaches [6]. Another option is to apply a modified primary screening as we described previously [7]. Briefly, the blocked filters are preincubated first with enzyme-conjugated anti-human IgG and stained. Signals are marked with a pencil. Subsequently, these membranes go through the standard process starting with incubation with the autologous serum, secondary antibody and thereafter staining. Only those signals which have not been marked after the first staining are harvested. If contamination is modest this pretesting by direct incubation with the anti-IgG antibody may be conducted on the level of secondary immunoscreening after individual phages have been picked and oligoclonalized.

12. Do not use moist plates. Prepare agar plates at least 24 h prior to use and preincubate at 37 °C for 1 h. Wipe dry with a paper tissue before pouring the top agar.
13. Use care and gloves when handling nitrocellulose filters. Do not allow filters to dry after plaque lift. Since inter-lot differences of nitrocellulose may affect the quality of the assay, we prefer to test batches prior to use. If top agar and plaques stick to the filters and are removed together with it, agar traces may hamper staining of positive clones. Top agar has to be removed without scratching. Washes of filters have to be thorough. We prefer to process each filter separately blotted side facing upward in individual Petri dishes for the critical steps of blocking, antibody incubation and staining to ensure adequate exposure. These Petri dishes do not need to be sterile and can be reused.
14. Patient sera can be used several times (up to 15 consecutive rounds of screening for a total of up to 300 plates). Reuse further reduces background due to absorption of nonspecific components rather than resulting in diminshed activity. This “maturation” in consecutive screening rounds increases sensitivity and specificity. Weak positive reactivities are detected in later rounds of the immunoscreening. In between screenings, sera should be stored at 4 °C and Thimerosal or NaN_3 should be added each week to prevent bacterial contamination.
15. Deliberately pick negative plaque as internal controls for the secondary round of immunoscreening. The objective of the secondary round is to confirm positivity and to monoclonalize positive phages
16. Ensure that at least one of the two titers for each clone has resulted in subconfluent phage plaques to allow safe monoclonalization. If the entire lawn is lytic with both dilutions this

clone should be repeated in the next round at a higher dilution. Moreover, ensure that you confirm positivity of each phage stock by discerning clearly positive and negative plaques in these oligoclonal stocks.

17. Obvious next steps are DNA sequencing and BLAST searches to determine identity of clones, determination of tissue distribution of the transcript and the encoded protein, determination of cell biological function and disease relevance, etc.
18. A two-step strategy is recommended for differential serology. First, a small-scale serological study based on the plaque lift assay is conducted to define those antigens without reactivity with control sera of healthy individuals but exhibiting reactivity with tumor sera. Subsequently, those antigens, which appear to be cancer related, are subjected to ELISA based on purified recombinant protein for further confirmation with a large panel of sera [8]. Since this process is labor-intensive, several modified protocols have been established based, for example, on lysate ELISA, arrays, etc. [9–12], including the SEROGRID approach described here [13].

References

1. Sahin U, Türeci O, Schmitt H, Cochlovius B, Johannes T, Schmits R, Stenner F, Luo G, Schobert I, Pfreundschuh M (1995) Human neoplasms elicit multiple specific immune responses in the autologous host. Proc Natl Acad Sci U S A 92:11810–11813
2. Rauch J, Gires O (2008) SEREX, Proteomex, AMIDA, and beyond: serological screening technologies for target identification. Proteomics Clin Appl 2:355–371
3. Krebs P, Kurrer M, Sahin U, Türeci O, Ludewig B (2003) Autoimmunity seen through the SEREX-scope. Autoimmun Rev 2(6):339–345
4. Old LJ, Chen YT (1998) New paths in human cancer serology. J Exp Med 187:1163–1167
5. Chomczynski P, Sacchi N (1987) Anal Biochem 162:156–159
6. Diatchenko L, Lukyanov S, Lau YF, Siebert PD (1999) Suppression subtractive hybridisation: a versatile method for identifying differentially expressed genes. Methods Enzymol 303:349–380
7. Türeci Ö, Schmitt H, Fadle N, Pfreundschuh M, Sahin U (1997) Molecular definition of a novel human galectin which is immunogenic in patients with Hodgkin's disease. J Biol Chem 272:6416–6422
8. Stockert E, Jäger E, Chen YT, Scanlan MJ, Gout I, Karbach J, Arand M, Knuth A, Old LJ (1998) A survey of the humoral immune response of cancer patients to a panel of human tumor antigens. J Exp Med 187:1349–1354
9. Lagarkova MA, Koroleva EP, Kuprash DV, Boitchenko VE, Kashkarova UA, Nedospasov SA, Shebzukhov YV (2003) Evaluation of humoral response to tumor antigens using recombinant expression-based serological mini-arrays (SMARTA). Immunol Lett 85:71–74
10. Scanlan MJ, Welt S, Gordon CM, Chen YT, Gure AO, Stockert E, Jungbluth AA, Ritter G, Jager D, Jager E, Knuth A, Old LJ (2002) Cancer-related serological recognition of human colon cancer: identification of potential diagnostic and immunotherapeutic targets. Cancer Res 62:4041–4047
11. Türeci Ö, Luxemburger U, Heinen H, Mack U, Ukena D, Sybrecht GW, Huber C, Sahin U (2006) CrELISA: a fast and robust enzyme linked immuno sorbent assay bypassing the need for purification of recombinant protein. Cancer Lett 236:64–71
12. Gnjatic S, Wheeler C, Ebner M, Ritter E, Murray A, Altorki NK, Ferrara CA, Hepburne-Scott H, Joyce S, Koopman J, McAndrew MB,

Workman N, Ritter G, Fallon R, Old LJ (2009) Seromic analysis of antibody responses in non-small cell lung cancer patients and healthy donors using conformational protein arrays. J Immunol Methods 341:50–58

13. Krause P, Türeci Ö, Micke P, Buhl R, Huber C, Sahin U (2003) SeroGRID: an improved method for the rapid selection of antigens with disease related immunogenicity. J Immunol Methods 283:261–267

Chapter 4

Antigen Discovery Using Whole-Genome Phage Display Libraries

Elisa Beghetto and Nicola Gargano

Abstract

In the last two decades phage display technology has been used for investigating complex biological processes and isolating molecules of practical value in several applications. Bacteriophage lambda, representing a classical cloning and expression system, has also been exploited for generating display libraries of small peptides and protein domains. More recently, large cDNA and whole-genome lambda-display libraries of human pathogens have been generated for the discovery of new antigens for biomedical applications. Here, we describe the construction of a whole-genome library of a common pathogen—*Streptococcus pneumoniae*—and the use of this library for the molecular dissection of the human B-cell response against bacterial infection and colonization.

Key words Bacteriophage lambda, Phage display library, Genomic DNA display, *Streptococcus pneumoniae*, Antigen discovery

1 Introduction

Lambda bacteriophage has been demonstrated to be the system of choice to display complex complementary deoxyribonucleic acid (cDNA) libraries. Libraries with a complexity of 10^7–10^8 independent clones can be easily constructed using efficient *in vitro* packaging systems.

The head decoration protein gpD (11.4 kDa, 405–420 copies per capsid) represents the ideal fusion partner, as it has been demonstrated that gpD can tolerate both amino- and carboxyl-terminal insertions of peptides and protein domains accessible for ligand interaction without interfering in phage replication and assembly of infective bacteriophages [1, 2].

The most commonly used vector for displaying large genomic and cDNA repertoires in lambda bacteriophage, named λKM4, is based on a "two-gene system", which is characterized by the presence of two copies of the *D* gene; a genomic copy harboring an

Kelly M. Fulton and Susan M. Twine (eds.), *Immunoproteomics: Methods and Protocols*, Methods in Molecular Biology, vol. 1061, DOI 10.1007/978-1-62703-589-7_4,

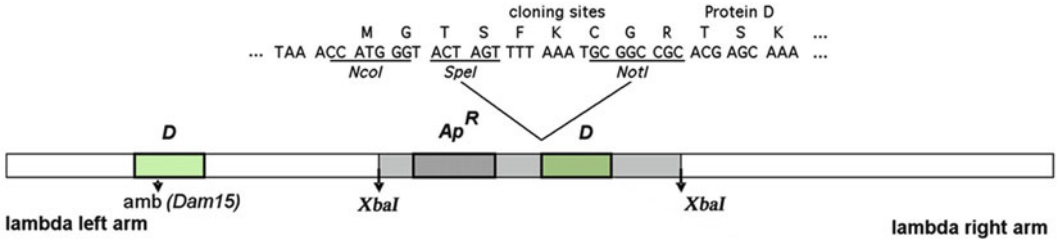


Fig. 1 Genetic map of λKM4 vector with sequence of the cloning sites. *Ap*^{R} ampicillin resistance, *amb amber* mutation

amber mutation (a stop codon causing the premature termination of gpD synthesis) at the 5′ end followed by an additional copy, controlled by an inducible promoter, containing a cloning cassette at the 5′ end (*see* Fig. 1) [3, 4]. Accordingly, bacteriophages grown on a suppressor bacterial strain (translating the amber mutation in a phenylalanine residue) display an array of wild-type gpD (encoded by the genomic copy) and recombinant gpD fused to foreign polypeptides (encoded by the additional copy). Such an array of wild-type proteins and recombinant fusion products is necessary to avoid the full display of recombinant proteins (all of the capsidic gpD copies are fused to the exogenous polypeptide) which can cause steric hindrance and disassembly of the viral capsid. The vector also harbors an antibiotic resistance gene (i.e., β-lactamase), allowing the growth of recombinant clones as antibiotic-resistant lysogenic colonies.

Most of the selection schemes for screening lambda libraries utilize a protocol which has been originally developed for filamentous phage applications [5]. Briefly, the target molecules linked to solid-phase matrices (i.e., polypropylene plates or tubes, sepharose matrices or magnetic beads) are challenged with a suspension of phage particles representing the library. After incubating the mixture for several hours, phage-ligand complexes are used to infect freshly cultured bacteria and the phage progeny released by host cells is directly screened for isolating individual clones or further processed for another round of affinity selection. In order to eliminate most of the low-affinity binders and unspecific interactions, the selection procedure generally includes two to three rounds of affinity selection.

In the last 10 years, the lambda-display approach has been successfully employed to identify and characterize antigens involved in the human immune response to infections caused by several pathogens, including protozoa, viruses, and bacteria [3, 6–11]. This was achieved by selecting the antigenic regions harboring B-cell epitopes via a direct challenge of pathogen-derived protein fragment display libraries with the whole antibody repertoire of infected individuals (i.e., antibodies present in plasma or serum).

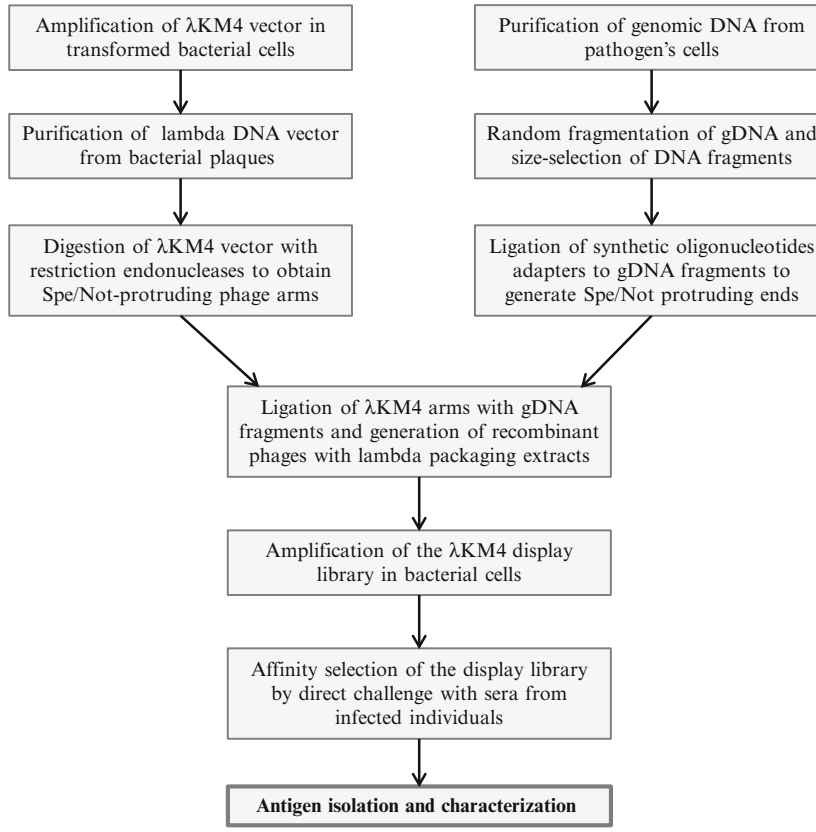


Fig. 2 Flow chart of library construction and antigen selection

In this chapter, the experimental design for investigating the human antibody response to *Streptococcus pneumoniae* (pneumococcus), a ubiquitous gram-positive bacterium causing invasive diseases such as pneumonia, sepsis, and meningitis [12, 13], will be described. In particular, the construction of a *S. pneumoniae* lambda-display library of genomic DNA (gDNA) fragments and the challenge of the library with sera from patients hospitalized for *S. pneumoniae* diseases will be detailed (*see* Fig. 2), followed by a preliminary molecular characterization of the selected recombinant clones, which represent B-cell antigens or antigenic regions containing B-cell epitopes recognized by antibodies of infected individuals.

2 Materials

Prepare all solutions using ultrapure water (prepared by purifying deionized water to attain a sensitivity of 18 MΩ cm at 25 °C) and analytical grade reagents. Prepare and store all reagents at room temperature (unless indicated otherwise). Diligently follow all waste disposal regulations when disposing waste materials.

2.1 Bacterial Strains, Vectors, and Human Sera

1. *Escherichia coli* BB4 and LE392 strains are commercially available (i.e., Invitrogen). Any bacterial strain susceptible to lambda bacteriophage infection is suitable for the construction of display libraries.
2. λKM4 vector is not commercially available but it can be provided by the authors of the present article for noncommercial purposes. Alternatively, it could be provided by other authors [4].
3. Human sera or plasma from *S. pneumoniae*-infected individuals collected from hospitals, blood banks, etc. (for the data presented here, sera have been provided by the Clinic of Infectious Diseases of Siena University, Italy).

2.2 Buffers and Reagents

1. Phosphate-buffered saline (PBS) (Gibco, Life Technologies).
2. 3 M Sodium acetate (CH_3COONa) stock solution: dissolve 408.3 g of CH_3COONa in 1 L of H_2O. Adjust pH to 5.2 by adding acetic acid. Autoclave and store the stock solution at room temperature up to 6 months.
3. Isopropyl alcohol.
4. Ethyl alcohol.
5. 0.5 M Ethylenediaminetetraacetic acid (EDTA) stock solution: dissolve 186 g of EDTA in 1 L of distilled, deionized H_2O. Adjust pH to 8.0 by using NaOH.
6. Adenosine-5′-triphosphate (ATP) (Invitrogen, Life Technologies).
7. Bovine serum albumin (BSA).
8. Deoxyribonucleotide triphosphates (dNTPs) (Invitrogen, Life Technologies).
9. Agarose gel electrophoresis: all reagents/devices (1.2 % Precast Agarose Gels, electrophoresis buffer, running device) for standard gel electrophoresis are commercially available (e.g., Sigma Aldrich).
10. SM buffer: dissolve 5.8 g of NaCl and 2.0 g of $MgSO_4[7H_2O]$ in 50 mL of 1 M Tris–HCl (pH 7.5) and 5.0 mL of 2 % gelatine with H_2O to a final volume of 1 L. Autoclave and store the solution at room temperature for a maximum of 6 months.
11. 2 % Maltose stock solution: dissolve 2 g of maltose in 100 mL of H_2O.
12. 1 M Magnesium sulfate ($MgSO_4$) stock solution: dissolve 246.5 g of $MgSO_4$ in 1 L of H_2O. Store at room temperature up to 12 months.
13. 5 M Sodium chloride (NaCl) stock solution: dissolve 292.2 g of NaCl in 1 L of H_2O.
14. Poly-ethylene glycol 8000 (PEG-8000) (Sigma Aldrich).
15. 7 % Dimethylsulfoxide (DMSO) solution.
16. Chloroform.

2.3 Bacterial Media

1. Luria Broth (LB): dissolve 10 g of NaCl, 10 g of bacto-tryptone, and 5 g bacto-yeast extract in 1 L of H_2O. Adjust to pH 7.0 with 5 N NaOH. All components purchased from Sigma Aldrich. Autoclave and store the medium at room temperature for a maximum of 4 months.
2. NZY Broth: Dissolve 5 g of NaCl, 2 g of $MgSO_4[7H_2O]$, 5 g of yeast extract, and 10 g of casein hydrolysate in H_2O (final volume of 1 L). Adjust the pH to 7.5 with NaOH. All components purchased from Sigma Aldrich. Autoclave and store the medium at room temperature for a maximum of 4 months.
3. NZY Top-Agar and NZY Agar: NZY Broth supplemented with 0.7 % (w/v) or 1.5 % (w/v) of agar (Sigma Aldrich), respectively.
4. Tryptic Soy Broth (TSB) (Difco Laboratories).

2.4 Synthetic Oligonucleotides

The following oligonucleotides can be purchased from commercial suppliers (i.e., Sigma Aldrich custom orders) and stored in H_2O at −20 °C in small aliquots until use. It should be noted that such oligonucleotides are specific for cloning DNA inserts into λKM4 vector [6], independently from the source (i.e., gDNA or cDNA) and the pathogen of interest (i.e., *S. pneumoniae*).

K185 5′-CTAGTCGTGCTGGCCAGC-3′

K186 5′-GCTGGCCAGCACGA-3′

K187 5′-CTAGTCGTGCTGGCCAGCT-3′

K188 5′-AGCTGGCCAGCACGA-3′

K189 5′-CTAGTCGTGCTGGCCAGCTG-3′

K190 5′-CAGCTGGCCAGCACGA-3′

K191 5′-TCTGGTGGCGGTAGC-3′

K192 5′-GGCCGCTACCGCCACCAGA-3′

K193 5′-TTCTGGTGGCGGTAGC-3′

K194 5′-GGCCGCTACCGCCACCAGAA-3′

K195 5′-TTTCTGGTGGCGGTAGC-3′

K196 5′-GGCCGCTACCGCCACCAGAAA-3′

λfor: 5′-GGGCACTCGACCGGAATTATCG-3′

λrev: 5′-CTCTCATCCGCCAAAACA GCC-3′

2.5 Laboratory Instruments

1. Spectrophotomer, allowing the measurement in both ultraviolet (UV) and visible wavelength spectra, are available from many suppliers (i.e., Eppendorf).
2. UV transilluminator (Vilber Lourmat).
3. Automated microplate ELISA reader having absorbance detection modules/filters for 450 and 620 nm wavelengths (Bio-Rad).

4. Automated ELISA washer (Tecan).
5. Power supply and device for agarose gel electrophoresis (all apparatus for standard electrophoresis can be purchased from Sigma Aldrich).
6. Other equipment for molecular biology laboratory: incubators (i.e., cabinet and shaking incubators for growing bacteria on plates and in flasks, respectively), centrifuges (for tubes and bottles), thermomixer (Eppendorf), pipettes (Gilson), etc.

2.6 DNA Purification Kits and Lambda-Packaging Kits

1. Large-scale preparation of bacteriophage λKM4 vector can be obtained by using commercially available kits (i.e., Lambda DNA Purification Kit purchased from Agilent Technologies).
2. Wizard SV Gel and Polymerase Chain Reaction (PCR) Clean-Up System (Promega) for purification of PCR products and other enzyme-processed DNA fragments.
3. The packaging of lambda bacteriophages for the construction of display libraries can be performed by using commercially available packaging kits (i.e., Ready-To-Go Lambda-Packaging Kit from GE Healthcare).

2.7 Enzymes and Buffers

All of the following enzymes as well as the corresponding reaction buffers are purchased from New England Biolabs:

1. *Spe*I restriction endonuclease.
2. *Not*I restriction endonuclease.
3. DNaseI-(RNase-free).
4. T4 DNA ligase.
5. T4 Polynucleotide Kinase.
6. T4 DNA polymerase.

2.8 Biopanning Reagents and Buffers

1. Dynabeads Protein G (Invitrogen).
2. Fetal Calf Serum (FCS) (Gibco, Life Technologies).
3. Biopanning coating buffer: 0.1 M NaH_2PO_4/Na_2HPO_4, pH 8.
4. Biopanning washing buffer: 1 % (v/v) Triton X-100, 10 mM $MgSO_4$ in PBS.
5. Biopanning blocking buffer: 5 % (w/v) nonfat dry milk in PBS, 0.25 % Triton X-100, 10 mM $MgSO_4$.

2.9 Phage Enzyme-Linked Immunosorbent Assay (ELISA) Reagents, Antibodies, and Buffers

1. Maxisorp multiwells plates (NUNC).
2. Polyclonal anti-lambda antibodies (developed in rabbits). Please note that the antibodies are not commercially available and should be made using custom services provided by antibody developers (e.g., Pierce, Genescript, etc.) using PEG-purified λKM4 bacteriophages as the antigen.

3. Secondary antibody: goat anti-human IgG Horse Radish Peroxidase (HRP) conjugated antibodies (Sigma Aldrich).
4. Chromogenic substrate: tetramethylbenzidine (TMB) liquid substrate system (Sigma Aldrich).
5. Phage-ELISA coating buffer: 50 mM $aHCO_3$ pH 9.6.
6. Phage-ELISA washing buffer: 0.05 % (v/v) Tween 20 in PBS.
7. Phage-ELISA blocking buffer: 5 % (w/v) nonfat dry milk, 0.05 % (v/v) Tween 20 in PBS.
8. Stop solution: 2 M H_2SO_4.

2.10 Immunoscreening Reagents, Antibodies, and Buffers

1. Nitrocellulose filters (Schleicher & Schuell).
2. Secondary antibody: goat anti-human IgG Alkaline Phosphatase (AP) conjugated antibodies (Sigma Aldrich).
3. Chromogenic substrates: nitrobluetetrazolium (NBT) and 5-bromo-4-chloro-3-indolyl phosphate (BCIP) substrates (Sigma Aldrich).
4. Immunoscreening washing buffer: 0.05 % Tween 20 in PBS.
5. Immunoscreening blocking buffer: 5 % nonfat dry milk, 0.1 % Tween 20 in PBS.

3 Methods

3.1 Preparation of Lambda-Display Transfer Vector (λKM4)

1. Organize a medium/large-scale preparation of λKM4 vector (i.e., 50–100 μg of DNA) for cloning the exogenous DNA fragments into the bacteriophage genome and generating high-complexity display libraries; grow the bacteriophage in *E. coli* BB4 cells cultured in top-agar medium, thus allowing the formation of phage plaques onto the lawn of bacterial plates.
2. Purify λKM4 DNA from phage plaques using commercial kits (i.e., Lambda DNA Purification Kit from Agilent Technologies); strictly adhere to manufacturer's instruction for growing conditions (initial amounts of BB4 cells and lambda bacteriophage, temperature and length of incubation, etc.).
3. Check the integrity of λKM4 genome by running 1 μg of the purified DNA with agarose gel electrophoresis (agarose gel: 1.2 % w/v).
4. Incubate for 2 h at 37 °C 30 μg of λKM4 vector DNA with *Spe*I (250 Units) and *Not*I (250 Units) restriction endonucleases in a final volume of 300 μL (*see* **Note 1**).
5. Purify bacteriophage DNA by using standard phenol/chloroform extraction methods [14].
6. Transfer the purified DNA solution (300 μL) in a 1.5 mL conical microfuge tube, then add 90 μL of 3 M CH_3COONa

pH 5.2, 560 µL of isopropanol and 510 µL of H_2O. Precipitate the digested DNA by incubating the tube on ice for 20 min.

7. Centrifuge the tube at 14,000×*g* for 15 min at 4 °C in a benchtop microfuge. Remove the solution, wash the DNA pellet with ice-cold 70 % ethanol, and suspend the DNA in 50 µL of 10 mM Tris–HCl pH 8.5 (*see* **Note 2**).
8. Store the purified DNA at –20 °C in small aliquots until use.

3.2 Preparation of Synthetic Adapters

1. Phosphorylate oligonucleotides K185, K187, K189, K192, K194, and K196 by mixing 1 µmol of each oligonucleotide with 8 Units of T4 Polynucleotide Kinase (PNK), 2 µL of 10×-PNK buffer, 2 µL of ATP in a final volume of 20 µL. Incubate the mixture at 37 °C for 60 min and then at 65 °C for 30 min to stop the enzymatic reaction (*see* **Note 3**).
2. Anneal the phosphorylated/unphosphorylated oligonucleotide pairs by mixing an equal amount of each partner (i.e., 1 µmol) to form six different adapters as the following:

K185 + K186	K187 + K188	K189 + K190
K191 + K192	K193 + K194	K195 + K196

Place the tubes containing the different mixtures in a water bath at 85 °C for 2 min and then leave them on the bench until room temperature is reached (i.e., 10–20 min).

3. Store the annealed adapters at –20 °C until use.

3.3 Preparation of S. pneumoniae DNA Fragments for Cloning into Lambda Vectors

1. Pick a single colony of *S. pneumoniae*, transfer the colony into 30 mL of TSB, and incubate the culture in a 5 % CO_2-enriched atmosphere at 37 °C for 12–16 h.
2. Purify the genomic DNA by using standard methods [14]. Alternatively, the bacterial DNA can be bought from commercial providers (i.e., the American Type Culture Collection; ATCC).
3. Determine the DNA concentration by reading its absorbance at 260 nm with a spectrophotomer.
4. Digest 5–10 µg of the genomic DNA with 1 Unit of Deoxyribonuclease I (DNase I)-RNase-free. Incubate the genomic DNA and the DNase enzyme for 20 min at 15 °C in a final volume of 100 µL of DNase buffer. Add EDTA at a final concentration of 50 mM to stop the reaction and put the tube on ice until next step (*see* **Note 4**).
5. Purify digested genomic DNA by using the Wizard SV Gel and PCR Clean-Up System (Promega). Elute the DNA in 50 µL of H_2O and determine the DNA concentration by reading the absorbance at 260 nm with a spectrophotomer.

6. Recess 3′-protruding termini of the digested genomic DNA fragments by using the T4 DNA polymerase. To this aim, mix 2.5 μg of DNA fragments with 9 Units of T4 polymerase, 0.5 μg of Bovine Serum Albumin (BSA) and 0.4 mM dNTPs in a final volume of 50 μL of T4 polymerase buffer. Incubate the mixture for 20 min at 12 °C and then for 20 min at 75 °C to stop the reaction.
7. Purify the blunt-end DNA fragments by phenol extraction and ethanol precipitation using standard methods [14]. Finally, suspend the DNA pellet in H_2O at a final concentration of 100–200 ng/μL and store the DNA at −20 °C until next step.
8. Incubate the purified blunt-end genomic DNA fragments with 20-fold molar excess of the six synthetic adapters bringing *Spe*I or *Not*I restriction sites (*see* Subheading 3.2) in the presence of the T4 DNA ligase. To this purpose, mix 1 μg of DNA fragments, 30 pmol of each different adapter, 5 μL of 10× T4 ligase buffer, 50 μg of BSA, and 3,000 Units of the T4 ligase enzyme in a final volume of 50 μL. Perform the ligation reaction overnight at 16 °C, then inactivate the T4 enzyme for 20 min at 65 °C.
9. Purify the DNA fragments by using the Wizard SV Gel and PCR Clean-Up System, as per the manufacturer's instructions.
10. Proceed with a further purification step of the DNA fragments ligated to the library adapters by using a 2 % agarose gel electrophoresis. Under UV lighting, cut out an agarose slice containing the DNA fragments with molecular lengths comprised between 300 and 1,000 bp (*see* **Note 5**). Finally, purify the size-selected DNA-adapter fragments from the agarose slice by using the Wizard SV Gel and PCR Clean-Up System, as per the manufacturer's instructions.

3.4 Lambda-Display Library Construction

1. Clone the size-selected genomic DNA-adapter fragments into *Spe*I–*Not*I digested λKM4 vector (*see* Subheading 3.1). To this aim, mix 0.5 μg of vector DNA, 10–20 ng of DNA fragments and 200 U of T4 ligase enzyme in a final volume of 10 μL. Perform the ligase reaction overnight at 16 °C, then inactivate the T4 enzyme for 20 min at 65 °C.
2. Incubate the whole ligase mixture with commercial lambda-packaging extracts. Following manufacturer's instructions calculate the number of extracts which are necessary for packaging the whole ligase reaction (10 μL), then remove the packaging extracts stored in the freezer (−80 °C) and place them on dry ice. Quickly thaw the packaging extracts until the content of the tubes just begins to thaw, and then add the DNA immediately. Stir the tube with a pipet tip to mix well and incubate the tube for 2 h at 22 °C.

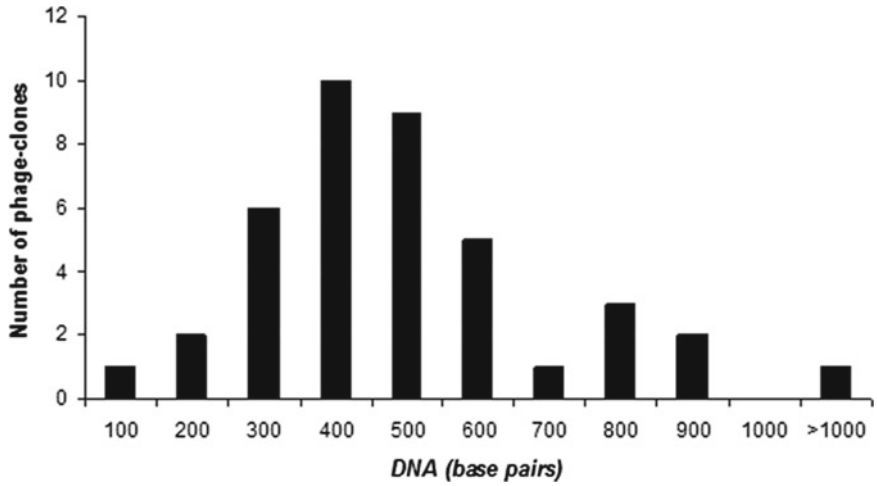


Fig. 3 Distribution of DNA fragments displayed on the surface of phage clones derived from a *S. pneumoniae* gDNA library, as obtained by amplifying DNA inserts from single phage plaques by PCR and analyzing the corresponding products by agarose gel electrophoresis

3. Add 500 µL of SM buffer to each tube and mix the content gently. Spin briefly to sediment the debris and store the supernatant containing the phages at 4 °C (supernatant may be stored at 4 °C for up to 1 month).
4. Pick one single colony of *E. coli* BB4 cells (alternatively the LE392 strain can be used) from the master plate and inoculate the colony in 100 mL of LB medium supplemented with 0.2 % maltose and 10 mM $MgSO_4$. Incubate the culture until the bacterial Optical Density (OD) reaches $OD_{600nm} = 1$; then centrifuge and suspend the bacterial pellet in 50 mL of SM buffer.
5. Incubate the bacterial cells (50 mL) with the whole phage supernatant (500 µL) containing the packaged phage particles (this procedure will allow an infection rate ≤1 phage per bacterial cell) for 20 min at room temperature.
6. Test the efficiency of the ligation/packaging/infection before plating the whole phage library and check also the percentage and the size of the DNA inserts cloned into lambda genome (*see* **Note 6** and Fig. 3).
7. Add 15 mL of top-Agar medium (pre-warmed to 42 °C) per 0.6 mL of infected cells and quickly adsorb the mixture to 15 cm NZY/agar-coated plates (a total of 80–90 plates are needed for each library).
8. Incubate the plates at 37 °C until bacterial plaques are clearly visible (i.e., 12–16 h).
9. Recover recombinant phages from bacterial plaques by adding 15 mL of SM buffer to each 15 cm plate and incubate the

plates for 4 h at room temperature in an orbital shaker. Transfer the SM buffer containing the eluted phages in conical tubes (50 mL) and remove cell debris by centrifuging the tubes at 3,300 × *g* for 20 min at 4 °C.

10. Transfer the supernatant into new tubes, then add NaCl to a final concentration of 1 M and incubate the tubes for 1 h on ice. After centrifugation at 3,300 × *g* for 20 min at 4 °C, collect the supernatant and add 10 % (v/v) of PEG-8000 to precipitate the phage particles.
11. After an overnight incubation at 4 °C, collect the bacteriophage particles by centrifugation at 3,300 × *g* at 4 °C for 30 min and finally suspend the pellet in 1/10 of the starting volume of SM buffer.
12. Repeat the precipitation step with 1 M NaCl and 10 % PEG-8000 incubating the tubes on ice for 1 h. Centrifuge the tubes at 3,300 × *g* for 30 min at 4 °C and then suspend the phage pellet in 1/20 of the starting volume of SM buffer.
13. Add DMSO at a final concentration of 7 % (v/v). Store the phage suspension in aliquots of 0.5–1 mL at −80 °C.

3.5 Biopanning of Lambda-Display Libraries with Human Sera

1. Pick one single colony of *E. coli* BB4 cells from the master plate and inoculate the colony in 20 mL of LB medium supplemented with 0.2 % maltose (w/v) and 10 mM $MgSO_4$. Grow bacterial cells at 37 °C until the bacterial density reaches an $OD_{600nm} = 1$. Then centrifuge and suspend the bacterial pellet in half the volume of SM buffer.
2. Take 100 μL of magnetic beads linked to Protein G (Dynabeads Protein G) for each affinity selection of the lambda library. Wash the beads twice with 1 mL of biopanning coating buffer, and then store the beads on ice in 0.1 mL of coating buffer. Add 10 μL of human serum (i.e., from *S. pneumoniae*-infected individuals) to the washed beads and incubate the beads suspension with rotation for 40 min at room temperature.
3. Wash the beads three times with 1 mL of biopanning washing buffer, then incubate the beads with 1 mL of biopanning blocking buffer for 1 h at room temperature.
4. Add 5×10^{10} plaque forming units (pfu) of recombinant lambda phages to the washed beads/serum suspension (final volume of 1 mL of blocking solution). Incubate the mixture for 3–4 h at room temperature with gentle rotation.
5. Wash the beads ten times with 1 mL of biopanning washing buffer and then add 1.2 mL of bacterial cells (BB4 freshly prepared as described above) to the washed beads/serum/phage mixture. Incubate the mixture for 20 min at room temperature.

6. Add 0.6 mL of the BB4/beads/serum/phage suspension to 15 mL of pre-warmed (42 °C) top-Agar medium, mix well and immediately adsorb the mixture onto 15 cm NZY-medium/agar-coated plates.
7. Incubate the plates overnight at 37 °C or until bacterial plaques are clearly visible (i.e. 12–16 h).
8. Recover recombinant phages from bacterial plaques by adding 15 mL of SM buffer to each 15 cm plate and incubate the plates for 4 h at room temperature in an orbital shaker. Transfer the SM buffer containing the eluted phages into conical tubes (50 mL) and remove cell debris by centrifuging the tubes at 3,300 × *g* for 20 min at 4 °C.
9. Transfer the supernatant into new tubes, then add NaCl to a final concentration of 1 M and incubate the tubes for 1 h on ice. After centrifugation at 3,300 × *g* for 20 min at 4 °C, collect the supernatant and add 10 % (v/v) of PEG-8000 to precipitate the phage particles.
10. After an overnight incubation at 4 °C, collect the bacteriophages by centrifugation at 3,300 × *g* for 30 min at 4 °C and finally suspend the pellet in 5 mL of SM buffer.
11. Store the purified phage suspension at 4 °C.
12. Repeat **steps 1–10** for the next round of biopanning (usually two to three rounds are necessary to sufficiently amplify the phage population selected by the human serum).

3.6 Phage-ELISA

1. Coat 96-well Maxisorb plates with either homemade or commercially available polyclonal anti-lambda antibodies, diluted to 0.1–1 μg/mL in ELISA coating buffer. Incubate overnight at 4 °C, then aspirate the solution and add 100 μL/well of ELISA blocking buffer. Incubate the plate for 1 h at 37 °C, then remove the blocking buffer and store the plates without washing at −20 °C until use.
2. Add the recombinant lambda phages selected from the display library to the coated plate (i.e., 10 μL/well of purified phage suspension from biopanning and 90 μL/well of ELISA blocking buffer) and incubate the plate overnight at 4 °C with gentle stirring.
3. Wash the plates using an automated ELISA-washer device with ELISA washing buffer.
4. Dilute 1–2 μL of human serum/plasma (i.e., serum from *S. pneumoniae*-infected individuals) with 10 μg/mL of FCS and 10 μg/mL of BB4 bacterial extract in ELISA blocking buffer (100 μL of final volume). Incubate the mixture for 30 min at room temperature.

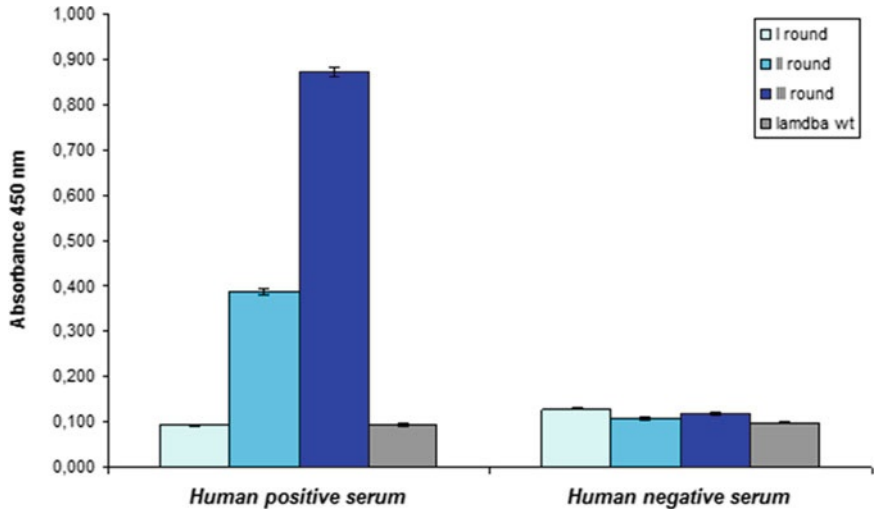


Fig. 4 Phage-ELISA reactivity of phage pools after multiple rounds of affinity selection (biopanning). Assayed are pools of phages after biopanning (I, II, and III rounds) and the λKM4 phage (lambda wild type (wt): negative control) with sera from *S. pneumoniae*-infected (positive serum) or -uninfected (negative serum) individuals

5. Add the previous mixtures to the wells and incubate the plate for 1 h at 37 °C with gentle stirring.
6. Wash the plate as in **step 3**.
7. Add 100 μL to each well of anti-human-IgG HRP-conjugated antibodies diluted 1:1,000 (1 μg/mL) in ELISA blocking solution and incubate the plate for 1 h at 37 °C with gentle stirring.
8. Wash the plates as in **step 3**.
9. Reveal the HRP-enzymatic activity by incubating the plates with the chromogenic substrate tetramethylbenzidine (TMB liquid substrate system) for 10–15 min at room temperature. Stop color development with 25 μL/well of 2 M H_2SO_4.
10. Read the plate by an automated ELISA reader as the difference between the absorbance at 450 and 620 nm (*see* Fig. 4).

3.7 Immunoscreening

1. Inoculate one single colony of *E. coli* BB4 cells in 20 mL of LB medium supplemented with 0.2 % (w/v) maltose and 10 mM $MgSO_4$, until the bacterial density reaches an $OD_{600nm} = 2$, then store the cells on ice.
2. Infect 0.2 mL of bacterial cells with tenfold serial dilutions of recombinant phages (i.e., from 10 pfu/mL up to 10,000 pfu/mL) derived from the affinity selection round (*see* Subheading 3.5) and leave the bacteria/phage mixture for 20 min at room temperature.

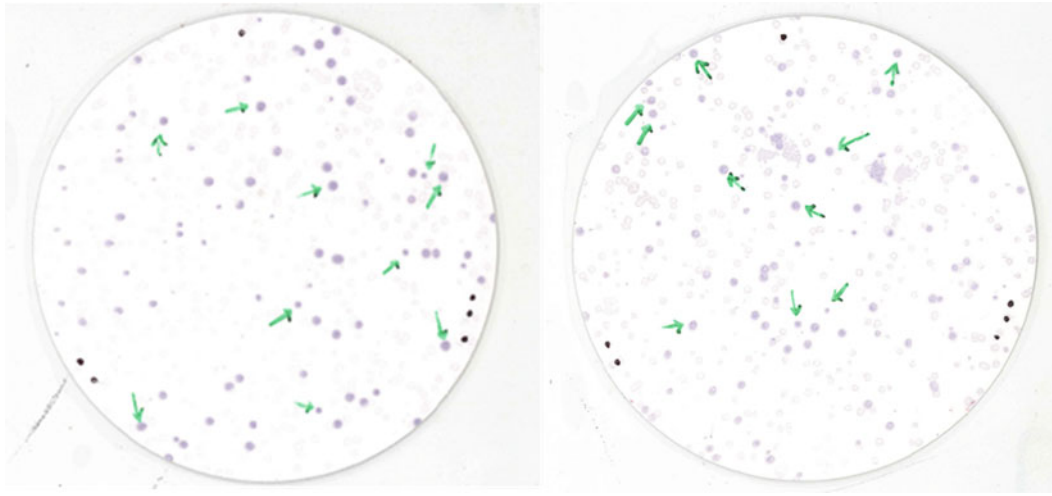


Fig. 5 Immunoscreening of phage pools selected from a *S. pneumoniae* gDNA display library with sera from infected individuals. Nitrocellulose filters are incubated with the same *S. pneumoniae* positive sera used for library selection, followed by incubation with anti-human IgG alkaline phosphatase-conjugated antibodies. Positive clones are finally revealed with alkaline phosphatase chromogenic substrates. *Arrows* indicate positive phage clones (*dark gray spots*)

3. Add 4–5 mL of pre-warmed top-Agar medium to the infected cells and adsorb the resulting mixture immediately onto 90 mm NZY-medium/agar-coated plates. Incubate the plates at 37 °C overnight or until phage plaques are clearly visible.
4. Put 90 mm nitrocellulose filters directly onto the surface of bacterial growth plates. Incubate the plates for 1 h at room temperature, then remove the nitrocellulose filters and store the plates (representing the master plates) at 4 °C.
5. Incubate the filters with 10 mL of immunoscreening blocking buffer for 1 h at room temperature.
6. Wash twice the filters with 10–15 mL of immunoscreening washing buffer.
7. Incubate the filters with human sera/plasma (i.e., serum from *S. pneumoniae*-infected individuals) diluted 1:100 in immunoscreening blocking buffer supplemented with 10 μL/mL of BB4 extract for 2 h at room temperature.
8. Wash the filters ten times with 10 mL of immunoscreening washing buffer.
9. Incubate the filters with anti-human IgG AP-conjugated antibodies, appropriately diluted (i.e., 1:1,000–1:10,000) in immunoscreening blocking buffer for 1 h at room temperature.
10. Reveal the AP-enzymatic activity by incubating the filters with the chromogenic substrates (NBT and BCIP substrates) (*see* Fig. 5).

3.8 Antigen Identification Through PCR and DNA Sequencing

1. Overlay the nitrocellulose filters to the master plates, then pick up positive phage plaques (detected by immunoscreening as described in Subheading 3.8) with a sterile tip and place the phages in a tube by pipetting up and down in 300 μL of SM buffer.
2. Add 50 μL of chloroform and leave the tubes for 15 min at 37 °C. After centrifugation at 11,000×*g* for 15 min at 4 °C, collect the supernatant and use 2 μL as a template for PCR reaction (final volume of 50 μL) to amplify the recombinant DNA fragments.
3. Use the primers λfor and λrev as specific oligonucleotides mapping, respectively, at the N- and C-terminus of the cloning sites in λKM4 vector.
4. Amplify the recombinant DNA fragments by using the following PCR conditions: 30 s at 94 °C, 30 s at 55 °C, and 30 s at 72 °C for 35 cycles.
5. Subject 5 μL of the PCR product to agarose gel electrophoresis (1.5 % of agarose gel), then purify the DNA fragments from the remaining PCR reaction mixture (45 μL) by using the Wizard SV Gel and PCR Clean-Up System kits. Store the purified PCR products at −20 °C.
6. Sequence the DNA fragments purified from PCR reactions by using standard methods [14].
7. According to the frame of the DNA fragment with respect to the fusion product with the *D* gene in λKM4, translate the corresponding nucleotide sequences into protein sequences and search for sequence similarities between the selected protein fragments and the corresponding sequences in the pathogen gene products (i.e., *S. pneumoniae*) using the available database (i.e., NCBI Blast; http://www.ncbi.nlm.nih.gov/BLAST/Blast.cgi?CMD=Web&PAGE_TYPE=BlastHome). This comparison allows the identification of protein fragments containing B-cell epitopes, as being selected by the challenge of the phage library with the antibody repertoire of infected individuals.
8. Use purified PCR products for digesting the DNA fragments with *Spe*I/*Not*I restriction endonuclease and cloning the digested fragments into bacterial expression vectors, thus allowing the production, purification, and further characterization of the selected antigen fragments as recombinant proteins, which can be used for diagnosis purposes and vaccine development.

4 Notes

1. When reaction with endonucleases is simultaneously performed by using two different enzymes, the supplier provides the appropriate reaction conditions (temperature, time, and buffer). In this specific case, *Spe*I/*Not*I digestion was carried out in

NEB buffer 3 supplemented with 0.1 mg/mL of bovine serum albumin.

2. During preparation and digestion of λKM4 vector, do not use vortex during pellet resuspension to avoid DNA breaking.
3. Phosphorylation of oligonucleotides representing the 5′-protruding end of adapters (K185, K187, K189, K192, K194, and K196) is necessary to promote ligation of DNA fragments into λKM4 vector, while the remaining nucleotides (K186, K188, K190, K191, K193, and K195) should not be phosphorylated to avoid ligation between blunt ends of adapters.
4. Before performing the large-scale digestion of the genomic DNA with DNaseI, digest a small amount of DNA (100–200 ng) and check the status of DNA fragmentation after different time intervals (i.e., 10, 15, and 20 min of incubation) by agarose gel electrophoresis. If necessary, adjust the length of incubation for the large-scale reaction.
5. When proceeding with the purification of DNA fragments from agarose gels, do not allow the samples to migrate too far since DNA fragments should be excised from the gel while avoiding picking up a large volume of agarose.
6. Before packaging the whole library, test a small amount of the ligation mixture (i.e., 0.2–0.5 μL) and analyze single phage clones by PCR and agarose gel electrophoresis (*see* Subheading 3.8) to determine the percentage of cloning and the size of DNA inserts.

References

1. Mikawa YG, Maruyama IN, Brenner S (1996) Surface display of proteins on bacteriophage lambda heads. J Mol Biol 262:21–30
2. Sternberg N, Hoess RH (1995) Display of peptides and proteins on the surface of bacteriophage lambda. Proc Natl Acad Sci USA 92:1609–1613
3. Beghetto E, Pucci A, Minenkova O, Spadoni A, Bruno L, Buffolano W, Soldati D, Felici F, Gargano N (2001) Identification of a human immunodominant B-cell epitope within the GRA1 antigen of *Toxoplasma gondii* by phage display of cDNA libraries. Int J Parasitol 31:1659–1668
4. Minenkova O, Pucci A, Pavoni E, De Tomassi A, Fortugno P, Gargano N, Cianfriglia M, Barca S, De Placido S, Martignetti A, Felici F, Cortese R, Monaci P (2003) Identification of tumor-associated antigens by screening phage-displayed human cDNA libraries with sera from tumor patients. Int J Cancer 106: 534–544
5. Parmley SF, Smith GP (1988) Antibody-selectable filamentous fd phage vectors: affinity purification of target genes. Gene 73: 305–318
6. Beghetto E, Spadoni A, Buffolano W, Del Pezzo M, Minenkova O, Pavoni E, Pucci A, Cortese R, Felici F, Gargano N (2003) Molecular dissection of the human B-cell response against *Toxoplasma gondii* infection by lambda display of cDNA libraries. Int J Parasitol 33:163–173
7. Santini C, Brennan D, Mennuni C, Hoess RH, Nicosia A, Cortese R, Luzzago A (1998) Efficient display of an HCV cDNA expression library as C-terminal fusion to the capsid protein D of bacteriophage lambda. J Mol Biol 282:125–135
8. Beghetto E, De Paolis F, Spadoni A, Del Porto P, Buffolano W, Gargano N (2008) Molecular dissection of the human B-cell response against cytomegalovirus infection by lambda-display. J Virol Methods 151:7–14

9. Beghetto E, Gargano N, Ricci S, Oggioni M, Garufi G, Peppoloni S, Pozzi G, Felici F (2006) Discovery of a novel *Streptococcus pneumoniae* antigen by screening a whole genome lambda-display library. FEMS Microbiol Lett 262:14–21
10. De Paolis F, Beghetto E, Spadoni A, Oggioni M, Montagnani F, Felici F, Gargano N (2007) Identification of a human immunodominant B-cell epitope within IgA1 protease of *Streptococcus pneumoniae*. BMC Microbiol 7:113
11. Beghetto E, De Paolis F, Montagnani F, Cellesi C, Gargano N (2009) Discovery of new *Mycoplasma pneumoniae* antigens by use of a whole-genome lambda-display library. Microbes Infect 11:66–73
12. Fedson DS, Anthony J, Scott G (1999) The burden of pneumococcal disease among adults in developed and developing countries: what is known and what is not known. Vaccine 17:S11–S18
13. Tuomanen EI (2000) Pathogenesis of pneumococcal inflammation: otitis media. Vaccine 19:S38–S40
14. Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning. A laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Chapter 5

Methods and Applications of Serological Proteome Analysis

Kelly M. Fulton, Shannon S. Martin, Lawrence Wolfraim, and Susan M. Twine

Abstract

The study of the humoral response to infectious diseases and chronic diseases, such as cancer, is important for many reasons, including understanding the host response to disease, identification of protective antigens, vaccine development, and discovery of biomarkers for early diagnosis. During the past decade, proteomic approaches, such as serological proteome analysis (SERPA), have been used to identify the repertoire of immunoreactive proteins in various diseases. In this chapter, we provide an outline of the SERPA approach, using the analysis of sera from mice vaccinated with a live attenuated tularemia vaccine as an example.

Key words SERPA, Serological proteome analysis, Western blotting, Immunoproteomics, Comparative immunoproteomics, *Francisella*, Tularemia, Live vaccine strain

1 Introduction

The humoral immune response to infection and chronic diseases has been studied for many years. For infectious diseases this has been accomplished by methods such as agglutination [1], ELISA [2], and 1D-Western blotting [3]. Early work was often unable to definitively identify the protein targets of the humoral immune response, due to limitations of protein identification technologies. More recently, 2D-polyacrylamide gel electrophoresis (2D-PAGE) combined with Western blotting has been exploited to separate protein antigens and elucidate those proteins reactive with immune sera. Combined with advances in genomics and mass spectrometry, this has allowed the facile identification of immunoreactive proteins. This has opened the door to rapid advances in determining protein antigens for vaccine development [4, 5], immune correlates, and biomarkers (e.g., *see* ref. 6) for disease diagnosis.

The antigen used in these studies is usually a cell lysate, or subproteome fraction (e.g., membrane) of target cell populations,

Kelly M. Fulton and Susan M. Twine (eds.), *Immunoproteomics: Methods and Protocols*, Methods in Molecular Biology, vol. 1061, DOI 10.1007/978-1-62703-589-7_5,

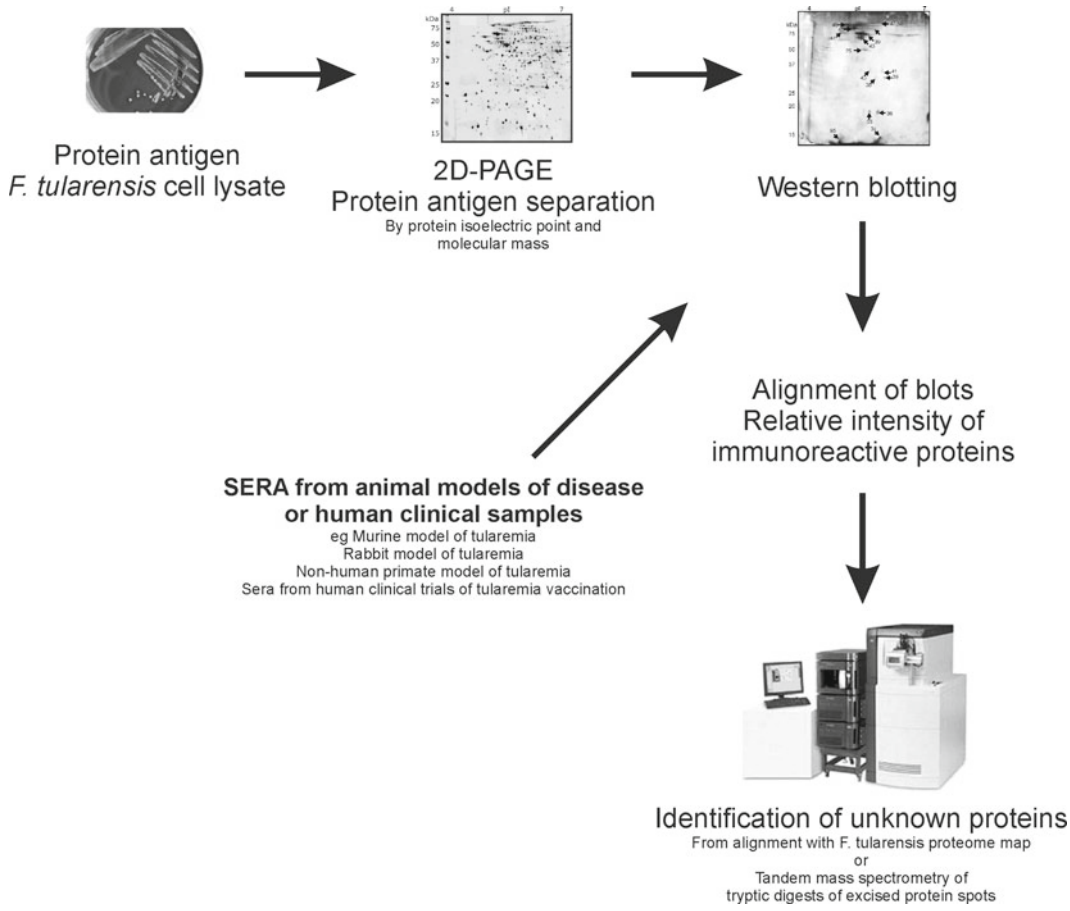


Fig. 1 Flow chart illustrating the approach used for SERPA. Protein antigens used for Western blotting were prepared from whole cell lysates of *F. tularensis*. Proteins were then separated in two dimensions: by protein isoelectric point and by molecular mass. Resolved proteins were transferred to PVDF membranes by electroblotting and probed with sera from mice, rabbits, NHPs, and humans. Western blots were aligned with protein stained membranes of 2D-PAGE and immunoreactive proteins identified from a proteome map. Unknown proteins were identified using standard proteomics approaches, as described in ref. 11

for example, *in vitro* grown bacteria. 2D-PAGE is able to resolve many proteins in the bacterial proteome to a single protein spot, retaining the native protein processing and posttranslational modifications. Subsequently, proteins are transferred to nitrocellulose or PVDF membrane and probed with primary sera and conjugated secondary antibody, as per traditional Western blotting. Proteins may be stained after transfer to a membrane, and the captured image used to align regions of immunoreactivity with areas of protein staining. Excising the identified immunoreactive proteins from a second protein stained 2D-PAGE, and subsequent digestion with trypsin allows identification of proteins using mass spectrometry based techniques (e.g., MS/MS). This process is shown in flowchart format in Fig. 1. As with any experimental approach,

2D-Western blotting has well documented disadvantages that are reviewed elsewhere [7]. Nevertheless, it remains one of the most accessible immunoproteomics approaches, and can be carried out in any laboratory equipped with protein electrophoresis and electroblotting equipment.

This chapter details a *Se*rological *P*roteome *A*nalysis (SERPA) protocol that can be applied to many models of disease. In our laboratory, this protocol was developed specifically for screening immune sera to support the development and characterization of tularemia vaccines. Tularemia, a disease of humans and other mammals, is caused by the intracellular pathogen, *Francisella tularensis*. Although humans are an accidental host, *F. tularensis* is increasingly being isolated from infected humans in the United States and several European countries [8]. *F. tularensis* has gained significant attention in the post-9/11 era as one of six organisms designated as high priority agents by the US Center for Disease Control and Prevention that could be exploited as agents of bioterror (category A pathogens). Combined, the low infectious dose and ease of dissemination of type A *F. tularensis* have made it a threat to both military personnel and civilians alike [9]. There is currently no licensed vaccine available in North America, although an attenuated type B strain, known as the live vaccine strain (LVS), has been used to vaccinate at-risk military personnel and laboratory workers. To further characterize the immune response induced by vaccination with LVS, we and others have exploited SERPA in an attempt to determine correlates of protection or markers of vaccination in animal models of disease and infected or vaccinated humans [10–19].

2 Materials

Prepare all solutions using deionized water with a resistance of ≥18 MΩ and analytical grade reagents. Prepare and store all reagents at room temperature (unless indicated otherwise). Ensure appropriate safety precautions are followed and dispose waste materials as per waste disposal guidelines. Sodium azide is not added to reagents.

2.1 Antigen Preparation

1. Cell lysis solution: 7 M urea, 2 M Thiourea, 4 % CHAPS, 1 % DTT, 0.5 % Amidosulfobetaine-14 (ASB-14), in 5.2 mL of MQ water (*see* **Note 1**).
2. Cell lysis solution can be prepared ahead and divided into 1 mL aliquots. Store at −20 °C (*see* **Note 2**).

2.2 Isoelectric Focusing (IEF) Components

1. Immobilized pH gradient (IPG) strips (for example 17 cm pH 4–7 strips from Bio-Rad).
2. IPG solution: 20 µL Biolytes 3–10 stock, 180 µL of cell lysis solution and 10 µg of Orange G.

3. Disposable rehydration tray.
4. Light Mineral Oil.
5. Paper electrode wicks.
6. Isoelectric focusing cell, for example, PROTEAN® Isoelectric focusing cell, including IEF tray from Bio-Rad.

2.3 SDS Polyacrylamide Gel Components

1. Resolving gel buffer: 1.5 M Tris–HCl.
2. 30 % Bis-acrylamide solution.
3. Sodium dodecyl sulfate (SDS): 10 % stock solution in water (*see* **Note 3**).
4. Ammonium persulfate (APS): 10 % solution in water (*see* **Note 4**).
5. *N*, *N*, *N*′, *N*′-tetramethylethylenediamine (TEMED).
6. SDS-PAGE running buffer: 1× Tris/Glycine/SDS (TGS) running buffer diluted from 10× stock (e.g., from Bio-Rad) with MQ water (*see* **Note 5**).
7. Precision Plus Prestained Dual Color molecular weight marker (Bio-Rad) (*see* **Note 6**).
8. Agarose solution: 50 mL ReadyPrep overlay agarose, 0.5 % in 1×TGS, 0.003 % Bromophenol blue (*see* **Note 7**).
9. Reducing equilibration solution: 0.05 g dithiothreitol (DTT), 0.1 g SDS, 0.68 mL 1 M Tris–HCl pH 8.8, 3.6 g urea, 3 g glycerol, in MQ water up to 5 mL (*see* **Note 8**).
10. Alkylating equilibration solution: 0.2 g iodoacetamide (IODO) (Sigma), 0.1 g SDS, 0.68 mL 1 M Tris–HCl, pH 8.8, 3.6 g urea, 3 g glycerol, in MQ water up to 5 mL (*see* **Note 9**).
11. PROTEAN® II xi 2-D Cell system (Bio-Rad), including: Lower buffer chamber (Buffer tank), cooling core/electrode assembly, cooling core gaskets, lid with electrical leads, outer and inner glass plates, gel casting stand, gel casting gaskets, sandwich clamps, IPG strip format combs, IPG strip format spacers, alignment card (*see* **Note 10**).
12. HC Powerpac™ Power Supply (Bio-Rad) (*see* **Note 11**).

2.4 Immunoblotting Components

1. Trans-blot® semi-dry transfer cell (Bio-Rad).
2. HC Powerpac™ Power Supply (Bio-Rad) (*see* **Note 11**).
3. Polyvinylidene fluoride (PVDF) membrane.
4. Extra thick filter paper.
5. Phosphate-buffered saline + tween (PBST) buffer: PBS with 0.5 mL/L of Tween 20.
6. PBST-M Solution: 400 mL PBST, 20 g fat free skim milk powder (*see* **Note 12**).

7. Primary antibody and secondary antibody (*see* **Note 13**).
8. Methanol.
9. Transfer buffer: 5.82 g Tris, 2.93 g Glycine, 200 mL Methanol, 0.375 g SDS, MQ water up to 1 L.
10. Fixing Solution: 20 mL methanol, 14 mL acetic acid in MQ water up to 200 mL per membrane.
11. Sypro Ruby Protein Blot Stain: 50 mL/membrane.
12. Enhanced chemiluminescence (ECL) detection kit (e.g., GE Healthcare): Prepared at a ratio of 40:1 for reagents A:B (200 µL B up to 8 mL with A).
13. Kodak X-ray film.
14. Developer solution: Dilute 217 mL of concentrated Kodak developer stock solution with 783 mL MQ water, as recommended by the manufacturer (*see* **Note 14**).
15. Fixer solution: Dilute 217 mL of concentrated Kodak fixer stock solution with 783 mL MQ water, as recommended by the manufacturer (*see* **Note 14**).
16. Gel imager, e.g., Bio-Rad FluorS with image analysis software such as PDQuest.

3 Methods

3.1 Sample Preparation

Unless otherwise specified, carry out all procedures at room temperature.

3.1.1 Serum

The serum used to illustrate the results of the current method were generated from animal models of the disease tularemia, caused by the bacterium *Francisella tularensis*. The description of how the serum was generated is described elsewhere [11]. The method is compatible with sera from humans or animals, as described in the accompanying notes.

3.1.2 Antigen Preparation

1. Thaw a vial of *F. tularensis* or O-antigen deficient strain *ΔwbtI* freezer stocks, prepared as described in ref. 10. Streak onto cysteine heart agar plates and grow overnight at 37 °C.
2. Scrap bacterial colonies from plate using inoculating loop or sterile slide and transfer to a sterile microcentrifuge tube containing 1 mL sterile MQ water.
3. Vortex bacteria to resuspend, then pellet by centrifugation at 1,350 × *g* in a bench top centrifuge.
4. Remove supernatant and replace with 1 mL of MQ water and repeat **step 3**.

5. Remove supernatant and add 12 times the pellet volume of Cell Lysis Solution and vortex until solution clears (*see* **Note 15**).
6. Quantify protein content of cell lysates using a protein quantification assay, such as the Bradford assay or similar [20].

3.2 Two Dimensional Polyacrylamide Gel Electrophoresis (2D PAGE)

3.2.1 Rehydration of IPG Strips

1. Transfer 100 µg of protein antigen into a clean microfuge tube and dilute to a total volume of 300 µL with cell lysis solution (*see* **Note 16**).
2. Add 15 µL Biolytes solution containing orange G dye (*see* **Note 16**).
3. Shake for 1 h at room temperature.
4. Centrifuge at 14,000 rpm in a benchtop centrifuge for 30 s to bring the solution to the bottom of the tube.
5. Pipette solution into rehydration tray and overlay with one 17 cm IPG strip, pH 4–7, ensuring contact between exposed gel and solution. Incubate strips for 1 h at room temperature (*see* **Note 17**).
6. Overlay each strip with 1.5 mL of mineral oil and incubate overnight at room temperature (*see* **Note 18**).

3.2.2 First Dimension Separation: Isoelectric Focusing (IEF)

1. Place paper wicks over electrodes of IEF focusing tray. Add 7 µL MQ water to each wick (*see* **Note 19**).
2. Remove rehydrated IPG strips from rehydration tray using forceps, and drain excess mineral oil. Lay strip onto IEF tray, gel side down ensuring good contact with both electrodes. Overlay each strip with mineral oil.
3. Place the tray containing strips into the IEF cell and carry out IEF using the following steps at 20 °C: 200 V for 1 h, 500 V for 1 h, ramp to 5,000 V over 5 h, focus to 80,000 Vh at 5,000 V, and maintain at 500 V until a cumulative total of 95,000 Vh has been reached (approximately 24 h total) (*see* **Note 20**).

3.2.3 Second Dimension Separation: PAGE

1. One day prior to commencing the SDS-PAGE separation part of the protocol, prepare gels.
2. Assemble large gel plates, as described by the manufacturer.
3. Prepare gel mix in a vacuum flask: 40 mL 30 % Bis-Acrylamide, 25 mL 1.5 M Tris–HCl, pH 8.8, 33.5 mL MQ water, 1 mL 10 % SDS (*see* **Note 21**).
4. Degas gel mix for 5 min and then add 400 µL of 10 % APS and 40 µL TEMED to 80 mL of gel mix. Pipette solution into assembled gels plates, taking care to avoid introducing air bubbles. Insert comb between the top of the two plates, and allow the gel solution to polymerize at room temperature for 1 h.
5. Transfer polymerized gels into an airtight bag containing 100 mL MQ water and store at 4 °C until use the following day.

6. Immediately prior to running the SDS-PAGE separation, place focused strips into equilibration tray, gel side up.
7. Equilibrate each focused IEF strip with 2 mL DTT solution at room temperature for 20 min (*see* **Note 22**).
8. Remove DTT solution and equilibrate each IEF strip with 2 mL IODO solution at room temperature for 20 min (*see* **Note 22**).
9. Heat agarose solution until melted and pipette 1 mL into each gel well. Place IEF strip into gel well, ensuring good contact with surface of gel. Allow agarose to set for 10 min at 4 °C.
10. Assemble gels into holding clamps (as appropriate for the gel assembly used) and fill the upper loading chamber with SDS-PAGE running buffer (*see* **Note 23**).
11. Add 10 µL molecular weight marker to the marker lane (*see* **Note 24**).
12. Add 1,100 mL of SDS-PAGE running buffer to tank before placing the plate assembly into tank (*see* **Note 25**).
13. Run at 30 mA/gel for approximately 4.25 h, with water cooling (*see* **Note 26**).

3.3 Western Blotting and Detection

3.3.1 Semi-dry Transfer

1. Incubate gel in 150 mL transfer buffer for 15 min.
2. Rehydrate PVDF membrane for 10 s in methanol (*see* **Note 27**) and then rehydrate three filter papers (cut to size) and membrane in transfer buffer for 15 min (*see* **Note 28**).
3. Place one filter paper on platinum anode of semi-dry blotter, removing bubbles.
4. Place PVDF membrane on top of filter paper, and subsequently overlay with the SDS-gel (*see* **Note 29**).
5. Place two filter papers directly on top of the gel and attach second electrode.
6. Transfer proteins from gel to PVDF membrane at 15 V for 1 h.
7. Remove PVDF membrane from blotting apparatus and air dry for 15 min (*see* **Note 30**).
8. Place protein side down in 200 mL blot stain fixing solution for 15 min.
9. Incubate membrane, protein side up in 200 mL MQ water for 5 min. Repeat an additional three times.
10. Incubate membrane with 50 mL Sypro Ruby Protein Blot stain for 15 min with agitation, protected from light.
11. Incubate membrane in MQ water for 1 min. Repeat an additional two times.
12. Record image using protein gel scanner, with UV illumination for 5 s (*see* **Note 31**).

13. Air dry membrane for 15 min.
14. Rehydrate membrane for 5 s in methanol and incubate with 50 mL PBST-M solution overnight at 4 °C with agitation (*see* **Note 32**).

3.3.2 Probing and Fluorescence Detection of Immunoreactivity

1. Incubate membrane in 100 mL PBST solution for 5 min, with gentle agitation. Repeat once. For all incubation steps, use gentle agitation, for example an orbital shaker.
2. To prepare the primary antisera, dilute 50 μL murine antisera in 50 mL PBST-M solution (1:1,000 dilution) (*see* **Note 13**).
3. Remove PBST, then add 50 mL of primary antisera solution and incubate for 1 h.
4. Remove solution from membrane and incubate membrane in 100 mL PBST solution for 5 min.
5. Add 100 mL fresh PBST solution for 30 min. Repeat an additional two times.
6. Incubate membrane in 100 mL fresh PBST for 5 min.
7. To prepare the secondary antibody, dilute 10 μL [goat] anti-mouse IgG, HRP-labelled antibody (Perkin Elmer Life and Analytical Sciences) in 50 mL PBST-M solution (1:5,000 dilution) (*see* **Note 13**).
8. Remove PBST and add 50 mL of secondary antibody solution, diluted appropriately, as denoted in this section (*see* **Note 13**). Incubate for 1 h.
9. Replace secondary antibody solution with 100 mL PBST solution and incubate with for 5 min.
10. Drain PBST and add 100 mL fresh PBST solution for 30 min. Repeat an additional two times.
11. Incubate membrane in 100 mL fresh PBST for 5 min.
12. Remove PBST and then add 8 mL of ECL fluorescence detection solution to the membrane for 3 min (*see* **Note 33**).

3.3.3 Developing Film Images and Alignment

Developing film images of Western blotted membranes (*see* **Note 34**).

1. Lay membrane between two transparency sheets.
2. Place membrane in autoradiography cassette.
3. In dark room, lay one sheet of Kodak Biomax Scientific imaging film on top of membrane. Close cassette and expose film for 1 min (*see* **Note 35**).
4. Incubate film in developer for 1 min.
5. Transfer to separate container of water and incubate film in water for 2 min.
6. Transfer film into fixer for a minimum of 5 min (*see* **Note 36**).
7. Transfer film to a second tray containing water for 2 min.
8. Air dry film.

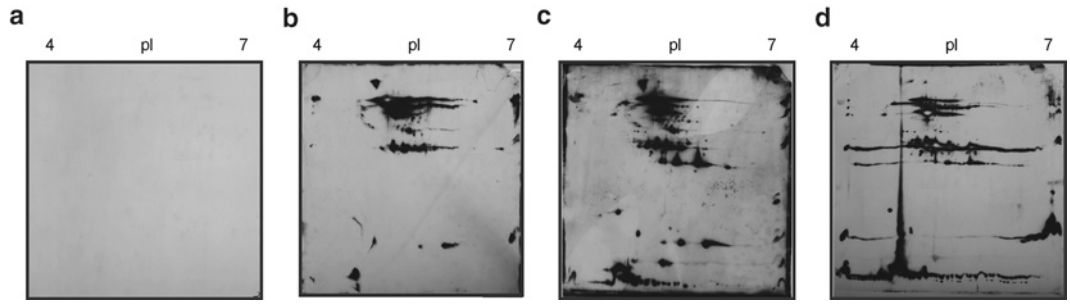


Fig. 2 2D Western blots probed with sera from mice vaccinated intranasally with *F. tularensis* LVS. Balb/c mice were immunized with lot 17 LVS or saline intranasally (~2.5×10^3 cfu per dose) and terminally bled. Shown are representative blots, probed with sera from individual mice (**a**) SHAM immunized, and LVS immunized mice, euthanized (**b**) 4 weeks post vaccination, (**c**) 6 weeks post vaccination, and (**d**) 8 weeks post vaccination

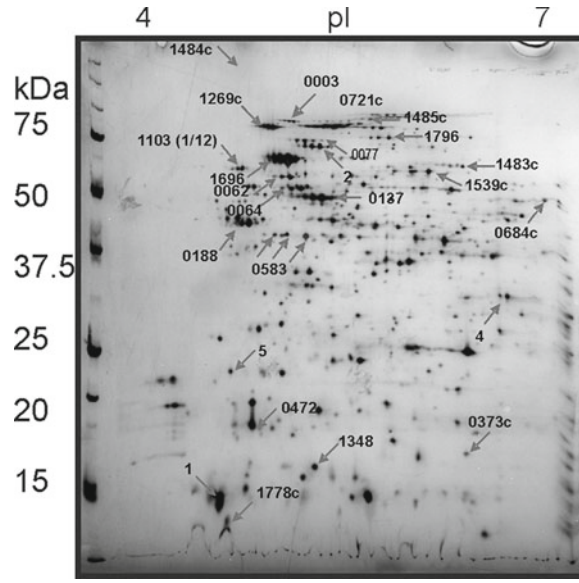


Fig. 3 2D-PAGE of total cell lysate of *F. tularensis*. 100 µg of protein was separated in the pH range 4–7. Proteins were visualized using non-fixing silver staining

9. Repeat **steps 1–8** for 30 s, 2 min, and 5 min film exposures (*see* **Note 35**).
10. Record image using protein gel scanner, with white light illumination for 3.5 s (*see* **Note 37**). Representative blots are shown in Fig. 2.
11. Scanned gel images can be aligned with the image of the protein stained membrane or an equivalent protein stained 2D-PAGE using image analysis software, such as PDQuest.
12. Protein identification can be carried out using now standard proteomics methods as outlined, for example, in ref. 10. Figure 3 shows a representative silver stained gel, indicating identified immunoreactive proteins corresponding to the Western blots in Fig. 2.

4 Notes

1. When adding water to dissolve the urea powder, note that the solution will initially be cold to touch. Mix frequently to dissolve completely.
2. Urea is used as a denaturant to increase the solubilzation of proteins, however it is not stable. Do not heat urea containing solutions or store above 20 °C. In the presence of heat, urea breaks down to form isocyanate, which leads to carbamylation of proteins. Protein carbamylaton will result in artifacts on 2D-PAGE. In addition, carbamylation of protein N-termini or lysine side chains can interfere with downstream enzymatic digestion and protein identification by mass spectrometry.
3. SDS is a respiratory, skin, and eye irritant. Weigh in fume hood. The stock solution may form crystalline precipitate in colder temperatures, warm slightly to redissolve.
4. APS is prepared fresh for each use.
5. For better reproducibility of 2D-PAGE, we use a commercial 10× stock solution, purchased, for example, from Bio-Rad. Dilute the stock in MQ water and stir slowly with magnetic stirrer to avoid formation of bubbles.
6. Prestained markers allow monitoring of protein transfer efficiency; at the end of the transfer, prestained markers will be visible on the membrane and little or no marker will be visible in the gel.
7. The agarose solution can be divided into 5 mL aliquots, allowed to solidify and stored at 4 °C. Prior to use, heat gently in water on a hot plate until liquefied.
8. Prepare Tris-HCl, urea, and glycerol in water for both solutions. Divide into two aliquots before adding DTT and IODO. Vortex both solutions thoroughly before adding SDS. Do not vortex once SDS is added to avoid foaming, instead mix gently by inverting a few times.
9. IODO is light sensitive, it should be made fresh and be kept in the dark.
10. 2D PAGE equipment can be purchased as a complete system from many vendors, for example, the PROTEAN® II xi 2-D Cell system (excluding the power supply) from Bio-Rad. Individual components may also be purchased separately.
11. Vendors such as Bio-Rad offer multiple power supply options. It is important to consider the outputs required for gels and blotting. For example, with the PROTEAN® II xi 2-D Cell, the power supply must be capable of outputting a constant

30 mA/gel (or 60 mA/port). Many basic power supplies are sufficient for 2D-PAGE. However, even though protein transfer is performed with a constant 15 V regardless of the number of Trans-blot® semi-dry transfer cells operating simultaneously, each cell will draw approximately 0.5 A. The output of many basic power supplies may be insufficient, and high current power packs, such as the Bio-Rad HC Powerpac™ or equivalent, is therefore recommended to accommodate up to four PROTEAN® II xi 2-D Cells or four Trans-blot® semi-dry transfer cells.

12. We buy commercially available skimmed milk (e.g., Carnation) powder for this purpose. After solubilizing, filter using a Whatmann filter to remove particulates.

13. Human
Primary antisera: Dilute 100 µL human antisera in 50 mL PBST-M solution (1:500 dilution).

Secondary antibody: Dilute 0.5 µL [goat] anti-human IgG, HRP-labelled antibody (Perkin Elmer Life and Analytical Sciences) in 50 mL PBST-M solution (1:100,000 dilution).

Nonhuman primate
Primary antisera: Dilute 100 µL monkey antisera in 50 mL PBST-M solution (1:500 dilution).

Secondary antibody: Dilute 1 µL [goat] anti-monkey IgG, HRP-labelled antibody (Fitzgerald Industries) in 50 mL PBST-M solution (1:50,000 dilution).

Rabbit
Primary antisera: Dilute 20 µL rabbit antisera in 50 mL PBST-M solution (1:2,500 dilution).
Secondary antibody: Dilute 10 µL [goat] anti-rabbit IgG, HRP-labelled antibody (Perkin Elmer Life and Analytical Sciences) in 50 mL PBST-M solution (1:5,000 dilution).

14. Based on the Kodak technical information sheets, the recommended replenishing rates are 20 mL of developer and 25 mL of fixer for each 20.3 cm × 25.4 cm film that is processed. When three times the original volume has been replenished, the solutions should be completely replaced. At minimum, developer and fixer solutions should be completely replaced every 3 months. However, when performing multiple experiments, solutions are generally replaced bi-weekly.

15. If cell lysis does not occur immediately, incubate at room temperature for 30 min with agitation. If solution has not cleared, add 10–20 µL of cell lysis solution, and continue until cells are lysed. Centrifuge to remove cell debris, and if dealing with pathogenic bacteria, plate 10 % of the sample on a suitable agar medium to check for sterility before removing from biocontainment facility.

Table 1
IEF programs for IPG strips of different lengths

IPG strip	IEF method	Target volt hours (Vh)	Approximate duration (h)
7 cm	30 min at 200 V 30 min at 500 V 2 h ramp to 6,500 V 3 h focusing at 6,500 V Maintain at 500 V	18,000–25,000	5–6
11 cm	45 min at 200 V 45 min at 500 V 4 h ramp to 4,000 V 35,000 VH focusing at 4,000 V Maintain at 500 V	45,000–55,000	20–24

16. Volumes noted are for large gel format 17 cm IPG strips. Typically, these can be loaded with 80–600 μg of protein. We find for analytical purposes 100 μg gives best resolution of gel spots. Other IPG strips can be purchased that are compatible with PAGE equipment from other manufacturers.
17. When overlaying strips, ensure that the rehydration tray is clean and dry. Remove plastic backing from the strip, with care not to damage the gel. Overlay strip onto rehydration solution carefully, using forceps to grip the plastic end of the strip. Take care not to create air bubbles as the strip is overlaid. There is a high degree of manual dexterity involved in doing this. We suggest practicing using the plastic backing from the strips to overlay on water spiked with orange G.
18. If you wish to avoid overlaying strips with mineral oil, an alternative solution is to fill empty wells of the rehydration tray with water and tightly seal the lid. The moist atmosphere prevents the strips from drying out. This is only applicable in cases where the number of wells permits. In addition, care must be taken not to allow water to overflow into wells containing the rehydrating IPG strips.
19. Moistening the IEF wicks helps remove salt from samples.
20. Different IEF programs are used for different IPG strip lengths, shown in Table 1.
21. Amounts indicated are sufficient for two large 12 % gels. For resolution of proteins between 15 and 150 kDa, a 12 % gel is usually appropriate. For better resolution of lower molecular weight proteins, a higher percentage acrylamide should be used. Conversely, for higher molecular weight proteins, a lower percentage acrylamide should be used. Gels can be stored for 1–2 days at 4 °C.

22. Typically, strips are placed on rocker or to agitate gently during equilibration, at room temperature.
23. The seal between the gel plates and the cooling core gaskets is important to ensure containment of the SDS-PAGE running buffer. To ensure the seal is sufficient, we recommend filling the upper buffer chamber half way initially to check for leaking. If there are no leaks, continue filling the chamber. If there are leaks, pour out the running buffer and disassemble the gel/cooling core assembly. Moistening the gaskets with water sometimes helps; remount the gels into the holding clamps and test again. If leaking persists, applying vacuum grease to the gaskets can improve the seal but is difficult remove from the glass plates. If the leaking is chronic, the gaskets may need to be replaced.
24. To load molecular weight marker, we recommend using pipette and gel loading tip. These tips have an elongated end which allows the marker to be loaded directly into marker well with no overflow. The molecular weight marker is always added after the SDS-PAGE running buffer to avoid displacing the marker out of the well.
25. Add running buffer to gel tank slowly, avoiding creation of bubbles. If bubbles are formed below the gel plates, use a spatula or similar device to move bubbles to one side of the gel tank.
26. At 30 mA/gel, the run time is usually 4.25 h. Monitor the gel and stop the run once the dye front has just migrated off of the gel.
27. SDS, present in denaturing gels and SDS–protein complexes, promotes protein elution from the gel during transfer. However, SDS hinders binding of proteins to the membrane. Methanol is added to the transfer buffer because it removes SDS from the SDS–protein complexes, improving protein-membrane binding. PVDF membranes are recommended over nitrocellulose for Western blotting because they are more stable in the presence of organic solvents. Unlike nitrocellulose, PVDF membranes must be activated by a pretreatment with 100 % methanol prior to equilibration in transfer buffer.
28. Shrinking or expanding of the gel or membrane during transfer will result in an apparent blurring of proteins transferred to the membrane. To avoid this, the gel must be adequately equilibrated in transfer buffer, and neither the gel nor the membrane must be permitted to dry out. The gel, membrane, and filter papers must also be prevented from drying to ensure proper conductivity during the transfer.
29. Proteins may begin to diffuse passively into the membrane immediately upon contact. Adjustment of the gel position once it has been laid should therefore be avoided to prevent transfer of the same protein to multiple spots on the membrane.

30. At all stages, it is important to limit the handling of the membrane. Forceps, rather than gloved hands, should be used. When handling the membrane, grasping only the edge or corner (even with forceps) will minimize the introduction of artifacts that may become visible when developing the film.
31. One gel can be reserved and silver stained in order to use as a reference proteome map. Staining of the PVDF membrane to visualize proteins, gives a reasonable reference map for aligning developed blot image on X-ray film with protein stained images, but blot staining is not as sensitive as silver staining of reference 2D-PAGE, so the silver-stained reference gel is also frequently used. The blot staining step is not essential. A non-fixing silver stain protocol that is compatible with potential downstream proteomics protocols is given here [21].
32. Blocking of blots is usually carried out overnight at 4 °C for convenience. It can also be carried out at room temperature for 1–2 h with similar results. Blocking and all subsequent incubation of the membrane are typically done on a rocker or orbital shaker to agitate gently.
33. Apply ECL solution immediately before exposing and developing film as the substrate will be used up and fluorescence signal will reduce over time.
34. If you have access to a scanner equipped with chemiluminescence, this can be a more rapid alternative to traditional X-ray film. Usually chemiluminescence works with HRP-fluorescence detection, and recommendations are usually provided by the manufacturer of the scanner.
35. When adapting this protocol, or working with a new set of sera, we generally use a range of exposure times, typically 10 s, 30 s, 1 min, 2 min, and 5 min. This range allows the optimal exposure time to be determined for a particular serum set. In addition, we routinely expose and develop films at 30 s, 1 min, and 5 min. Although comparisons across the dataset were always made using 1 min exposure, shorter or longer exposures sometimes helped in alignment of immunoreactive areas with protein stained gels, for example where a shorter exposure helps separate an intensely immunoreactive region into several distinct areas.
36. The X-ray films are coated in silver halides which are converted to metallic silver during developing to produce an image. In addition to making the image permanent, the fixing step also dissolves unconverted silver halides from the rest of the film. Since these silver halides are light sensitive, residual silver halides, as a result of insufficient fixing time, will result in discoloration of the film once it is removed from the dark room. Therefore a minimum of 5 min should be used for the fixing step.

37. Alignment can be carried out using open source or commercially available gel alignment software, for example, PDQuest (Bio-Rad). This is highly recommended if measurement of relative intensity of each immunoreactive area by densitometry is required. If appropriate software is not available, then a scanned image of an equivalent silver-stained 2D-PAGE can be aligned manually with the developed blot.

Acknowledgments

This work was supported by Federal funds from the National Institute of Allergy and Infectious Diseases, National Institutes of Health, Department of Health and Human Services, Contract No. HHSN266200500041C and in part by the National Research Council, Canada. The authors thank Freyja Lynn, Drs Vicki Pierson, Kristin Debord, Patrick Sanz (National Institutes of Allergy and Infectious Diseases), Dr J. Wayne Conlan, Luc Tessier and Marianne Savicky (National Research Council, Canada), and Gretchen Stup (DynPort Vaccine Company) for their contributions throughout this work. The authors also thank Patricia Massel for method testing and critical reading of the manuscript.

References

1. Engelfried JJ, Spear F (1966) Modified agglutination test for *Pasteurella tularensis*. Appl Microbiol 14:267–270
2. Carlsson HE, Lindberg AA, Lindberg G, Hederstedt B, Karlsson KA, Agell BO (1979) Enzyme-linked immunosorbent assay for immunological diagnosis of human tularemia. J Clin Microbiol 10:615–621
3. Bevanger L, Maeland JA, Naess AI (1988) Agglutinins and antibodies to *Francisella tularensis* outer membrane antigens in the early diagnosis of disease during an outbreak of tularemia. J Clin Microbiol 26:433–437
4. Klade CS (2002) Proteomics approaches towards antigen discovery and vaccine development. Curr Opin Mol Ther 4:216–223
5. Vytvytska O, Nagy E, Bluggel M, Meyer HE, Kurzbauer R, Huber LA, Klade CS (2002) Identification of vaccine candidate antigens of *Staphylococcus aureus* by serological proteome analysis. Proteomics 2:580–590
6. Parida SK, Kaufmann SH (2010) The quest for biomarkers in tuberculosis. Drug Discov Today 15:148–157
7. Kilmury SL, Twine SM (2010) The *Francisella tularensis* proteome and its recognition by antibodies. Front Microbiol 1:143
8. Eliasson H, Broman T, Forsman M, Back E (2006) Tularemia: current epidemiology and disease management. Infect Dis Clin North Am 20:289–311, ix
9. Dennis DT, Inglesby TV, Henderson DA, Bartlett JG, Ascher MS, Eitzen E, Fine AD, Friedlander AM, Hauer J, Layton M, Lillibridge SR, McDade JE, Osterholm MT, O'Toole T, Parker G, Perl TM, Russell PK, Tonat K (2001) Tularemia as a biological weapon: medical and public health management. JAMA 285: 2763–2773
10. Fulton KM, Zhao X, Petit MD, Kilmury SL, Wolfraim LA, House RV, Sjostedt A, Twine SM (2011) Immunoproteomic analysis of the human antibody response to natural tularemia infection with Type A or Type B strains or LVS vaccination. Int J Med Microbiol 301: 591–601
11. Twine SM, Petit MD, Fulton KM, House RV, Conlan JW (2010) Immunoproteomics analysis of the murine antibody response to vaccination with an improved *Francisella tularensis* live vaccine strain (LVS). PLoS One 5:e10000
12. Havlasova J, Hernychova L, Brychta M, Hubalek M, Lenco J, Larsson P, Lundqvist M, Forsman M, Krocova Z, Stulik J, Macela A

(2005) Proteomic analysis of anti-*Francisella tularensis* LVS antibody response in murine model of tularemia. Proteomics 5:2090–2103

13. Havlasova J, Hernychova L, Halada P, Pellantova V, Krejsek J, Stulik J, Macela A, Jungblut PR, Larsson P, Forsman M (2002) Mapping of immunoreactive antigens of *Francisella tularensis* live vaccine strain. Proteomics 2:857–867
14. Janovska S, Pavkova I, Hubalek M, Lenco J, Macela A, Stulik J (2007) Identification of immunoreactive antigens in membrane proteins enriched fraction from *Francisella tularensis* LVS. Immunol Lett 108:151–159
15. Janovska S, Pavkova I, Reichelova M, Hubaleka M, Stulik J, Macela A (2007) Proteomic analysis of antibody response in a case of laboratory-acquired infection with *Francisella tularensis* subsp. *tularensis*. Folia Microbiol (Praha) 52: 194–198
16. Huntley JF, Conley PG, Rasko DA, Hagman KE, Apicella MA, Norgard MV (2008) Native outer membrane proteins protect mice against pulmonary challenge with virulent type A *Francisella tularensis*. Infect Immun 76: 3664–3671
17. Huntley JF, Robertson GT, Norgard MV (2010) Method for the isolation of *Francisella tularensis* outer membranes. J Vis Exp
18. Kovarova H, Stulik J, Macela A, Lefkovits I, Skrabkova Z (1992) Using two-dimensional gel electrophoresis to study immune response against intracellular bacterial infection. Electrophoresis 13:741–742
19. Macela A, Stulik J, Hernychova L, Kroca M, Krocova Z, Kovarova H (1996) The immune response against *Francisella tularensis* live vaccine strain in Lps(n) and Lps(d) mice. FEMS Immunol Med Microbiol 13:235–238
20. Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72:248–254
21. Zhao L, Liu C, Sun Y, Ban L (2012) A rapid and simplified method for protein silver staining in polyacrylamide gels. Electrophoresis 33: 2143–2144

Chapter 6

Pre-absorbed Immunoproteomics: A Novel Method for the Detection of Bacterial Surface Proteins

Guangjin Liu, Wei Zhang, and Chengping Lu

Abstract

Bacterial surface proteins are often investigated as potential vaccine candidates and biomarkers of virulence. In this chapter, a novel method for identifying bacterial surface proteins is presented, which combines immunoproteomic with immunoserologic techniques. Immunoproteomics, involving the separation of proteins by two-dimensional electrophoresis (2-DE) and Western blotting, has become an increasingly popular method for identifying immunoreactive proteins. In conventional serological technique, cross-absorption is a powerful method used to minimize cross reaction during agglutination assays. The serum pre-absorption process in our method was developed from cross-absorption but modified to remove antibodies that recognized bacterial surface antigens, thereby generating pre-absorbed sera. These pre-absorbed sera were used in Western blotting after 2-DE to find bacterial surface protein antigens. This new method has been proven to be a useful tool for identifying surface proteins, and aid in the development of new vaccine subunits and disease diagnostics.

Key words Bacteria, Immunoproteomic, Immunoserology, Pre-absorbed

1 Introduction

Surface proteins of pathogenic bacteria can serve as protective antigens and virulence markers, though they can be technically challenging to identify. Several biochemical and microbiological techniques have been employed to characterize bacterial surface proteins, including multidimensional protein identification [1], stable isotope labeling [2], biotinylation approaches [3], surface shaving approaches [4], genome analyses, and protein and antibody arrays [5]. During the last decade, immunoproteomics has become an increasingly popular method used for identifying immunoreactive proteins. One technique involves the separation of proteins by two-dimensional electrophoresis (2-DE) and Western blotting. Though host antibodies primarily recognize proteins on the surface of a bacterium, non-surface proteins can also become immunogenic after proteolytic digestion in host antigen presenting cells (APCs).

Kelly M. Fulton and Susan M. Twine (eds.), *Immunoproteomics: Methods and Protocols*, Methods in Molecular Biology, vol. 1061, DOI 10.1007/978-1-62703-589-7_6,

Thus, distinguishing between antibodies that recognize surface and non-surface proteins is an important consideration when designing immunoproteomics experiments to identify potential vaccine candidates. Cross-absorption is a powerful method used in conventional serological techniques to minimize cross reaction during agglutination assay [6]. The traditional cross-absorption was performed between pathogen and homologous sera, or pathogen and heterologous sera to increase the accuracy of agglutination diagnoses. In our method, the intact bacteria was incubated with homologous hyperimmune sera to remove antibodies that recognized bacterial surface antigens and this produced novel "pre-absorbed" antisera. The pre-absorbed sera could provide a means of differentiating outer and cytoplasmic bacterial proteins in the immunoproteomic analysis. Protein spots that appeared in the blot probed with untreated sera, but that were absent in the blot treated with pre-absorbed sera, were assumed to be surface proteins.

Streptococcus suis is a zoonotic pathogen that can cause meningitis, pneumonia, septicemia, and arthritis in humans and animals [7]. The research on identifying *S. suis* surface protein could help to develop effective vaccines and treatments. We used untreated and "pre-absorbed" antisera of *S. suis* to probe 2-DE blots of *S. suis* cell lysates. Protein spots that appeared in the blot probed with untreated serum, but that were absent in the blot treated with pre-absorbed serum, were presumed to be surface proteins and identified using matrix-assisted laser desorption ionization–time of flight mass spectrometry (MALDI-TOF MS). Bioinformatics predictions and immunofluorescence assay verified that the identified proteins were actually located on the bacterial cell surface. A schematic diagram of the surface protein detection method is shown in Fig. 1 [8]. This method may prove useful for the development of new vaccine subunits and disease diagnostics.

2 Materials

2.1 2-DE Materials

1. *S. suis* strain and culture media: *S. suis* strain HA9801 cultured in Todd Hewitt broth (THB), pH 7.8, 37 °C.
2. Phosphate-buffered saline (PBS): 137 mM NaCl, 2.7 mM KCl, 10 mM $Na_2HPO_4{\cdot}2\ H_2O$, 2.0 mM KH_2PO_4, pH 7.4.
3. Mutanolysin working solution: 30 mM Tris-HCl (pH 7.5), 3 mM $MgCl_2$, 25 % sucrose, 125 U/mL mutanolysin.
4. Solution B: 7 M urea, 2 M thiourea, 4 % 3-[(3-cholamidopropyl) dimethylammonio]-1-propanesulfonate (CHAPS), and 65 mM dichlorodiphenyltrichloroethane (DDT).
5. 10 % (w/v) trichloroacetic acid (TCA).
6. 13 cm gel strips: Immobiline DryStrip (pH 4–7), supplied by GE Healthcare, Inc.

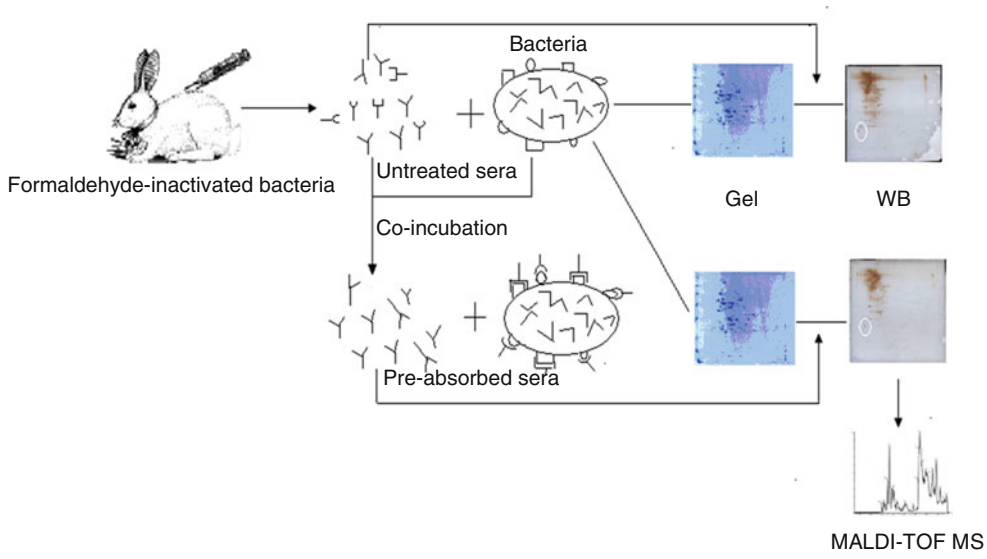


Fig. 1 Schematic diagram of the surface protein detection assay. One sample of the *S. suis* antiserum was pre-absorbed with whole-cell *S. suis* to remove the antibodies that recognize outer surface proteins ("pre-absorbed"). Then, untreated and "pre-absorbed" antisera were used to probe 2-DE gels of *S. suis* proteins. Spots that appear in the blot probed with untreated antiserum, but that were absent from the blot probed with "pre-absorbed" antiserum, were evaluated to identify the surface proteins. From ref. 8

7. 2-DE Clean-up kit (for example, from GE Healthcare, Inc.).
8. Rehydration solution: 2D starter Kit Rehydration Buffer (Bio-RAD, USA).
9. Equilibration buffer 1: 375 mM Tris–HCl (pH 8.8), 6 M urea, 2 % sodium dodecyl sulfate (SDS), 2 % DDT.
10. Equilibration buffer 2: 375 mM Tris–HCl (pH 8.8), 6 M urea, 2 % SDS, 2.5 % iodoacetamide.
11. SDS-polyacrylamide gel electrophoresis (PAGE) molecular weight pre-stained standard (for example, from Invitrogen).
12. Colloidal Coomassie brilliant blue G-250 dye: 8 % $(NH_4)_2SO_4$, 0.8 % H_3PO_4, 0.08 % Colloidal Coomassie brilliant blue G-250, 20 % CH_3O in distilled water (dH_2O).
13. Ultrasonic Processor.
14. Isoelectric focusing (IEF) equipment: for example, the Ettan IPGphor-3 IEF system (GE Healthcare, Inc.).
15. SDS-PAGE electrophoresis system: supplied by GE Healthcare, Inc.
16. 12 % SDS-polyacrylamide gel: 20 mL 30 % (w/v) polyacrylamide, 12.5 mL 1.5 M Tris–HCl (pH 8.8), 0.5 mL 10 % SDS, 0.5 mL 10 % APS (make fresh each time), 20 μL of *N,N,N′,N′*-tetramethylethylenediamine (TEMED), add dH_2O to total volume of 50 mL for making two gels.

17. SDS Running Buffer: 25 mM Tris, 192 mM glycine, 0.1 % SDS.
18. 0.5 % (w/v) agarose sealing buffer: 0.5 % (w/v) agarose and 0.002 % (w/v) bromophenol blue in SDS Running Buffer.
19. TOF Ultraflex II mass spectrometer from Bruker Daltonics, Inc.

2.2 Western Blotting Material

1. Immunologic adjuvants: Montanide ISA 206 VG (SEPPIC Co. Ltd).
2. Rabbit hyperimmune bacteria sera.
3. Polyvinylidene fluoride (PVDF) membranes.
4. Ponceau S dye: 0.1 % Ponceau S, 5 % acetic acid in dH_2O.
5. TBST solution: 50 mM Tris–HCl buffer (pH 7.4), 0.05 % Tween 20.
6. Blocking solution: 5 % skim milk in TBST.
7. Horseradish peroxidase (HRP)-goat anti-rabbit serum.
8. 3,3′-Diaminobenzidine (DAB) coloration.
9. Semi-dry blotting apparatus (for example, the TE77 from GE Healthcare).
10. Gel scanner (for example, Umax scanner from GE Healthcare, Inc.).

3 Methods

3.1 S. suis Protein Sample Extraction

Protein precipitations were performed according to Winterhoff [9], but with some modifications (*see* **Note 1**).

1. Exponential-phase bacterial cultures of 100 mL were centrifuged at 10,000 × *g* for 15 min at 4 °C and washed twice in PBS.
2. For Gram-positive bacteria only, pellets were resuspended in 4 mL Mutanolysin working solution and incubated for 90–120 min at 37 °C (*see* **Note 2**).
3. The spheroplasts of Gram-positive bacteria or Gram-negative bacteria pellets were collected in solution B and sonicated in an ice bath for 50 cycles (5 s on at 100 W, followed by 10 s off).
4. After 30 min incubation at 25 °C, unbroken cells were removed by centrifugation at 10,000 × *g* for 15 min at 4 °C.
5. Proteins in the supernatant were precipitated in 10 % pre-chilled TCA and incubated in ice bath for 30 min. After centrifugation at 10,000 × *g* for 10 min at 4 °C, the pellet was resuspended in 10 mL of pre-chilled acetone and washed twice.
6. The final pellet was dried in air.

3.2 Isoelectric Focusing

1. Isoelectric focusing (IEF) was performed using an Ettan IPGphor-3 IEF system and 13 cm pH 4–7 Immobiline Drystrip gel strips.
2. Prior to rehydration, the precipitated proteins were treated using a 2-DE Clean-up kit, according to the manufacturer's instructions, to remove contaminants that can interfere with electrophoresis.
3. Each of three immobilized pH gradient (IPG) strips was rehydrated overnight at room temperature (RT) in 150 μL of rehydration solution with 200 μg of protein.
4. IEF was carried out at 20 °C for 12 h (maximum voltage of 8,000 V and maximum current of 50 μA per IPG strip; total 28,000 Vh).

3.3 2D SDS-PAGE

1. Prior to 2D SDS-PAGE, each IPG strip was washed in equilibration buffer 1 for 15 min followed by equilibration buffer 2 for 15 min.
2. Each IPG strip plus an SDS-PAGE molecular weight pre-stained standard was loaded onto a homogeneous 12 % polyacrylamide gel and sealed with 0.5 % (w/v) agarose.
3. Electrophoresis was performed at 15 °C using an initial voltage of 110 V for 30 min, followed by 220 V until the tracking dye had reached the bottom of the gel.
4. Three gels were performed at the same time and one gel was stained with colloidal Coomassie brilliant blue G-250 dye and washed by dH_2O (*see* Fig. 2a).
5. The other two gels were prepared for Western blotting, as described in Subheading 3.6.

3.4 Preparation of Hyperimmune Sera

1. Rabbits, negative for *S. suis*, were immunized with formaldehyde-inactivated *S. suis* strain HA9801, using Montanide ISA 206 VG as the adjuvant.
2. Two doses of 1.0×10^9 cells/rabbit were administered by intramuscular injections at 3 week intervals.
3. Sera from immunized and control rabbits were collected before the first and after the second immunizations.
4. The titers of the sera were evaluated using indirect enzyme-linked immunosorbent assay (ELISA) [10].

3.5 Preparation of "Pre-absorbed" Sera

1. Exponential cultures of *S. suis* HA9801 were centrifuged at $3{,}000 \times g$ for 15 min at 4 °C, and then washed twice in PBS.
2. A total of 1.0×10^8 bacteria were suspended in 100 μL of HA9801 hyperimmune sera, incubated for 2 h at 37 °C, and then overnight at 4 °C (*see* **Note 3**).

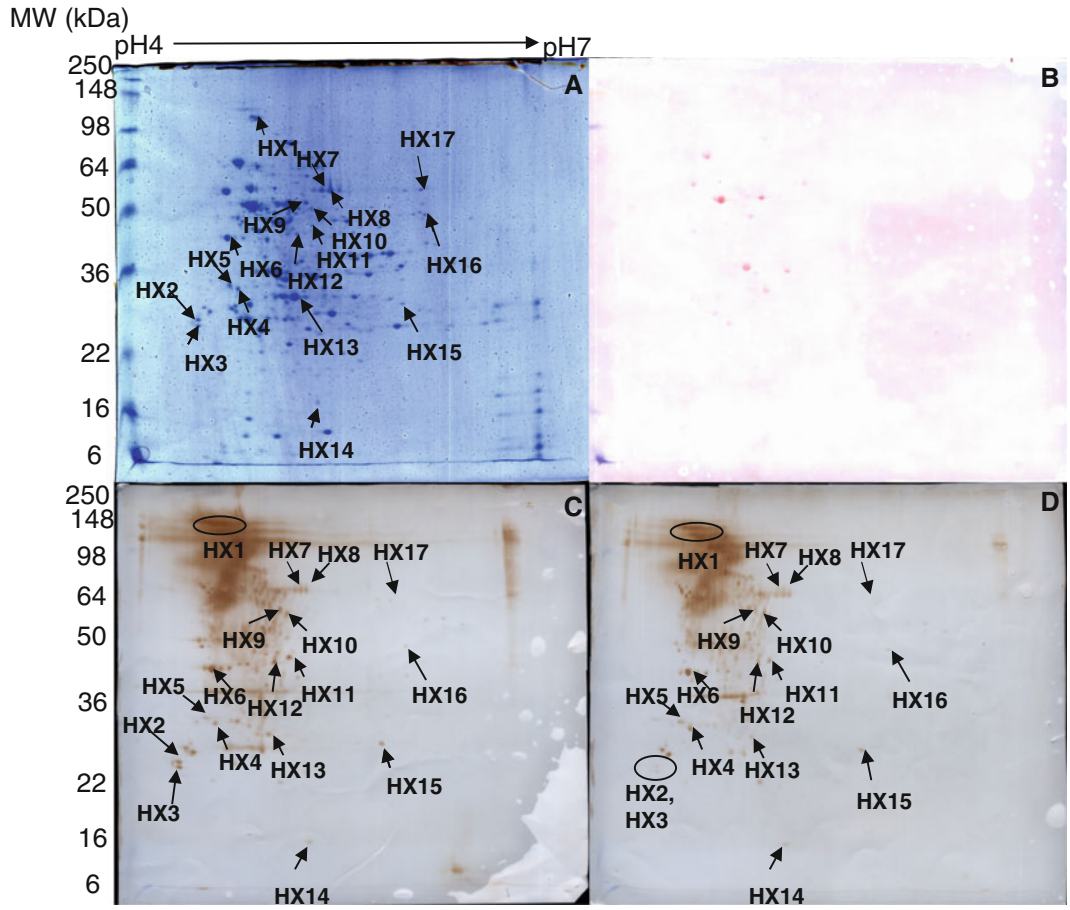


Fig. 2 2-DE gel and Western blot analyses of *S. suis* HA9801. (**a**) HA9801 total cell proteins (pH 4–7, 13 cm), stained with colloidal Coomassie brilliant blue G-250. (**b**) 2-DE blot of HA9801 stained with Ponceau S. (**c**) 2-DE blot of HA9801 proteins probed with untreated antiserum. (**d**) 2-DE blot of HA9801 proteins probed with "pre-absorbed" antiserum. From ref. 8

3. Bacteria were pelleted by centrifugation at 10,000×*g* for 30 min.
4. The supernatant, lacking antibodies specific for bacterial surface antigens, was collected and used for Western blotting as the "pre-absorbed" sera.

3.6 Western Blotting

1. Protein samples from each SDS-PAGE gel were transferred onto PVDF membranes for 2 h at 0.65 mA/cm^2 using a semi-dry blotting apparatus.
2. Membrane-bound proteins were detected by staining with Ponceau S. For this, the PVDF membrane was submerged in the Ponceau S dye with gentle agitation for 5 min.
3. The membrane was washed several times with dH_2O until the protein spots were visible, and then digitally scanned using Umax scanner (*see* Fig. 2b) [8].

4. Subsequently, the Ponceau S stain was removed from the membranes by rinsing gently in dH_2O. After removing Ponceau S, the membrane was blocked with blocking solution for 2 h at RT or overnight at 4 °C.
5. The blocked membrane was incubated with HA9801 hyperimmune serum or HA9801 "pre-absorbed" serum (1:500 dilution in blocking solution) for 2 h at RT (*see* **Note 4**).
6. Three washes with TBST of 10 min were carried out.
7. The membrane was incubated with horseradish peroxidase-goat anti-rabbit serum (1:10,000 dilution in blocking solution) for 1 h at RT.
8. Membrane was washed three times with TBST for 10 min each.
9. Membrane was then fully soaked in DAB coloration until optimal color was obtained.
10. The comparison of membranes probed with untreated and pre-absorbed blotting is manual. In Fig. 6.2c [8], membrane probed with the untreated sera showed many immune-reactive protein spots. Meanwhile an identical blot was probed using the "pre-absorbed" sera. Some of the spots, such as HX1, HX2, and HX3, which were less distinct or had disappeared (Fig. 2d [8]), were presumed to be surface proteins (*see* **Note 5**).

3.7 MALDI-TOF MS and Database Searching

1. Proteins identified from the 2-DE blots as potential surface proteins were excised from duplicate SDS-PAGE gels and used for in-gel trypsin digestion [11] and MALDI-TOF MS analysis.
2. Peptide mass fingerprinting (PMF) data were analyzed using the MASCOT server (http://www.matrixscience.com).
3. MASCOT searches were used to determine which peptides were to be considered significant and used for the combined peptide scores.
4. The extent of sequence coverage, number of matched peptides, and the score probability obtained from the PMF data were all used to validate protein identifications. Low-scoring proteins were either verified manually or rejected.

3.8 Bioinformatics Analysis

1. Sequences of the identified proteins were searched using the BLASTX server (http://blast.ncbi.nlm.nih.gov/Blast.cgi) to find homologous sequences.
2. The PSORT server (http://www.psort.org/), LocateP (http://www.cmbi.ru.nl/locatep-db), and Gpos-mPLoc (http://www.csbio.sjtu.edu.cn/bioinf/Gpos-multi/) programs were used to predict subcellular localizations of the proteins (*see* **Note 6**).

4 Notes

1. One immunoproteomics technique involving 2-DE and western blotting has been applied to a wide of range of organisms including *M. tuberculosis*, *Streptococcus pneumonia* [12], *Staphylococcus epidermidis* [13], and *Candida albicans* [14] for identifying immunoreactive proteins. In many cases this has led to the identification of novel antigens that have been demonstrated to be protective in animal models. However 2-DE has some limitations, such as the inability to analyze very large or very small, acidic, basic, or highly hydrophobic proteins [15], and proteins that are only expressed in vivo. So the amount of proteins detected in 2-DE was always less than the actual proteins in sample.
2. This step is carried out in Gram-positive bacterial protein sample extraction and can be skipped in Gram-negative bacteria. Because Gram-positive bacteria possesses thick cell wall consisting of many layers of peptidoglycan and teichoic acids, proteomic-based approaches for Gram-positive bacteria have been hampered by the preparation of protein samples. Mutanolysin, which is purified from the culture supernatant of *Streptomyces globisporus*, is an efficient reagent that can be used to obtain protoplasts of Streptococcus mutants [16], as it is highly effective for inducing the lysis of bacterial cells without any associated proteolytic activity. The working concentration of mutanolysin we used to generate 100 mL *S. suis* spheroplasts is 125 U/mL in mutanolysin working solution. The mutanolysin concentration should vary depending on the bacterium itself. It needs to make most of bacteria spheroplasts and the mixed liquid clearer after incubated for 90–120 min at 37 °C.
3. Preparation of "pre-absorbed" sera involves depletion of antibodies that bind surface exposed antigens on bacterial cells. The hyperimmune sera were added to the intact bacteria sediment to make a final concentration of approximately 10 %. The amount of the intact bacteria sediment needs to be optimized depending on the bacterium itself and the titer of serum in order to sufficiently remove antibodies for surface antigens in the sera.
4. In order to get more distinct immunoreactive spots, the sera dilution as primary antibody in WB may be adjusted (1:200 to 1:1,000 in blocking solution) according to the sera titer, the HRP-second antibody titer and the amount of proteins blotted on the membrane.
5. Immunoproteomics contain a series of reactions: sample preparation, IEF, SDS-PAGE, membrane transfer, probed with sera. The potential error of each step could be accumulated and affect the final result. In order to minimize the error,

the experiments were repeated at least in triplicate, and the separation profiles were consistent and highly reproducible. We performed the pre-absorbed immunoproteomics approach several times, only selecting the highly reproducible spots that disappeared or faded.

6. These bioinformatics software programs focus on the analysis of protein transmembrane regions and protein subcellular localizations to help confirm the trial result. But validation experiments must be performed, such as immunofluorescence assay, to confirm the protein is indeed located on the surface of bacteria.

References

1. Wolters DA, Washburn MP, Yates JR III (2001) An automated multidimensional protein identification technology for shotgun proteomics. Anal Chem 73:5683–5690
2. Zhu H, Pan S, Gu S, Bradbury EM, Chen X (2002) Amino acid residue specific stable isotope labeling for quantitative proteomics. Rapid Commun Mass Spectrom 16:2115–2123
3. Hempel K, Pane-Farre J, Otto A, Sievers S, Hecker M, Becher D (2010) Quantitative cell surface proteome profiling for SigB-dependent protein expression in the human pathogen *Staphylococcus aureus* via biotinylation approach. J Proteome Res 9:1579–1590
4. Tjalsma H, Lambooy L, Hermans PW, Swinkels DW (2008) Shedding & shaving: disclosure of proteomic expressions on a bacterial face. Proteomics 8:1415–1428
5. Cahill DJ (2001) Protein and antibody arrays and their medical applications. J Immunol Methods 250:81–91
6. Mittal KR, Higgins R, Lariviere S (1982) Evaluation of slide agglutination and ring precipitation tests for capsular serotyping of *Haemophilus pleuropneumoniae*. J Clin Microbiol 15:1019–1023
7. Gottschalk M, Segura M (2000) The pathogenesis of the meningitis caused by *Streptococcus suis*: the unresolved questions. Vet Microbiol 76:259–272
8. Zhang W, Liu G, Tang F, Shao J, Lu Y, Bao Y, Yao H, Lu C (2011) Pre-absorbed immunoproteomics: a novel method for the detection of *Streptococcus suis* surface proteins. PLoS One 6:e21234
9. Winterhoff N, Goethe R, Gruening P, Rohde M, Kalisz H, Smith HE, Valentin-Weigand P (2002) Identification and characterization of two temperature-induced surface-associated proteins of *Streptococcus suis* with high homologies to members of the Arginine Deiminase system of *Streptococcus pyogenes*. J Bacteriol 184: 6768–6776
10. Robertson ID, Blackmore DK (1989) Occupational exposure to *Streptococcus suis* type 2. Epidemiol Infect 103:157–164
11. Kussmann M, Nordhoff E, Rahber-Nielsen H et al (1997) Matrix-assisted laser desorption/ ionization mass spectrometry sample preparation techniques designed for various peptide and protein analytes. J Mass Spectrom 32:593–601
12. Ling E, Feldman G, Portnoi M, Dagan R, Overweg K, Mulholland F, Chalifa-Caspi V, Wells J, Mizrachi-Nebenzahl Y (2004) Glycolytic enzymes associated with the cell surface of *Streptococcus pneumoniae* are antigenic in humans and elicit protective immune responses in the mouse. Clin Exp Immunol 138:290–298
13. Sellman BR, Howell AP, Kelly-Boyd C, Baker SM (2005) Identification of immunogenic and serum binding proteins of *Staphylococcus epidermidis*. Infect Immun 73:6591–6600
14. Fernandez-Arenas E, Molero G, Nombela C, Diez-Orejas R, Gil C (2004) Low virulent strains of *Candida albicans*: unravelling the antigens for a future vaccine. Proteomics 4:3007–3020
15. Bunai K, Yamane K (2005) Effectiveness and limitation of two-dimensional gel electrophoresis in bacterial membrane protein proteomics and perspectives. J Chromatogr B Analyt Technol Biomed Life Sci 815:227–236
16. Siegel JL, Hurst SF, Liberman ES, Coleman SE, Bleiweis AS (1981) Mutanolysin-induced spheroplasts of Streptococcus mutants are true protoplasts. Infect Immun 31:808–815

Chapter 7

Identification of the Antigen Content of Electroimmunoprecipitates

N. Helena Beyer and Niels H.H. Heegaard

Abstract

Polyclonal antibodies including purified antibody fractions and animal or human antisera may react with unknown antigens or antigens other than their main specificity in reactions that are best visualized by gel electroimmunoprecipitation methods, e.g., when analyzing complex antigen mixtures. The great advantage of gel immunoprecipitation approaches is that each immunoprecipitate contains antigen in a pure form and that the precipitate is separated by position, shape, and size from other precipitates in the complex patterns of crossed immunoelectrophoresis. The identification of the antigen content of such immunoprecipitates is important but challenging because of the very stable, high affinity complex formation leading to precipitation in the gels. Here, we present detailed step-by-step recipes for identifying the antigen content of electroimmunoprecipitates.

Key words Antigen identification, Mass spectrometry, Immunoelectrophoresis, Immunoprecipitates, Electroimmunoprecipitation, Antigen, Antibody, Dissolution of immunoprecipitates

1 Introduction

Immunoelectrophoretic (IE) techniques are used for the quantitative determination of specific antigens that form immunoprecipitates with antibodies in gels [1]. The method works with most proteins if precipitating (polyclonal) antibodies are available and if the protein has a pI lower than that of the immunoglobulins. This ensures a differential electrophoretic mobility of antibody and antigen and hence electrophoretically mediated mixing in the gel. In comparison with the in-solution immunoprecipitation methods that are widely used in biochemistry, the IE techniques offer added possibilities for separation and quantification of multiple antigens in one operation. As an example, in crossed IE (CIE) [2] neither antibody nor sample needs to be pure. Thus, CIE can separate and quantify several antigens by using mixtures of antibodies in one analysis. For straightforward quantification of a protein in a crude

Kelly M. Fulton and Susan M. Twine (eds.), *Immunoproteomics: Methods and Protocols*, Methods in Molecular Biology, vol. 1061, DOI 10.1007/978-1-62703-589-7_7,

mixture, so-called rocket IE (RIE) [3], a one-dimensional gel immunoelectrophoresis method, is simple, but requires a monospecific antibody.

In all electroimmunoprecipitation techniques, the identity of antigens in unknown precipitates may be an issue. An example is when more than one precipitate emerges with a supposedly monospecific antibody reacting with sample mixtures, or when a particular precipitate appears for example in samples from a specific group of patients (i.e., a possible biomarker). To determine the identity of antigens in such immunoprecipitates IE-immunoblotting methods have been devised [4–6], but these approaches are dependent on other antibodies and thus require an educated guess as to the identity of the antigen in question. Additional techniques for identifying unknown antigens in precipitates also are based on assumptions about the antigens and on the use of purified proteins and/or additional specific antibodies. These techniques include tandem crossed IE [7], the use of intermediate gels with known antigens or specific antibodies [8], and immunofixation [9].

For identification of an unknown antigen, an orthogonal analytical method would be useful. An example is the excision of immunoprecipitates followed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) which is an approach that gives an estimate of antigen molecular weight [10] and offers possibilities for further characterization. This approach does not appear to have been systematically optimized regarding the best method of extracting and recovering the antigen part of the excised immunoprecipitates. Also, electroimmunoprecipitates may be excised and analyzed directly by mass spectrometry after in-gel tryptic digestion. This is a simple and robust method but due to the inaccessibility of some antigens in immunoprecipitates it may suffer from insufficient tryptic fragmentation yields and thus a low sensitivity and antigen identification success rate (*See* also ref. [11]).

Here, we present an optimized approach using RIE and CIE precipitates and SDS-PAGE with densitometry of protein bands and peptide mass fingerprinting (PMF) by mass spectrometry for identification of extracted proteins. The optimized extraction procedure is simple, sensitive, and robust. It allows the identification of antigens from electroimmunoprecipitates down to 135 ng of applied protein and may be used for characterization of antibody specificity.

2 Materials

Use MilliQ (MQ)-H_2O or similar for preparation of buffers and gels unless otherwise specified. Chemicals were of the highest grade available [high performance liquid chromatography (HPLC-grade)] except as noted.

2.1 Immuno-electrophoresis

2.1.1 Casting of Immune Gels

1. Levelling table.
2. 99 % Ethanol.
3. Tap water.
4. Detergent for dish washing in the laboratory.
5. KimWipes or other lint-free paper towels.
6. Electric heating plate with magnetic stirring.
7. Water bath.
8. Steel blades or knives.
9. 5× Tris-Tricine buffer (1 L) pH 8.6: 800 mL of MQ-H_2O, 26 mM Tricine (*N*-Tris(hydroxymethyl)methyl glycine), 74 mM Tris-base (2-amino-2(hydroxyl,ethyl)propane-1,3-diol), 1.8 mM calcium lactate. 5 N HCl to pH 8.6, then bring final volume to 1 L with MQ-H_2O.
10. 1× Electrophoresis buffer working dilution (2 L): 400 mL 5× stock Tris-Tricine buffer, up to 2 L in MQ-H_2O.
11. 2.5 g Agarose HSA 100 Litex (BioWhitaker Molecular).
12. Glass plates (thickness of 1 mm) or Gelbond film (Lonza Cologne, Walkersville, MD, USA).
13. Plate sizes: 5 cm × 5 cm, 5 cm × 7 cm, 7 cm × 7 cm, 7 cm × 10 cm, 10 cm × 10 cm, or 11 cm × 20.5 cm depending on the experiment (*see* **Note 1**).
14. Humidity chamber (plastic box with lid, line bottom with moist paper towels).
15. Gel puncher for sample application wells, 2.5 mm in diameter corresponds to 5 µL sample volumes.
16. Template (*see* Fig. 1). For an additional example, please *see* fig. 1.4 in ref. 12 for punching out sample application wells (*see* **Note 2**).

2.1.2 Running of Immuno Gels

1. Electrophoresis buffer from casting of gels.
2. Electrophoresis module (e.g., the SAS-MX isoelectric focusing module from Helena Biosciences). For additional details *see* ref. 12:
   - Electrophoresis apparatus.
   - Cooling thermostat in circulating water bath.
   - Power supply with four channels, individually adjustable.
   - Test probe for direct measurement/adjustment of the potential gradient in the agarose gel.
   - Filter paper wicks, five pieces of Whatman No. 1 (Whatman, Florham Park, NJ, USA) for low-voltage electrophoresis (2–3 V/cm, corresponding to a total of 70–110 V), must be same width as immunoelectrophoresis plates for even distribution of voltage.

Fig. 1 Template for placement of wells and gel slabs of 1st dimension electrophoresis on 9×13 cm plates. The template is placed under the glass plate supporting the 1.5 mm thick gel. After 1st dimension electrophoresis, the 5 mm gel along the edges is removed and the four gel slabs (2×11 cm), indicated by dotted lines, are cut out individually. Placement of the anode for the 1st dimension electrophoresis is indicated. (Color figure online)

2.1.3 Washing, Pressing, and Drying of Immuno Gels

1. Timer.
2. Phosphate-buffered saline (PBS) washing solution: 0.082 M NaCl, 0.043 M Na_2HPO_4, 0.0098 M KH_2PO_4, in MQ-H_2O, pH 7.4.
3. Whatman No. 1, moistened (Whatman, Florham Park, NJ, USA).
4. Soft absorbent tissue paper.
5. Glass plate (thickness of 8–10 mm).
6. Hot air stream for drying.

2.1.4 Visualization and Extraction of Electroimmunoprecipitates

1. Timer.
2. Plate holder.
3. Container for staining solution.
4. Aqueous staining solution, Coomassie Brilliant Blue: 0.1 % (w/v) Coomassie Brilliant Blue R-250, in MQ-H_2O.

5. Standard staining solution, Coomassie Brilliant Blue: 0.1 % (w/v) Coomassie Brilliant Blue R-250, 45 % (v/v) methanol, 5 % (v/v) acetic acid, in MQ-H_2O.
6. Container for destaining solution.
7. Aqueous destaining solution: MQ-H_2O.
8. Standard destaining solution: 20 % (v/v) of 99 % ethanol, 10 % (v/v) acetic acid, in MQ-H_2O.
9. Soft absorbent tissue paper.
10. Extraction buffer: 0.06 M Tris–HCl, 10 % (w/v) SDS in HPLC-grade H_2O, pH 7.

2.2 Running of SDS-PAGE

1. Spatula.
2. Power supply.
3. XCell Surelock mini-cell (Invitrogen Life Technologies, Carlsbad, USA).
4. 4–20 % Tris-glycine gels (Invitrogen Life Technologies, Carlsbad, USA).
5. Low range unstained SDS-PAGE molecular weight standard (Bio-Rad Laboratories, Hercules, USA) or SeeBlue Plus2 Pre-Stained molecular weight standard (Invitrogen Life Technologies, Carlsbad, USA).
6. Sample buffer, final concentration in sample, 20 µL/sample: 8 mM Tris–HCl, pH 8.0, 1.6 % (w/v) SDS (BDH Chemicals, Poole, UK), 2 % (w/v) Glycerol, 20 mM Dithiothreitol (DTT), 0.5 µg Pyronin G.
7. Tris-glycine gel running buffer: 25 mM Tris-base, pH 8.6, 192 mM Glycine, 0.1 % (w/v) SDS.

2.3 Staining of SDS-PAGE

2.3.1 Coomassie Brilliant Blue Staining

- Comassie Brilliant Blue Stain (*see* Subheading 2.1.4).
- Comassie Brilliant Blue Destaining (*see* Subheading 2.1.4).

2.3.2 Mass Spectrometry-Compatible Silver Staining According to Ref. 13

1. Fixating solution: 45 % (v/v) methanol, 5 % (v/v) acetic acid, 45 % (v/v), MQ-H_2O.
2. MQ-H_2O.
3. Sensitizing solution (50 mL): 0.02 % Sodium thiosulfate ($Na_2S_2O_3 \cdot 5\ H_2O$) in MQ-H_2O.
4. Staining solution (50 mL): 0.1 % Silver nitrate ($AgNO_3$) in MQ-H_2O (prepared just before use, MQ-H_2O chilled at 4 °C).
5. Developing solution (50 mL): 2 % Sodium carbonate (Na_2CO_3), 15 µL of 37 % formaldehyde, in MQ-H_2O.
6. Stop solution (50 mL): 1 % Acetic acid in MQ-H_2O.

2.4 In-Gel Digestion with Trypsin

1. Gloves.
2. Lint-free wipes.
3. Ice.
4. 99 % Ethanol for wiping of surfaces and equipment.
5. Laminar air flow (LAF) hood.
6. Benchtop centrifuge.
7. Vacuum centrifuge.
8. Incubator, preferably benchtop and shaking.
9. Microcentrifuge tubes: 1.7 or 0.65 mL SafeSeal, thin-walled PCR tubes (Sorenson BioScience, Salt Lake City, UT, USA).
10. 100 % Acetonitrile, HPLC-grade.
11. Washing: H_2O, HPLC-grade, 200 µL.
12. Destaining/washing (50–100 µL): 100 % Acetonitrile.
13. Ammonium bicarbonate (NH_4HCO_3) buffer: 100 mM NH_4HCO_3, in HPLC-grade H_2O.
14. Reducing solution: 10 mM DTT, 100 mM NH_4HCO_3, in MQ-H_2O.
15. Alkylating solution: 55 mM (10.2 mg/mL) Iodoacetamide, 100 mM NH_4HCO_3, in MQ-H_2O.
16. Digesting buffer: 50 mM NH_4HCO_3.
17. Digesting solution with trypsin: 12.5 ng/µL Trypsin, TPCK-treated, sequencing grade, modified trypsin [Stratagene (La Jolla, CA, USA)], in 50 mM NH_4HCO_3.
18. Resuspension solution (20 µL per sample).
    For matrix-assisted laser desorption ionization mass spectrometry (MALDI-MS):
    - 0.1 % (v/v) Triflouroacetic acid (Rathburn, Walkerburn, UK), in HPLC-grade H_2O.

    For electrospray ionization mass spectrometry (ESI)-MS:
    - 1 % (v/v) HPLC or analytical-grade Formic acid, in HPLC-grade H_2O.

2.5 Sample Preparation for MALDI Mass Spectrometry

1. C_{18} Stage-tips (Proxeon Biosystems, Odense, Denmark) or C_{18} Zip-tips (Eppendorf AG, Hamburg, Germany) or POROS 20 R2 (Applied Biosystems, Foster City, CA, USA).
2. GeLoader tips (Eppendorf AG, Hamburg, Germany).
3. Plastic syringe for micropurification.
4. KimWipes.
5. Washing buffer for MALDI-MS: 0.1 % (v/v) Trifluoroacetic acid, in HPLC-grade H_2O.

6. Washing buffer for ESI-MS: 1 % (v/v) HPLC or analytical-grade Formic acid, in HPLC-grade H_2O.
7. Elution—Matrix for MALDI-MS: 6 g/L α-cynano-4-hydroxycinnamic acid, 33 % (v/v) Methanol, 33 % (v/v) Acetonitrile, 0.1 % (v/v) Trifluoroacetic acid (Rathburn, Walkerburn, UK), HPLC-grade H_2O.
8. Elution for ESI-MS: 1 % (v/v) HPLC or analytical-grade Formic acid, 50 % (v/v) Methanol, in HPLC-grade H_2O.
9. Peptide standard I (Bruker, Bremen, Germany) for calibration in MALDI-MS, applied according to the manufacturers recommendation.
10. Polished steel target plate (Bruker, Bremen, Germany).

2.6 Samples and Antibodies

1. Samples may be sera or other complex biological fluids.
2. Antibodies must be polyclonal for immunoprecipitation to occur.

3 Methods

For the entire workflow, refer to Fig. 2 for extraction of precipitates prior to in-gel digestion with trypsin (modified from ref. 11). Briefly, electroimmunoprecipitation is performed in agarose gels with subsequent Coomassie-based visualization of precipitates. Precipitates of interest are excised with a scalpel from either wet or dried gel and extracted overnight or preferably for 24 h (*see* **Note 3**) in SDS-containing extraction buffer. Contents of extracts are separated using SDS-PAGE. After visualization of proteins in the SDS-PAGE gel, protein bands are excised and in-gel digested with trypsin. After in-gel digestion, samples are micropurified and prepared for mass spectrometric analysis. Alternatively, the immunoprecipitates are directly in-gel digested and then subjected to mass spectrometry analysis (right track in Fig. 2). For an estimate of total time of analysis, *see* **Note 3**.

We found no signs to indicate when direct enzymatic digestion of electroimmunoprecipitates (right track) would be successful or when antigen identification could only be accomplished using extraction (left track). It may be as simple as being related to the size of the antigen relative to the size of the antibody or it may be related to the nature and complexity of the precipitates (*see* **Note 4**). In most cases it will be necessary to apply an overnight extraction to get enough material for subsequent PMF-based antigen identification.

About 50 % of the protein applied for immunoelectrophoresis was recovered in the corresponding band in the SDS-PAGE gel. The reason for the low recovery is thought to result from the nature

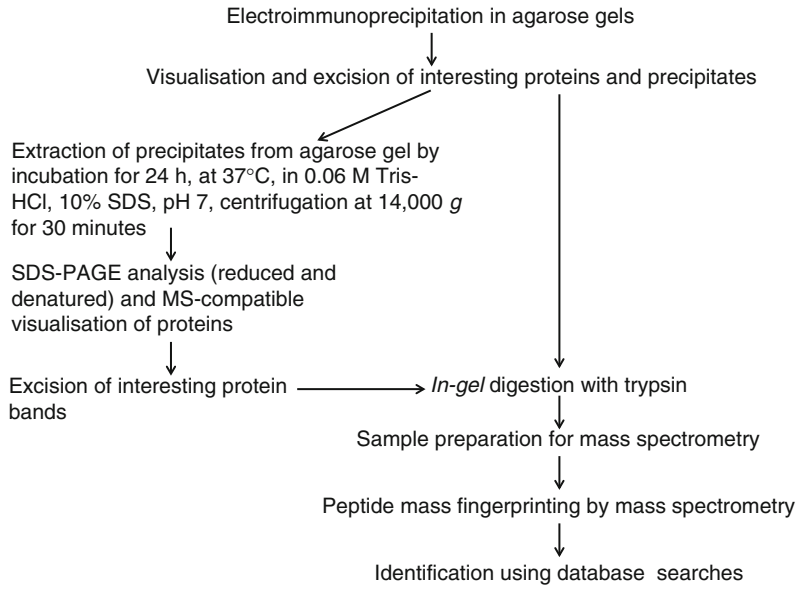


Fig. 2 The workflow involved in identification of electroimmunoprecipitated antigens. Visualized precipitates of interest are excised. Several precipitates may be pooled if the antigen is present in low amounts. Excision may be followed by extraction, SDS-PAGE analysis and in-gel digestion (*left track of figure*) or direct in-gel digestion with trypsin (*right track of figure*). After in-gel digestion, samples are prepared for peptide mass fingerprinting using mass spectrometry. Data are used for parent protein (antigen in original immunoprecipitate) identification by database searches. Reproduced with modification from Beyer et al. [11] with permission from Elsevier

of formation of the electroimmunoprecipitates making them exceedingly difficult to dissolve completely (*see* **Notes 4** and **5**).

The identification of antigens from electroimmunoprecipitates from CIE demanded more material compared to RIE antigen identification, which may result from overlapping of the precipitates. This could perhaps be overcome by the use of crossed line electroimmunoprecipitation using intermediate gels or other CIE modifications [1, 8].

3.1 Immunoelectrophoresis

3.1.1 Casting of 1 % (w/v) Agarose Gels

1. Wash glass plates by hand for two cycles in detergent and rinse with tap water.
2. Rinse in ethanol.
3. Wipe dry with KimWipes or other lint-free paper towels.
4. Place dry glass plates or Gelbond film (hydrophilic side facing upwards) on the levelling table.
5. Add 2.5 g of agarose to 250 mL working dilution of electrophoresis buffer.
6. Mix by stirring on a heating plate until boiling and the agarose solution is homogenous and clear.

7. Place in water bath at 56 °C. The solution is ready to use when the gel solution temperature is 56 °C. Gel solution stability is 3 months at 4 °C or a maximum of three cycles of boiling for use.
8. Add the proper volume of agarose gel (0.2 mL agarose/cm^2), ready to use (i.e., max. 56 °C) to a tube equilibrated to the temperature of the water bath (*see* Subheading 2.1.1).
9. Add antibodies to the gel solution, e.g., 1/100 dilution (v/v, 2 μL/cm^2) depending on the antibody and sample.
10. Mix by gently turning tube upside-down 2–3 times.
11. Carefully pour agarose-antibody gel onto the plate.
12. The agarose gel must be evenly distributed.
13. Eliminate air bubbles by a gentle touch of a pipette tip.
14. Leave for gelation.
15. After gelation, place in humidity box at 4 °C for a maximum of 24 h if not used for analysis right after gelation. Preferably, the gel should be used immediately after casting.

3.1.2 Sample Application and Running of Gel

1. Place the agarose gel on template for sample application wells.
2. Punch out sample application wells with sucking gel puncher. Punchers are used in diameters of 2.0, 2.5, 3.0, and 4.0 mm corresponding to sample volumes of 1–15 μL in gels of 1.5 mm thickness.
3. Dilute samples in electrophoresis buffer as appropriate for the volume of the punched hole in the gel.
4. Apply samples into the application wells by pipetting.
5. Add electrophoresis buffer to buffer chambers, 1 L per chamber.
6. Wet a total of ten filter paper wicks (Whatman No. 1 paper) cut to the same width as the immunoelectrophoresis plates (for even distribution of voltage) in electrophoresis buffer, five pieces in each buffer chamber.
7. Place gels on the electrophoresis apparatus.
8. Gently place the 2× five filter paper wicks on 1 cm of the top and bottom of the gel. Ensure even contact between gel and filter paper wick.
9. For placing of gels and adjustment of voltage prior to immunoelectrophoresis:
   - Place electrophoresis lid and electrodes and turn on voltage to adjust to the proper field strength (usually 2–3 V/cm, corresponding to a total of 70–110 V; *see* Subheading 3.1.2, **step 10**). Ensure right polarity (movement usually toward the anode (+)).

- Once the appropriate voltage for the experiment is determined, turn off power supply, remove the electrophoresis lid and electrodes. Place a glass plate bridging the wicks and gel to ensure proper placing of gel and wicks and to prevent condensation of water on the gel. Field strength cannot be measured after the glass plate is in place. Reassemble the electrophoresis chamber by placing the lid and electrodes again.

10. Turn on voltage as determined above and perform immunoelectrophoresis.
    - For rocket immunoelectrophoresis (RIE) according to ref. 1: 2 V/cm for 16 h at 10 °C.
    - For crossed immunoelectrophoresis (CIE) according to ref. 2 performed in two dimensions:
      - First dimension: 10 V/cm for about 1 h or until a bromophenol blue-labelled albumin marker has migrated about 4.5 cm from the application well.
      - After running the first dimension, cut out the gel lane containing the first dimension separation and place it across the top of a second antibody-containing agarose gel. Allow the gels to set before performing the second dimension separation at 2 V/cm for 16 h at 10 °C.
11. Stop immunoelectrophoresis by turning off the power supply and remove gels.

3.1.3 Washing and Pressing of Agarose Gels (Pressing Times Are Not Critical)

1. Press gels for 20 min by squirting with MQ-H_2O, and layering Whatman No. 1 paper (no air bubbles), soft absorbent tissue paper and a heavy glass plate. Pressing increases removal of non-precipitated proteins and thus minimizes background staining of the gel.
2. Wash gels for 20 min in PBS or isotonic NaCl.
3. Press gels for 20 min as above.
4. Wash gels for 20 min in MQ-H_2O.
5. Press gels for 20 min as above.
6. Wash gels for 20 min in MQ-H_2O.
7. Press gels for 20 min as above.
8. Dry in a stream of hot air.

3.1.4 Staining and Destaining of Agarose Gels

1. Stain for 1 h using aqueous CBB stain.
2. Destain in water until the background gel without precipitates appears completely clear (*see* Fig. 3).

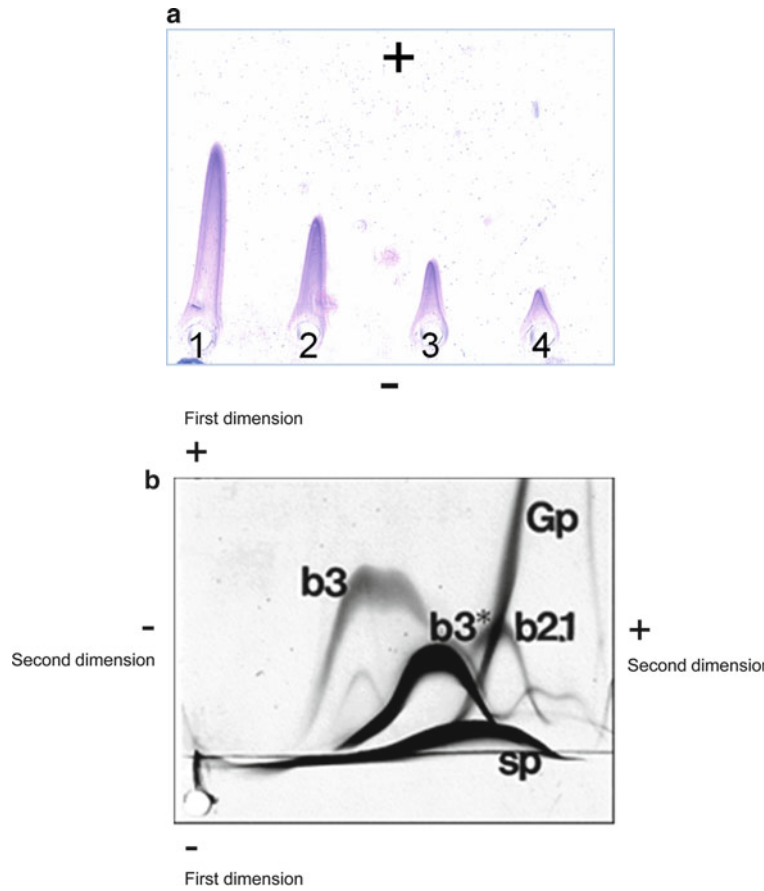


Fig. 3 (**a**) Standard CBB-staining of transthyretin/anti-transthyretin rocket immunoelectrophoretic precipitates. Antibody dilution was 1:150, applied serum was a dilution-series 1:20 (1), 1:40 (2), 1:80 (3), and 1:160 (4), according to ref. 11. Reproduced from Beyer et al. [11] with permission from Elsevier. (**b**) Crossed immunoelectrophoresis of 30 mg Triton X-100-solublized human erythrocyte membrane proteins. The second dimension gel contains 7 mL/cm^2 of a polyclonal rabbit anti-erytherocyte membrane protein antibody. b2.1, band 2.1 protein (ankyrin); b3/b3*, band 3 protein (anion transporter, free and (*) in complex with ankyrin); *Gp* glycophorin, *sp* spectrins [14]

OR

1. Stain for 15 min using standard CBB stain.
2. Destain in standard destaining solution until the background gel without precipitates appears completely clear (*see* Fig. 3).
3. Scan the gel in ethanol-wiped plastic wrap using a gel scanner or similar equipment.

3.2 Extraction of Immunoprecipitates

1. Find IE gel, RIE or CIE, new or up to several years old (*see* **Note 6**).
2. Excise precipitates using a clean scalpel.

3. Place excised gel material in a microcentrifuge tube.
4. Incubate with 50–100 µL of extraction buffer at 37 °C for about 16–24 h (*see* **Note 7**).
5. Centrifuge for 30 min at 14,000 × *g*, according to [11].
6. Transfer supernatants to clean tubes.

3.3 SDS-PAGE

1. Make Tris-glycine running buffer.
2. Place 1 or 2 of the 4–20 % Tris-glycine gels in the XCell Surelock mini-cell.
3. Add Running buffer (600–800 mL).
4. Reduce 20 µL of the electroimmunoprecipitate extracts with 5 µL of reducing sample buffer.
5. Denature by heating at 100 °C for 5 min.
6. Load 15 µL samples in the gel wells.
7. Load molecular weight standards for SDS-PAGE if required.
8. Run electrophoresis at 120–150 V for 1–2 h according to ref. 15.
9. Stop by turning off power supply.
10. Visualize protein bands by mass spectrometry-compatible staining, Coomassie Brilliant Blue staining, mass spectrometry-compatible silver staining ([13] or other staining (*see* Fig. 4)).
11. Scan the gel using a gel scanner or similar equipment.

3.4 In-Gel Digestion with Trypsin According to Ref. 13

3.4.1 Washing

1. Excise protein bands or electroimmunoprecipitates of interest and place in microcentrifuge tubes.
2. Add 200 µL HPLC-H_2O.
3. Mix for 10–15 min.
4. Remove liquid from gel pieces.
5. Add 50 µL of 100 % acetonitrile.
6. Mix for 5–15 min (until gel-plugs are white).
7. Remove liquid from gel pieces.
8. Dry down in a vacuum centrifuge (open lids) for a maximum of 30 min.
9. Turn on the heating block to 56 °C if reduction and alkylation is necessary.

3.4.2 Reduction and Alkylation

1. Add 50 µL of 10 mM DTT or DTE in 100 mM NH_4HCO_3 to previously silver-stained protein bands of interest.

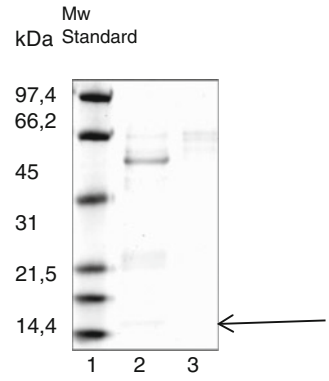


Fig. 4 Mass spectrometric compatible silver-stained SDS-PAGE of extracted proteins from immunoelectrophoretically produced precipitates from Fig. 2a. *Lane 1*—Protein Low Molecular Weight Standard, *lane 2*—Extract of precipitates, *lane 3*—Blank from agarose gel. *Arrow* indicates the transthyretin protein band. Reproduced from Beyer et al. [11] with permission from Elsevier

2. Place in heating block at 56 °C and leave for 45 min.
3. Cool to room temperature and remove liquid.
4. Add 50 µL of 55 mM iodoacetamide in 100 mM NH_4HCO_3 to the microcentrifuge tube. Leave in dark for 30 min.
5. Remove liquid.
6. Add 10 µL of 100 % acetonitrile (washing step).
7. Mix for 5–15 min (until gel-plugs are white).
8. Remove liquid.
9. Add 15–25 µL of 100 mM NH_4HCO_3 to the microcentrifuge tube.
10. Mix for 5 min.
11. Remove liquid.
12. Add 30 µL of 100 % acetonitrile.
13. Mix for 5–15 min (until gel-plugs are white).
14. Remove liquid.
15. Dry down in a vacuum centrifuge (open lids) for a maximum of 30 min.

3.4.3 Digestion

For bands excised from silver-stained or Coomassie Brilliant Blue-stained SDS-PAGE gels:

1. Add 20–30 µL of digestion solution with trypsin. Gel-plugs must just be covered with liquid.

For electroimmunoprecipitates:

1. Add 1 μg of trypsin per sample in digestion buffer. Gel-plugs must just be covered with liquid.

For protein bands from SDS-PAGE or electroimmunoprecipitates:

1. Leave on ice for a minimum of 45 min but up to 2 h to ensure that trypsin is evenly distributed in the gel-plug before digestion begins.
2. Remove excess liquid.
3. Add 20–30 μL or enough digestion buffer for the gel-plugs to be covered.
4. Leave samples at 37 °C for 4–16 h to digest.
5. Pipette supernatant to a new microcentrifuge tube for further analysis.

3.5 Sample Preparation for Mass Spectrometry

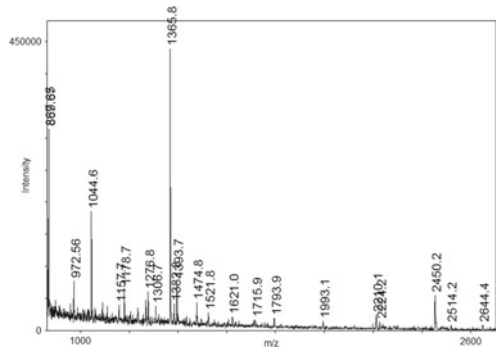
Micropurification of peptides is performed according to refs. 16, 17. Columns may be packed in the laboratory or be bought premade (Stage-tips (C18) or Zip-tips (C18)). Packing of columns in the laboratory may done by restricting GeLoader tips and applying a suspension of POROS R20 matrix into the restricted GeLoader tips until a small column is formed (*see* refs. 16, 17).

1. Activate and equilibrate the microcolumn, the Stage-tips (C18) or Zip-tips (C18) according to the manufacturer's recommendations. Activate columns made in the laboratory with 50 % methanol in 0.1 % triflouroacetic acid or 5 % formic acid, followed by washing in 0.1 % triflouroacetic acid, or 5 % formic acid. They are then ready to use for micropurification of samples.
2. To micropurify, apply the sample onto the microcolumn, and wash with 10–30 μL of 0.1 % triflouroacetic acid or 5 % formic acid.
3. After washing, elute sample peptides as described below.
4. For *MALDI-MS*: Elute samples with 0.8 μL of matrix (α-cynano-4-hydroxycinnamic acid, 6 g/L in 33 % methanol, 33 % acetonitrile, 0.1 % trifluoroacetic acid) directly onto a polished steel target plate. Leave to air-dry and analyze.

   For *ESI-MS*: Elute samples with 1 μL of 50 % methanol in 5 % formic acid for ESI analysis. Samples may also be analyzed using online LC-ESI-MS on reversed phase microcolumns and thus micropurification may omitted.

3.6 Mass Spectrometry

1. For reliable identification, use internal calibration and fragmentation.
2. Identification of peptides from PMF is based on online database searches in the latest available version of the NCBI database using the MASCOT search engine at http://www.matrixscience.com/cgi/nph-mascot.exe?1 (*see* Fig. 5).



gil2098256
Mass:13806 **Score: 74**
Expect: 0.0055
Queries matched: 6
Sequence Coverage: 74%

Chain B, Tertiary Structures Of
Three Amyloidogenic Transthyretin Variants And
Implications For Amyloid Fibril Formation

Matched peptides shown in **Bold Red**

1 **GPTGTGESKR PLMVKVLDAV RGSPAINVAV HVFRKAADDT WEPFASGK**TS
51 ESGELHGLTT EEEFVEGIYK VEIDTKSYWK **ALGISPFHEH AEVVFTANDS**
101 **GPRRYTIAAL LSPYSYSTTA VVTNPKE**

Start - End	Observed	Delta	Miss	Sequence
1 - 21	2210.1364	-0.0840	2	-.GPTGTGESKRPLMVKVLDAVR.G
22 - 34	1365.8295	0.0779	0	R.GSPAINVAVHVFR.K
35 - 48	1521.8027	0.0928	1	R.KAADDTWEPFASGK.T
36 - 48	1393.6995	0.0845	0	K.AADDTWEPFASGK.T
81 - 103	2450.2073	0.0095	0	K.ALGISPFHEHAEVVFTANDSGPR.R
104 - 127	2644.3640	-0.0108	2	R.RYTIAALLSPYSYSTTAVVTNPKE.-

Fig. 5 MALDI-MS, Peptide mass fingerprint of in-gel digested TTR from Fig. 2. Identification of the antigen TTR from one excised band from SDS-PAGE, trypsin in-gel digestion, MALDI-MS, and subsequent database searches. Match indicating 74 % sequence coverage

4 Notes

1. The size of the gel plate and consequently the immuno gel chosen for an experiment depends on the number of samples and the complexity of both the polyclonal antibody and the samples. Generally, when doing RIE of a few samples and a specific antibody against a single antigen molecule (as the RIE in Fig. 3a), small to medium size gels (5 cm × 5 cm to 10 cm × 10 cm) are often sufficient for IE precipitates of the antigen. When more than four samples are to be applied to the gel, larger gels (10 cm × 10 cm to 10 cm × 21 cm) are recommended. When more complex samples/antibodies (e.g., serum samples and anti-serum antibodies), or several dimensions are applied, medium gels may be used (5 cm × 7 cm to 10 cm × 10 cm).
2. Template for punching out sample application wells:
   Place the gelated agarose gel on the glass plate or Gelbond film on top of the template and use the sucking gel puncher to make sample application wells in the gel. Punchers typically used have diameters of 2.0, 2.5, 3.0, and 4.0 mm corresponding to sample volumes of 1–15 μL in gels of 1.5 mm thickness. The template may be applied to larger gels by replication.

3. Immunoelectrophoresis runs overnight for the best result. For extraction followed by SDS-PAGE, one day (and overnight) is used for extraction, one day for SDS-PAGE, staining and excision of protein bands. One day for start of overnight in-gel digestion and about a day for sample preparation for mass spectrometry and analysis, as this approach may result in many samples. Thus a total of 5 days should be considered as analysis time from sample to result. Data analysis depends on programs available.
4. Immunoprecipitates created by electrophoretic mixing of antigens and polyclonal antibodies are very stable structures. This is due to the high affinity and polyvalency of antibodies derived from hyperimmunized animals. Additionally, the immune complex stability is probably further enhanced by the enrichment of high affinity interactions in electrophoresis since antigen molecules binding with low affinity are removed by the electrical field when dissociating.

   This leaves very stably bound antigen covered with large antibody molecules binding to and crosslinking multiple epitopes in the final electroimmunoprecipitate. Such precipitates, especially in the dried gel, are stable indefinitely in the dark at room temperature. The accessibility of the antigen for further characterization is very limited and explains the relatively low yields upon direct trypsin-mediated digestion of electroimmunoprecipitates and the importance of optimizing immune complex dissociation when using the extraction procedure.
5. One factor is that the IgG bands on the SDS-PAGE gels were clearly stained much more intensely than the antigen bands, indicating that much more IgG than antigen was present in the extracts. This could reflect two things. If IgG and the antigen precipitate in a 1:1 molar ratio, then, in our examples, the stainability of the antibody part of the precipitate is ten times that of the antigen due to the difference in size.

   The much stronger intensity of the IgG protein bands on SDS-PAGE could also reflect that more than one IgG molecule is present per antigen molecule in electroimmunoprecipitates. This is likely with polyclonal antibodies that bind to multiple epitopes on antigen molecules.

   The success of the procedure probably varies with the strength of association between antibody and antigen. Thus, the sometimes difficult recovery is linked to the pronounced stability of the electroimmunocomplexes. It is well known that polyclonal antisera contain a spectrum of antibody affinities and electroimmunoprecipitation, as noted above, appears to select for the highest antibody affinities. Non-electrophoretic immunoprecipitation methods result in less stable immunoprecipitates as low, medium, and high affinity antibodies will all

participate in the precipitation of the antigen and since incubation times are typically on the hour scale with no separation involved. In immunoelectrophoresis, however, precipitate is formed in overnight experiments and under the continued influence of an electrical field. Thus, it would be expected that immunoprecipitates from, for example, single radial immunodiffusion experiments [19] would be easier to dissociate than the corresponding antibody–antigen system precipitated by electroimmunotechniques. Thus, the present procedure will also be readily usable for immunoprecipitates produced by other methods.

6. Both old and new electroimmunoprecipitates have been used (*see* Fig. 3). The method works even when gels have been stored for years. For example, CIE of sera from systemic lupus erythematosus (SLE) patients from an earlier study [18] where serum amyloid protein (SAP) and transthyretin (TTR) were electroimmunoprecipitated with specific antibodies [11].
7. The optimal method was to excise the CBB-stained precipitate and extract it with a minimal volume of 0.06 M Tris–HCl (pH 7), 10 % SDS, for 16–24 h at 37 °C. We found that at least 16 h of extraction is necessary. In addition, a temperature of 37 °C was better than higher temperatures, although extraction at room temperature could be investigated. Addition of ionic detergent, such as 10 % SDS (w/v) was crucial for extraction and less detergent resulted in a lower yield. Extraction solution could be either 0.06 M Tris–HCl at pH 7 or 0.1 M Glycine, but Tris–HCl was preferred for compatibility with SDS-PAGE. The step of centrifugation was also crucial. Using this method, 135 ng TTR applied to an immunoelectrophoretic gel could be identified.

   To increase sensitivity, precipitates from up to eight separate CIEs may be pooled in a single microcentrifuge tube for extraction and subsequent PMF-based antigen identification.

References

1. Axelsen NH, Bock E (1983) Electroimmunoassay (rocket immunoelectrophoresis). In: Axelsen NH (ed) Handbook of immunoprecipitation-in-gel techniques. Blackwell, Oxford, pp 103–106
2. Grubb AO (1983) Crossed immunoelectrophoresis. In: Axelsen NH (ed) Handbook of immunoprecipitation-in-gel techniques. Blackwell, Oxford, pp 113–124
3. Laurell CB (1966) Quantitative estimation of proteins by electrophoresis in agarose gel containing antibodies. Anal Biochem 15: 45–52
4. Bjerrum OJ, Hagen I (1983) Biomolecular characterization of membrane antigens. In: Bjerrum OJ (ed) Electroimmunochemical analysis of membrane proteins. Elsevier, Amsterdam, pp 112–126
5. Koch C, Skjodt K, Laursen I (1985) A simple immunoblotting method after separation of proteins in agarose gel. J Immunol Methods 84:271–278
6. Bjerrum OJ, Selmer JC, Lihme A (1987) Native immunoblotting – transfer of membrane-proteins in the presence of non-ionic detergent. Electrophoresis 8:388–397

7. Krøll J (1983) Tandem crossed immunoelectrophoresis. In: Axelsen NH (ed) Handbook of immunoprecipitation-in-gel techniques. Blackwell, Oxford, pp 135–140
8. Axelsen NH (1983) Intermediate gel immunoelectrophoresis. In: Axelsen NH (ed) Handbook of immunoprecipitation-in-gel techniques. Blackwell, Oxford, pp 141–149
9. Vandvik B, Nordal HJ, Vartdal F, Nilsen RE, Norrby E (1983) Imprint immunofixation of antibodies separated by agarose electrophoresis or by electrofocusing. In: Axelsen NH (ed) Handbook of immunoprecipitation-in-gel techniques. Blackwell, Oxford, pp 33–37
10. Norrild B, Bjerrum OJ, Vestergaard BF (1977) Polypeptide analysis of individual immunoprecipitates from crossed immunoelectrophoresis. Anal Biochem 81:432–441
11. Beyer NH, Schou C, Houen G, Heegaard NHH (2008) Extraction and identification of electroimmunoprecipitated proteins from agarose gels. J Immunol Methods 330(1–2): 24–33
12. Svendsen PJ, Weeke B, Johansson B-G (1983) Chemicals, solutions, equipment and general procedures. In: Axelsen NH (ed) Handbook of immunoprecipitation-in-gel techniques. Blackwell, Oxford, pp 3–20
13. Shevchenko A, Wilm M, Vorm O, Mann M (1996) Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. Anal Chem 68:850–858
14. Heegaard NHH (1990) Immunochemical characterization of interactions between circulating autologous IgG and normal erythrocyte membrane proteins. Biochim Biophys Acta 1023:239–246
15. Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227:680–685
16. Gobom J, Schuerenberg M, Mueller M, Theiss D, Lehrach H, Nordhoff E (2001) Alpha-cyano-4-hydroxycinnamic acid affinity sample preparation. A protocol for MALDI-MS peptide analysis in proteomics. Anal Chem 73(3):434–438
17. Kussmann M, Lässing U, Stürmer CA, Przybylski M, Roepstorff P (1997) Matrix-assisted laser desorption/ionization mass spectrometric peptide mapping of the neural cell adhesion protein neurolin purified by sodium dodecyl sulfate polyacrylamide gel electrophoresis or acidic precipitation. J Mass Spectrom 32(5):483–493
18. Sen JW, Recke C, Rahbek L, Skogstrand K, Heegaard NH (2002) Structural, quantitative and functional comparison of amyloid P component in sera from patients with systemic lupus erythematosus and healthy donors. Scand J Immunol 56:645
19. Mancini G, Carbonara AO, Heremans JF (1965) Immunochemical quantitation of antigens by single radial immunodiffusion. Immunochemistry 2:235

Chapter 8

An Immunoproteomics Approach to Screen the Antigenicity of the Influenza Virus

Kevin M. Downard

Abstract

The structure and antigenicity of protein antigens of the influenza virus are screened in a single step employing an immunoproteomics approach. Matrix-assisted laser desorption ionization mass spectrometry (MALDI-MS) coupled to gel electrophoresis is used both to identify viral antigens and screen their antigenicity. Earlier evidence that antigen–antibody complexes can survive on MALDI targets has allowed both the primary structure and antigenicity of viral strains to be rapidly screened with the specific localization of protein epitopes. The approach is anticipated to have a greater role in the future surveillance of the virus and should also aid in the development of immunogenic peptide constructs as alternatives to whole virus for vaccination.

Key words Influenza, Flu, Virus, Surveillance, Antigenicity, Mass spectrometry, Proteomics, Immunoproteomics

1 Introduction

The influenza virus is a leading cause of death resulting in the loss of some 500,000 lives every year [1, 2]. It further inflicts illness and suffering the world over and is responsible for major productive losses, economic and healthcare burdens. Despite a global surveillance strategy [3], the widespread availability and administration of vaccines against the virus [4, 5] and the development of a new generation of anti-viral drugs [6], rates of infection continue unabated. With the very real likelihood of future pandemics [7, 8], new rapid surveillance approaches [9–14] are of value to screen and evaluate the threat of circulating strains.

Humans are primarily infected with type A and B influenza. Influenza A viruses also infect birds and other mammals, while type B influenza is almost exclusively a human pathogen. Inactivated forms of common circulating strains of these types form the basis of current vaccines administered against the virus. Viral strains are further subtyped [15, 16] according to the nature of the two

Kelly M. Fulton and Susan M. Twine (eds.), *Immunoproteomics: Methods and Protocols*, Methods in Molecular Biology, vol. 1061, DOI 10.1007/978-1-62703-589-7_8,

surface protein antigens hemagglutinin (H) and neuraminidase (N). The role of the hemagglutinin antigen is to bind the virus particle to sialic acid receptors on the surface of the host cell, while the neuraminidase enzyme releases new budding virus particles that emerge from the host cells and which promote further infection.

The first line of defense offered by the immune response is afforded by antibodies [17]. These are secreted from naive or memory B-cells of the host that bind to and inactivate the viral antigens. Alterations in the sequence and structure of the viral antigens, as a consequence of antigenic drift and shift, help the virus to elude detection and inactivation. These changes are associated with errors produced during the replication of the virus, that lead to mutations in the encoding genes, and the trading of genetic material from one strain to another when humans are infected simultaneously with multiple forms. As humans have no immunity to a strain sufficiently diverged from that to which they were exposed, through natural infection or vaccination, such antigenic drifts and shifts can lead to a local epidemic and even a global pandemic [18].

A worldwide, influenza surveillance strategy was launched by the World Health Organization in 1952 [3]. The primary goal of this international initiative is to identify and characterize emerging strains of the virus in humans and some animals and recommend the composition of the annual vaccine in the northern and southern hemispheres [3]. Effective vaccines rely on a close match between the current vaccine composition and the antigenicity of circulating strains. As little as a single point mutation in the gene sequence encoding hemagglutinin can result in an existing vaccine being ineffective [19]. The surveillance of the virus in some animals and birds is designed to complement the human studies and aid in the understanding of the ecology of the influenza strains of relevance to human health. More than 175,000 patient samples are collected annually of which approximately 2,000 viruses are submitted to the WHO collaborating centers for antigenic and genetic analyses [3].

The hemagglutination inhibition (HI) assay is the most common assay employed to screen viral isolates [20]. Viruses are first grown in the allantoic fluid of chicken eggs prior to analysis. The inhibition of hemagglutination employs a standardized quantity of the hemagglutinin antigen. The virus is mixed with a suspension of red blood cells in a solution of antisera raised to reference strains. The onset of hemagglutination inhibition is identified by the buttoning of red blood cells in the base of a well of the microtiter plate. This results from the attachment of antibody to the hemagglutinin protein that, in turn, prevents the virus from binding with receptors of the erythrocytes.

The HI assay provides no molecular detail. Changes in the sequence of the surface antigens, therefore, are probed employing genetic approaches. Since the virus genome is single-stranded

RNA, a DNA copy (cDNA) must be prepared using a reverse transcriptase prior to PCR amplification and hybridization sequencing experiments. Microarrays of oligonucleotide probes [21] to known subtypes of the hemagglutinin and neuraminidase antigens have recently been developed to screen viral isolates. Where a new or sufficiently diverged strain is identified, reverse transcriptase-polymerase chain reaction (RT-PCR) sequencing is performed [22]. A primer specific for the hemagglutinin (HA) gene is annealed to the denatured RNA template and extended with reverse transcriptase to synthesize cDNA. After cDNA synthesis, an aliquot of the RT reaction is used for the PCR sequencing cycle. Separation of the PCR products enables the gene sequence to be read [23]. Although automated, genotyping of the influenza virus takes some hours to days to perform.

There are merits in developing an immunoproteomics approach in which sequence differences among surface antigens are characterized first hand. The use of mass spectrometry to screen both the primary structure *and* antigenicity of the virus in a single step in terms of the protein antigens to achieve this was first reported in 1999 [9, 10]. The approach [9], originally applied to whole virus [10], was later advanced for gel-purified antigens [11, 12] in order to improve sequence coverage and the likelihood of epitope identification (Fig. 1). Important to the success of the approach is the realization that peptides representing an antigenic determinant can remain bound to antibody throughout immunological treatment, MALDI sample deposition and ablation [9–13], and mass spectral analysis [14] providing certain experimental conditions are implemented. The antigenicity of component antigens can then be assessed by a comparison of a pair of their MALDI mass spectra [9–12] before and after their treatment with monoclonal antibodies. The methods involved in this immunoproteomics approach [11–13] are described herein both to highlight the advantages of the method and to outline the experimental protocols.

2 Materials

2.1 Viral Isolates

Inactivated viruses were obtained from commercial sources or the World Health Organization Collaborating Centre for Reference and Research on Influenza, Melbourne. Viruses were inactivated with 0.005 % merthiolate and purified by ultracentrifugation using a 10–40 % sucrose gradient (*see* **Note 1**).

2.2 Monoclonal Antibodies

Monoclonal antibodies obtained commercially were derived from the hybridization of myeloma and spleen cells of mice immunized with the relevant serotype or prepared in a miniPERM bioreactor in serum-free medium (CSL Limited, Melbourne).

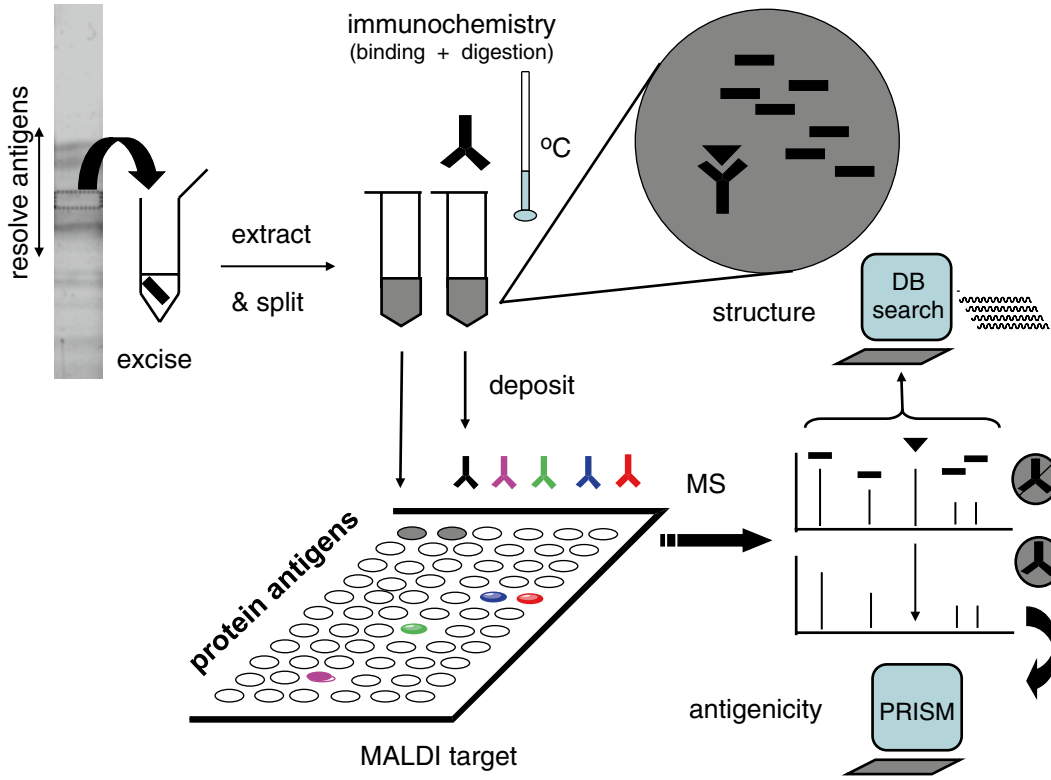


Fig. 1 Schematic representation of the immunoproteomics approach to survey both the primary structure and antigenicity of antigens of the influenza virus. Gel-resolved antigens are treated with monoclonal antibodies prior to or following proteolytic digestion. The products are analyzed directly by MALDI mass spectrometry. Antigens are identified through searches of general or specialized flu databases [29] with the mass fingerprint data. The PRISM algorithm [24] aids in the identification of epitopic peptides that bind antibody

2.3 Gel Electrophoresis

1. Acrylamide.
2. Ammonium persulfate.
3. Gel running buffer: 25 mM Tris–HCl, 192 mM glycine, pH 8.3 with 0.1 % (w/v) sodium dodecylsulfate added for denaturing gels.
4. Mini gel box apparatus and power supply.
5. Gel stain solution: 0.25 % (w/v) Coomassie blue (R-250) in 53 % (v/v) water, 40 % (v/v) methanol, and 7 % (v/v) acetic acid.
6. Destain solution: 50 % (v/v) acetonitrile and 10–50 mM ammonium bicarbonate.
7. Digital gel scanner.

2.4 Enzymatic Digestion and Peptide Recovery

1. Buffer solution: 50 mM ammonium bicarbonate, pH 8.
2. Sequence-grade protease (e.g., trypsin).
3. Peptide-N_4-(acetyl-β-glucosaminyl)-asparagine amidase (PNGase F).

4. Trifluoroacetic acid.
5. Bath sonicator.
6. Centrifugal vacuum concentrator.

2.5 Immunochemistry

1. Phosphate-buffered saline (PBS) buffer: 20 mM phosphate, 100 mM NaCl, pH 7.8.

2.6 Mass Spectrometry

1. MALDI-based mass spectrometer (*see* **Note 2**).
2. Matrix solution: saturated solution of α-cyano-4-hydroxycinnamic acid in 70 % (v/v) acetonitrile.

2.7 Data Analysis

1. Influenza sequence database (ISD), Los Alamos National Laboratory, http://www.flu.lanl.gov.
2. PRotein Interactions from the Spectra of Masses (PRISM software [24]) This software is available for not-for-profit institutions from the author upon request and signing a license agreement.

3 Methods

3.1 Gel Electrophoresis

The inclusion of a stage of separation of viral antigens [11–13], over the use of whole virus [9], improves the sequence coverage achieved in the MALDI mass map [25]. The presence of additional antigens and components can result in the overlap or suppression of ion signals for a particular antigen. This impacts on the ability to correctly identify the antigen, establish differences in its sequence from a like-antigen of a reference strain, and follow the binding of antigen to antibody by monitoring the reduction in an epitopic peptide signal relative to nonbinding peptides.

In order to identify and characterize discontinuous epitopes, antigen separation is achieved using native gel electrophoresis without the use of sodium dodecylsulfate. Gradient gels afford better separations for this application. Protein antigen bands are subsequently identified using a sacrificial stain-aligned band in order to prevent the denaturation of recovered antigen with the stain, and during the destaining of bands in organic solvents. Whole antigens are recovered from the native gel using a protocol developed for this purpose that employs bath sonication [26]. This contrasts with an alternate second approach that can be used for the detection of linear, continuous epitopes. In this instance, protein antigens can be digested in gel following their separation by SDS-PAGE and subsequently treated with monoclonal antibodies ahead of mass spectrometric analysis.

1. Solutions of influenza strains (30 µg of total virus at 1 µg/µL) in gel buffer were loaded in separate lanes of a commercial or cast 8–17.5 % polyacrylamide gel and run with gel buffer, in

the presence or absence of 0.1 % w/v of sodium dodecylsulfate, for 2–3 h at 150 V (*see* **Note 1**).

2. Gels, or a single reference lane of the gel, were stained in stain solution overnight and subsequently destained for 2 h. Despite the electrophoretic separation stage, some bands may still contain more than one antigen particularly where native gel electrophoresis is employed. Even in the case of denaturing polyacrylamide gel electrophoresis (PAGE), the hemagglutinin and nucleoprotein antigens of type H1N1 strains share similar molecular weights and are found to migrate to similar positions on the gel. Nonetheless, even the incomplete purification of antigens from whole virus improves their sequence coverage in the mass map that subsequently aids in the identification of epitopic domains.

3.2 Immunobinding and Proteolytic Digestion

1. The bands containing the viral antigens hemagglutinin or neuraminidase were excised, cut into small pieces and transferred into separate prewashed (with acetonitrile) tubes for extraction of the protein either before or after digestion (*see* **Note 1**).
2. Whole antigens were recovered by sonication [26] and with the aid of trifluoroacetic acid (1 % by volume) in the case of in-gel digested protein (*see* **Note 3**). The gel pieces were sonicated in a bath sonicator for 3×15 min with the samples repeatedly cooled over ice to prevent heating of the solution. Protein recoveries of between 60 and 80 % have been achieved for model proteins [26].
3. The integrity of the extracted protein structures following this treatment was confirmed by electrospray ionization mass spectrometry (ESI-MS) and circular dichoism (CD) spectroscopy [26].
4. Antigens were treated with monoclonal antibodies prior to [27] or after enzymatic digestion subject to the nature of the predicted epitope. Treatment with monoclone precedes limited digestion of the immune complex [27] where the nature of the epitope is unknown. Solutions of monoclonal antibodies in PBS buffer were treated with antigen at a 2:1 antibody to antigen mole ratio for a period of 24 h at 4 °C, following antigen recovery from a native gel [26] or after its digestion in gel (*see* **Note 4**).
5. After resuspension in ammonium bicarbonate buffer, sequence-grade endoproteinase (e.g., trypsin) is added at an enzyme to substrate ratio of 1:50 to effect the limited proteolysis of antibody–antigen complexes, or digestion of antigen ahead of antibody treatment, for a period of 15 h at 37 °C.

3.3 Mass Spectrometry and MS Data Analysis

1. A portion of the antibody reaction mixture and untreated control (1 μL) was diluted in a saturated solution of α-cyano-4-hydroxycinnamic acid (3 μL) in acetonitrile and water

(70:30 v/v) and deposited onto the sample stage (1 μL per spot) (*see* **Note 5–7**).

2. The samples were analyzed by MALDI-MS using a 337 nm nitrogen laser operating at approximately 10 μJ (*see* **Note 2**) employing time-delayed ion extraction (delay time ~10 μs) on a linear or reflecting time-of-flight mass analyzer [28] (*see* **Note 2**). The MALDI mass spectra are acquired in seconds affording a significant advantage over the time required for analysis in other analytical and spectroscopic approaches. The resulting spectra, each acquired from an average of some 100 laser shots from across the target in order to ensure the representative nature of the peptides within the sample and also improve ion statistics, are compared manually and/or with the aid of a computer program developed in-house for this purpose [24]. Known as PRISM, as it identifies PRotein Interactions from the Spectra of Masses, the algorithm features an easy to use, intuitive graphical user interface (*see* Subheading 2.7).
3. The identification of protein epitopes or determinants is established based on the selective reduction in the relative area of their ion signals of the spectra recorded for the no-antibody control sample versus the antibody-treated sample. An absolute reduction of at least 10–15 % in relative area is required to establish the identity of an antigenic peptide(s) in order to compensate for experimental errors associated with the variation of mass spectra obtained from different samples (*see* **Note 8**). The data is input as a list of mass or *m/z* values for the peptide ions and their abundances. Each pair of spectra is processed to remove any *m/z* values of ions found in only one spectrum (within a specified *m/z* error; default = 0.1) and that are of low signal-to-noise relative to the most abundant base peak (default 1 %). Spectra of the antigen-only sample are compared with *m/z* values for proteolytic peptides predicted for the protein based on translated gene or protein sequences derived from the Universal Protein Database (UniProt) or ISD [29]. The algorithm then compares changes in the area under each of the ion signals relative to a constant peak. The constant peak is determined by first measuring the ratio of areas of all adjacent peaks $n + 1/n$ (in terms of *m/z*) in both spectra and establishing the ion peak *n* for which the absolute difference is the smallest. The value for the relative area of the constant peak to the most abundant base peak in the control (no-antibody) spectrum is then used to establish an imaginary peak in the spectrum of the antibody-treated sample. This is achieved such that the relative area of the constant peak to the base peak in the control spectrum is equal to the ratio of area of the constant peak to the imaginary peak in the spectrum of the antibody-treated sample. This establishes a common reference peak with which

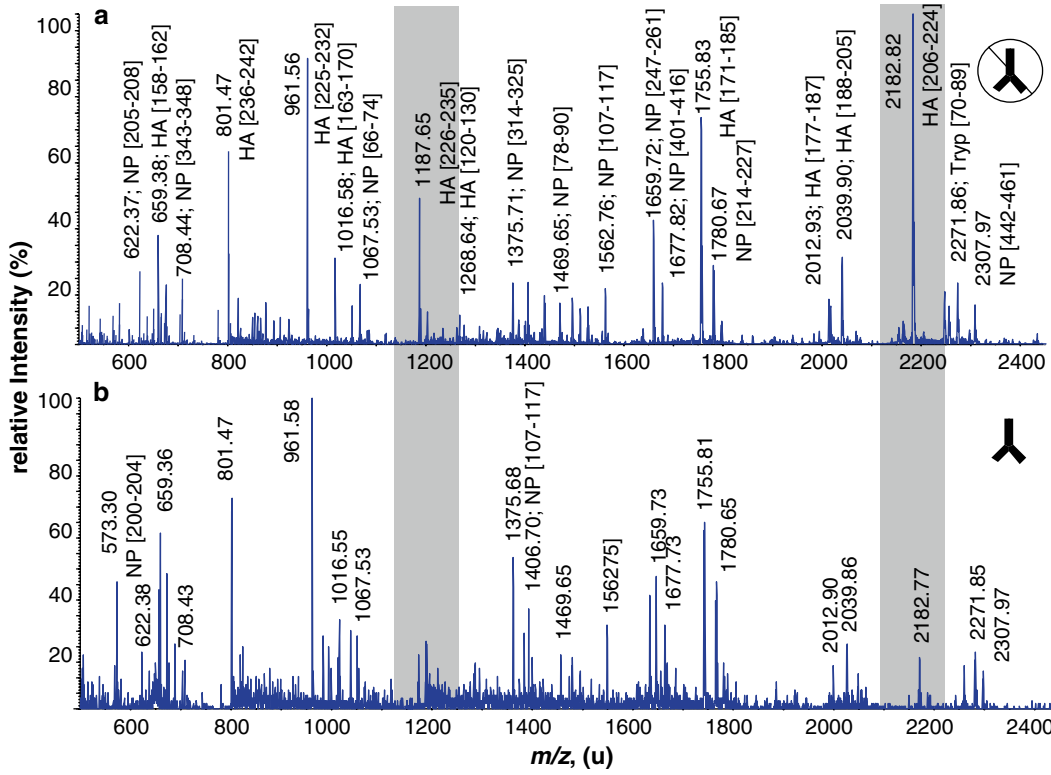


Fig. 2 MALDI mass spectra of gel-recovered hemagglutinin after tryptic digestion of the Beijing 262/95 type A strain (**a**) without antibody, and (**b**) after 24 h incubation with a monoclonal antibody raised to an H1N1 serotype [11]

to compare relative areas across the two different mass spectra. Areas under the ion signals are computed relative to the base peak in spectrum of the control and relative to the imaginary peak within the spectrum of the antibody-treated sample. Absolute changes in relative area among common ions in both spectra that are greater in value than 10 % (default value) to 15 % represent possible antibody-binding peptides. The *m/z* values for these ions are output together with the identity of the constant peak to enable the spectra to be independently assessed.

4. Replicate experiments from common and diverged strains are used to verify the results. The latter also helps establish critical binding residues where differences among antigen sequences within an epitope may impact on the ability of the peptide to bind monoclone.

3.4 Application

To illustrate a typical set of mass spectra and their analysis, MALDI mass spectral data for the gel-recovered hemagglutinin of H1N1 influenza strain (Beijing 262/95) before and after treatment with a monoclonal antibody raised the antigen of this serotype are shown in Fig. 2 [11]. Note that recovered antigen also contains

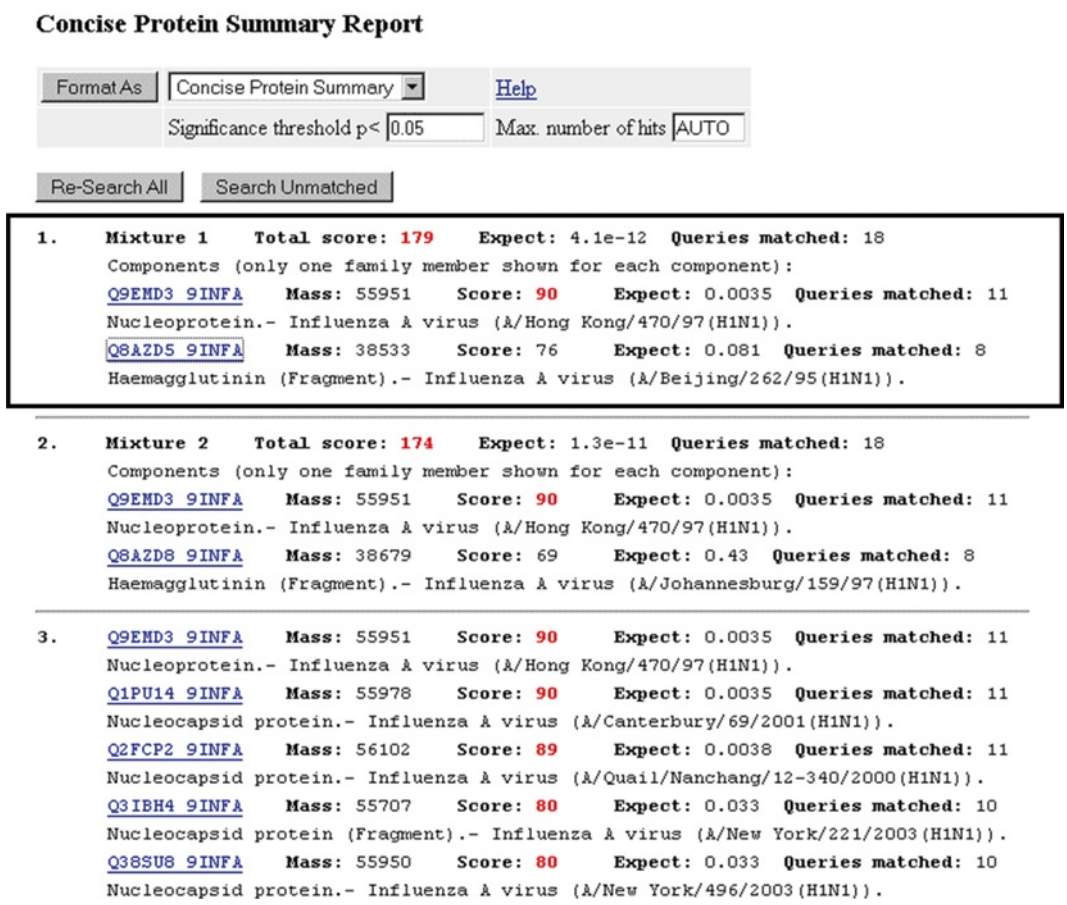


Fig. 3 Mascot peptide mass fingerprint search output for the three highest scored protein candidates obtained from a search of the MSDB database with the mass spectral data from Fig. 2a. The MSDB database is a composite, non-identical protein sequence database compiled from entries of the PIR, Trembl, GenBank, Swiss-Prot, and NRL3D databases at Imperial College, London

some nucleoprotein (NP) that was not completely separated by gel electrophoresis. Each spectrum provides a mass "fingerprint" or map of the viral antigen(s) that can be used in its own right to search protein databases [25] to confirm or identify differences in the sequences of the component antigens. A Mascot peptide mass fingerprint search of the MSDB nonredundant database correctly establishes the identity of the antigens in the band as a mixture of both nucleoprotein (of a diverged Hong Kong H1N1 strain) and the hemagglutinin antigen (Fig. 3). Differences in amino acid sequence can easily be confirmed by generating tandem (or MS/MS) spectra for any or all other peptides including epitopic peptides [9–12]. On most mass spectrometers, this can be performed in an automated manner where ions of a specified m/z value, or those above a signal threshold, are isolated in the first mass analyzer and subjected to dissociation in MS/MS experiments [30].

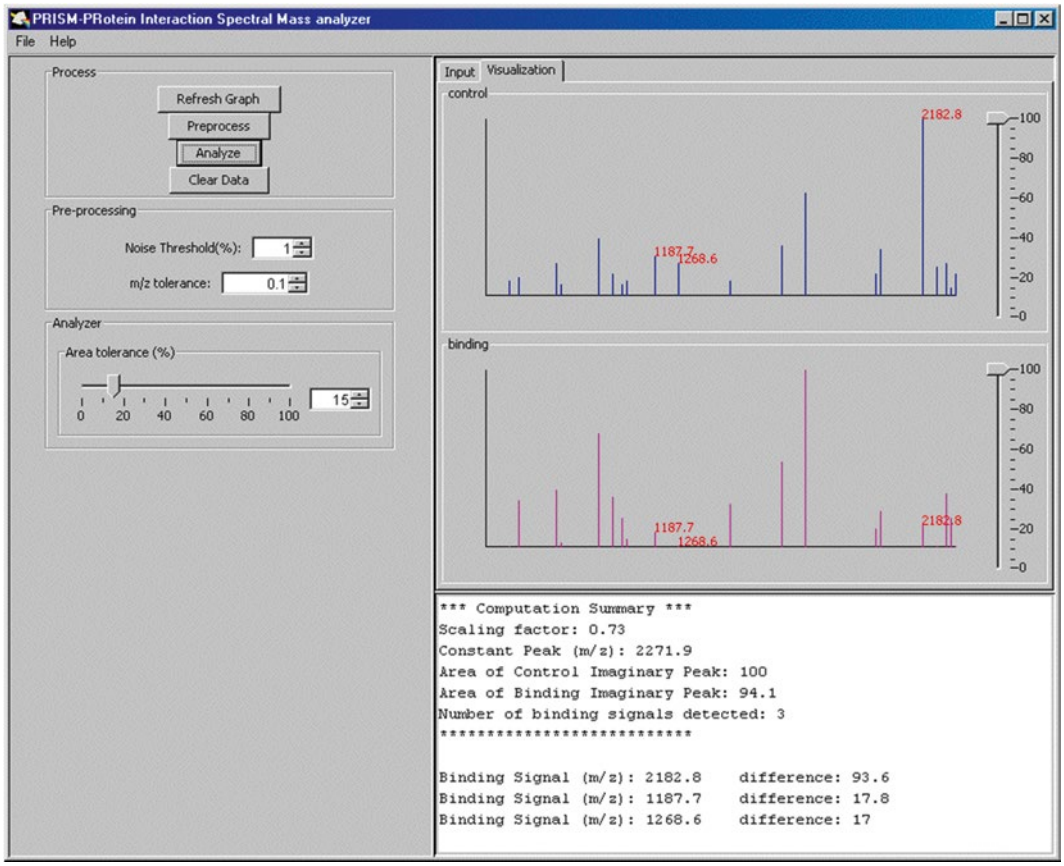


Fig. 4 Graphical user interface (GUI) of the PRISM algorithm [24] showing the output for the mass spectral data shown in Fig. 2

Despite an incomplete separation of the antigens by gel electrophoresis, the peptide ion comprising residues 206–224 of hemagglutinin at *m/z* 2,182.8 is seen to significantly reduce in its relative area (94 % absolute change) in the presence of antibody due to its interaction with it and the preservation of the peptide-complex on the MALDI target. All other peptide ion signals remain relatively unchanged within experimental area (some ≤15 %) with the exception of the peptide ions at *m/z* 1,187.65. This peptide corresponds to residues 226–235 of the hemagglutinin antigen that flanks the C-terminal end of the epitope and also binds to a lesser degree. The location of the epitope is in accord with earlier results obtained for a diverged H1N1 strain using whole virus digests [10].

The ability of the PRISM algorithm to identify the epitopic peptide from such pairs of mass spectra without human intervention is illustrated in Fig. 4. The two mass spectra shown in Fig. 2 are input into the algorithm as *m/z* values versus peak area data. After noise reduction and the removal of ions not common to both spectra, the algorithm compares the relative area of ions across both

spectra using the approach described above. Using a cut-off of 15 % (*see* **Note 8**), only three peptide ion signals are found whose relative areas change by a level that exceeds this value. These include the two peptides identified above, validating the performance of the algorithm, and an additional peptide comprising residues at 120–130 of the hemagglutinin antigen (*M* 1268.6) which shares sequence homology with the epitopic peptide at *m/z* 2,182.82 [11].

The successful application of this immunoproteomics approach to this [11] and other type A H1N1 strains has been shown for other strains of the virus including three diverged H3N2 type A strains [12] and two type B strains. It produces results in accord with traditional hemagglutination inhibition assays [31].

The approach has been extended to examine the rates of binding of epitopic peptides with antibodies in time-course experiments [32]. Furthermore, it has recently been adapted to study the binding of anti-viral inhibitors to influenza neuraminidase [33].

4 Notes

1. Deactivation of live or attenuated strains of the influenza virus using gamma radiation or formalin treatment should be performed to enable safe handling of specimens [34]. Other containment procedures should be followed in accordance with regulations within each laboratory and its jurisdiction.
2. Any MALDI-based mass spectrometer from TOF, hybrid-TOF, and other scanning mass analyzers can be employed in this approach where appropriate consideration is given in the case of the latter to analyzer scan rates and their dynamic mass range.
3. Endoproteinases should be chosen with care in order to effect site-specific limited digestion of the antigen–antibody complex or antigen alone so as to refine epitopic domains without cleaving epitopic peptides. It has been shown in one study [9] that where digestion was performed with chymotrypsin a linear epitope (identified following a tryptic digestion) within H1 hemagglutinin was cleaved and no binding peptides were detected. If no epitopic peptides are detected in one experiment, a second protease should be chosen to assess this issue.
4. Proteolytic digestion should aim to achieve the maximum antigen sequence coverage in order to follow changes in the relative area of ion signals for peptides across the entire protein. The use of two experiments with two different endoproteases can assist with this goal. In practice, 100 % coverage is rarely achieved even for highly purified protein standards in mass mapping experiments [25].

5. Laser powers should be kept as low as possible in order not to disrupt peptide–antibody complexes on the MALDI target. The higher extraction potentials required to directly detect immune complexes [14] are not required where such complexes are preserved on the target surface.
6. The spectra acquired should be representative of the total peptide composition on the MALDI target. In this respect, the plated sample should be as homogeneous as possible and/or sample should be ablated from across the surface by moving the laser position or target.
7. MALDI matrix additives such as acetic acid and trifluoroacetic acid must be avoided since these have been shown to induce dissociation of peptide–antibody complexes leading to the detection of epitopic peptides in the antibody-treated sample [9].
8. A 10–15 % absolute change in the relative area of ion signals has been demonstrated to be an appropriate cut-off value to efforts to detect binding peptides in order to allow for experimental fluctuations that cause slight ion intensity variations in mass spectral data for different samples [9–12].

Acknowledgments

The author acknowledges the contributions of Bethny Morrissey, Dr. Alexander Schwahn, Dr. Margaret Streamer, and Joshua Ho as indicated by the cited works. The author thanks Robert Shaw and Dr. Ian Barr of the WHO Collaborating Centre for Reference and Research on Influenza and Dr. Elizabeth Pietrzykowski of CSL Limited for the supply of some virus strains and monoclonal antibodies, respectively. Funding for this work was provided by the Australian Research Council (DP0449800 and DP0770619 to the author) and the University of Sydney.

References

1. Nicholson KG, Wood JM, Zambon M (2003) Influenza. Lancet 362:1733–1745
2. Barry JM (2004) The great influenza: the epic story of the deadliest plague in history. Penguin Viking, New York
3. World Health Organisation Global Influenza Surveillance Network. http://www.who.int/csr/disease/influenza/surveillance/en/index.html
4. Stephenson I, Nicholson KG (2001) Influenza: vaccination and treatment. Eur Respir J 17: 1282–1293
5. Girard MP, Cherian T, Pervikov Y, Kieny MP (2005) A review of vaccine research and development: human acute respiratory infections. Vaccine 23:5708–5724
6. Garman E, Laver G (2004) Controlling influenza by inhibiting the virus's neuraminidase. Curr Drug Targets 5:119–136
7. Webster RG, Walker EJ (2003) Influenza: the world is teetering on the edge of a pandemic that could kill a large fraction of the human population. Am Sci 91:122
8. Horimoto T, Kawaoka Y (2005) Influenza: lessons from past pandemics, warnings from current incidents. Nat Rev Microbiol 3:591–600
9. Kiselar JG, Downard KM (1999) Direct identification of protein epitopes by mass spectrometry

without immobilization of antibody and isolation of antibody–peptide complexes. Anal Chem 71:1792–1799

10. Kiselar JG, Downard KM (1999) Antigenic surveillance of the influenza virus by mass spectrometry. Biochemistry 38:14185–14191
11. Morrissey B, Downard KM (2006) A proteomics approach to survey the antigenicity of the influenza virus by mass spectrometry. Proteomics 6:2034–2041
12. Morrissey B, Streamer M, Downard KM (2007) Antigenic characterisation of H3N2 subtypes of the influenza virus by mass spectrometry. J Virol Methods 145:106–114
13. Downard KM, Morrissey B (2007) Fingerprinting a killer – surveillance of the influenza virus by mass spectrometry. Analyst 132:611–614
14. Kiselar JG, Downard KM (2000) Preservation and detection of specific antibody-peptide complexes by matrix-assisted laser desorption ionization mass spectrometry. J Am Soc Mass Spectrom 11:746–750
15. Zhang WD, Evans DH (1991) Detection and identification of human influenza viruses by the polymerase chain reaction. J Virol Methods 33:165–189
16. Wright KE, Wilson GA, Novosad D, Dimock C, Tan D, Weber JM (1995) Typing and subtyping of influenza viruses in clinical samples by PCR. J Clin Microbiol 133:1180–1184
17. Ruben FL, Jackson GG, Gotoff SP (1973) Humoral and cellular response in humans after immunization with influenza vaccine. Infect Immun 7:594–596
18. Wood JM (2002) Selection of influenza vaccine strains and developing pandemic vaccines. Vaccine 20:B40–B44
19. Kodihalli S, Justewicz DM, Gubareva LV, Webster RG (1995) Selection of a single amino acid substitution in the hemagglutinin molecule by chicken eggs can render influenza A virus (H3) candidate vaccine ineffective. J Virol 69:4888–4897
20. Julkunen I, Pyhala R, Hovi T (1985) Enzyme immunoassay, complement fixation and hemagglutination inhibition tests in the diagnosis of influenza A and B virus infections. Purified hemagglutinin in subtype-specific diagnosis. J Virol Methods 10:75–84
21. Li J, Chen S, Evans DH (2001) Typing and subtyping influenza virus using DNA microarrays and multiplex reverse transcriptase PCR. J Clin Microbiol 39:696–704
22. O'Connell J (ed) (2002) RT-PCR protocols, methods in molecular biology, vol 193. Humana, NJ
23. Poddar SK (2002) Influenza virus types and subtypes detection by single step single tube multiplex reverse transcription-polymerase chain reaction (RT-PCR) and agarose gel electrophoresis. J Virol Methods 99:63–70
24. Ho JWK, Morrissey B, Downard KM (2007) A computer algorithm for the identification of protein interactions from the spectra of masses (PRISM). J Am Soc Mass Spectrom 18:563–566
25. Pappin DJC, Hojrup P, Bleasby AJ (1993) Rapid identification of proteins by peptide-mass fingerprinting. Curr Biol 3:327–332
26. Mackun K, Downard KM (2003) Strategy for identifying protein–protein interactions of gel-separated proteins and complexes by mass spectrometry. Anal Biochem 318:60–70
27. Sheshberadaran H, Payne LG (1988) Protein antigen–monoclonal antibody contact sites investigated by limited proteolysis of monoclonal antibody–bound antigen: protein "footprinting". Proc Natl Acad Sci 85:1–5
28. Cotter RJ (1997) Time-of-flight mass spectrometry: instrumentation and applications in biological research. American Chemical Society, Washington, DC
29. Macken C, Lu H, Goodman J, Boykin L (2001) The value of a database in surveillance and vaccine selection. In: Osterhaus ADME, Cox N, Hampson AW (eds) Options for the control of influenza IV. Elsevier Science, Amsterdam, pp 103–106
30. Downard KM (2004) Chapter 4 – tandem mass spectrometry, in Mass spectrometry – a foundation course. Royal Society of Chemistry, Cambridge
31. Morrissey B, Downard KM (2008) Kinetics of antigen-antibody interactions employing a MALDI mass spectrometry immunoassay. Anal Chem 80:7720–7726
32. Schwahn AB, Downard KM (2009) Antigenicity of a type A influenza virus through comparison of hemagglutination inhibition and mass spectrometry immunoassays. J Immunoassay Immunochem 30:245–261
33. Swaminathan K, Downard KM (2012) Antiviral inhibitor binding to influenza neuraminidase by MALDI mass spectrometry. Anal Chem 84:3725–3730
34. Polley JR (1969) Preparation of stable noninfective influenza virus antigens for typing by hemagglutination–inhibition. Can J Microbiol 15:203–207

Chapter 9

In Vivo Microbial Antigen Discovery (InMAD) to Identify Diagnostic Proteins and Polysaccharides That Are Circulating During Microbial Infections

Sindy J. Chaves, Kathleen Schegg, Thomas R. Kozel, and David P. AuCoin

Abstract

Immunoassays employed at the point-of-care (POC) are often useful for diagnosing acute infections. Many of these assays rely on identification of microbial antigens that are secreted or shed during infection. However, determining which microbial antigens are best to target by immunoassay can be the most difficult aspect of developing a new diagnostic product. Here we describe a novel technique termed "*In vivo* Microbial Antigen Discovery" or "InMAD" for identification of microbial antigens that may be targeted for the diagnosis of infectious diseases.

Key words Immunoassay, Diagnostic antigens, InMAD, Diagnosis

1 Introduction

Immunoassay for detection of microbial antigens in patient samples can lead to a rapid diagnosis of infection and timely administration of appropriate antibiotics. However, targeting antigens by immunoassay is not a trivial endeavor. Patient samples often contain low levels of microbes. Hence, only trace amounts of microbial antigens may be present within patient samples that contain an overwhelming amount of host proteins. Therefore, a critical step in diagnostic development relies on identifying microbial antigens that are shed into body fluids during infection at concentrations that are sufficient for detection. For this purpose we developed a novel strategy aimed at target identification termed "*In vivo* Microbial Antigen Discovery" or "InMAD" [1, 2]. To date we have used this technique to identify diagnostic antigens for *Burkholderia pseudomallei* (melioidosis) and *Francisella tularensis* (tularemia). Both of these pathogens present diagnostic challenges; most notably, they accumulate to very low levels within blood.

Kelly M. Fulton and Susan M. Twine (eds.), *Immunoproteomics: Methods and Protocols*, Methods in Molecular Biology, vol. 1061, DOI 10.1007/978-1-62703-589-7_9,

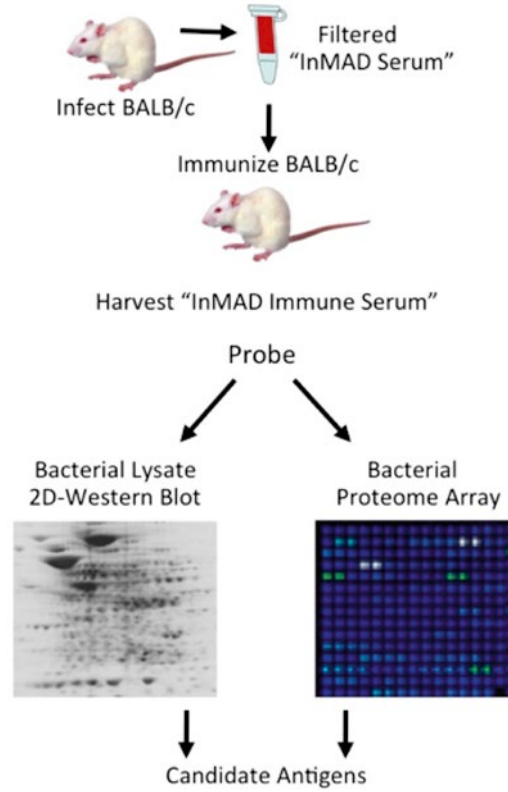


Fig 1 InMAD strategy for identifying secreted/shed microbial antigens. BALB/c mice are infected with a pathogen, and InMAD serum is harvested when mice become moribund. Filtering the InMAD serum removes whole microbes, and secreted/shed microbial antigens in the sample pass through the filter. Syngeneic mice are then immunized with the InMAD serum combined with adjuvant. InMAD immune serum is harvested from the immunized mice and used to probe microbial cell lysates by Western blot or a bacterial proteome array

B. pseudomallei, for example, has been shown to accumulate in patient blood samples at only one bacterium per milliliter [3].

The InMAD technique involves harvesting blood from an animal model of infection (Fig. 1). The serum is isolated and filtered to remove any whole microbial cells, leaving behind soluble microbial antigens released during infection. The "InMAD serum" is combined with adjuvant and used to immunize BALB/c mice. The immunized mice will generate an antibody response specific to the microbial antigens present in the serum. Filtering of the serum reduces the development of an antibody response to microbial antigens that are strictly cell-associated; these antigens may not make good diagnostic targets. The mice are bled and the resulting "InMAD immune serum" is used to probe one- and two-dimensional Western blots prepared from cell lysates of the microbe of interest. Reactive microbial antigens are identified by mass spectrometry. Alternatively, the InMAD immune serum can be used to probe a

microbial proteome array for rapid identification of reactive antigens [1, 4, 5]. Theoretically, the reactive antigens should be present in the original InMAD serum samples used to immunize the mice; therefore, these antigens represent candidate diagnostic antigens. These candidate diagnostic antigens can then be targeted with monoclonal antibodies by antigen-capture ELISA, lateral flow immunoassay or other antibody-based diagnostic platforms. The results obtained using InMAD serve as proof-of-concept that the technique is a powerful discovery platform to identify circulating proteins and polysaccharides that are present during microbial infections [1].

2 Materials

Prepare all solutions using ultrapure water and analytical-grade reagents. Prepare and store all reagents at room temperature (unless indicated otherwise). Filter all reagents (unless specified otherwise). Follow all waste disposal regulations when discarding reagents.

2.1 Immunization Components

1. Syringes (1 mL), 22 G1 gauge needles, and emulsification adaptor (BD, Franklin Lakes, NJ).
2. Titermax® Gold (Titermax, Norcross, GA).
3. Sterile 1× Dulbecco's phosphate-buffered saline (DPBS), pH 7.4 (GIBCO®, Grand Island, NY).
4. BD Microtainer® serum separator tubes (BD).
5. InMAD serum (pooled and filtered serum harvested from an animal model of infection, *see* **Note 1**).
6. Naïve BALB/c mice (*see* **Note 2**).

2.2 SDS-Polyacrylamide Gel Components

1. SDS-PAGE running buffer: 25 mM Tris–HCl, 192 mM Glycine, 0.1 % SDS, pH 8.3 (Cell Signaling Technology, Danvers, MA).
2. Mini Protean precast gels: 10–20 % SDS-PAGE gradient gels or 10 % SDS-PAGE gels depending on the molecular weights of the proteins of interest (Bio-Rad, Hercules, CA).
3. 2× Laemmli Sample buffer (Sigma-Aldrich, St. Louis, MO).
4. Boil-proof microcentrifuge tubes (VWR, Radnor, PA).
5. Prestained protein standard: Precision Plus Protein™ Standards Dual Color (Bio-Rad).

2.3 1D-Western Components

1. Western Blot Transfer buffer: 0.025 M Tris, 0.2 M glycine, and 20 % methanol, pH 8.5 (Cell Signaling Technology).
2. 10× Tris-buffered saline (TBS): 1.5 M NaCl, 0.1 M Tris–HCl, pH 7.4.

3. Polyvinylidene difluoride (PVDF) membrane, 0.2 μm pore size (Bio-Rad).
4. 0.1 % Tween 20 in TBS (TBST).
5. Blocking Solution: 5 % nonfat dry milk (NFDM) in TBST. Store at 4 °C for 1 week or prepare fresh batch when needed.
6. Two plastic containers with lids (just large enough for incubating PVDF membrane in blocking and washing buffers).
7. Super Signal West Pico Chemiluminescent Substrate (Thermo Scientific, Rockford, IL).
8. Miniblotter 20 SL (Immunetics, Boston, MA, USA).

2.4 2-D Gel Electrophoresis and Western Blot Components

1. 11 cm Immobilized pH gradient (IPG) strips covering a pH range of 3–10 NL (nonlinear) (Bio-Rad).
2. DeStreak Rehydration Solution (GE Healthcare).
3. pH 3–10 Ampholytes (40 % w/v) (Sigma Aldrich).
4. Sypro Ruby gel stain (Invitrogen, Carlsbad, CA).
5. Sypro Ruby blot stain (Invitrogen).
6. Dithiothreitol (DTT) solution: 2 % DTT, 6 M urea, 2 % SDS, 0.05 M Tris–HCl, pH 8.8, 20 % glycerol. This solution must be prepared on the day of use.
7. Iodoacetamide solution 2.5 % iodoacetamide 6 M urea, 2 % SDS, 0.05 M Tris–HCl, pH 8.8, 20 % glycerol. This solution must be prepared just before use.
8. Nitrocellulose membranes (0.2 μm, Bio-Rad).
9. Protean IEF Cell (Bio-Rad).
10. Agarose: dissolve 0.5 % electrophoresis grade agarose in 1× SDS-PAGE running buffer. Add a few grains of Bromophenol blue to give a medium blue color.
11. 8–16 % Criterion™ TGX™ Precast Gels (IPG plus 1-well comb, Bio-Rad).
12. 2-D Gel transfer buffer: 25 mM Tris, 192 mM glycine, 0.005 % SDS, 20 % methanol

3 Methods

Carry out all procedures at room temperature unless otherwise specified.

All animal work must be carried out in accordance with the recommendations in the Guide for the Care and Use of Laboratory Animals from the National Research Council. The Institutional Animal Care and Use Committee of the respective universities or facilities carrying out this procedure must also approve the protocol.

3.1 Immunization with InMAD Serum

1. Prior to the immunizations, collect pre-immune serum by retro-orbital bleed or submandibular puncture (100–200 µL of blood) from each mouse that will be immunized. Collect the blood in BD Microtainer® serum separator tubes (*see* **Note 3**). This serum will serve as a negative control.
2. Immunize each mouse with 5–50 µL of InMAD serum. Bring the InMAD serum up to 100 µL with PBS and combine it with 100 µL of TiterMax gold adjuvant. Scale up depending on the number of mice being immunized. Using the emulsification adapter emulsifies this mixture back and forth through two syringes (*see* **Note 4**).
3. Inject the emulsified product into BALB/c mice via the subcutaneous route (200 µL per mouse) (*see* **Note 5**).
4. Collect blood (maximum 200 µL) from mice at 4, 6, and 8 weeks post-immunization by retro-orbital bleeding or cardiac puncture (1 mL) and place in BD Microtainer® serum separator tubes. Sampling at 4, 6, and 8 weeks is done to determine the optimal time-point to harvest InMAD immune serum by cardiac puncture.
5. Once the blood is collected, allow it to clot for about an hour, and centrifuge the tubes at 4 °C for 5 min (2,300 × *g*). Collect the serum in a separate tube and label it (this serum is termed InMAD immune serum).
6. Keep the serum samples frozen at –20 °C if they will be used within a week. If not, keep the serum samples frozen at –80 °C.

3.2 SDS-PAGE

1. Resuspend pelleted microbial cells (enough cells to yield 50–100 µg following lysis) in 250 µL of (2×) Laemmli Buffer and heat sample(s) in a boiling water bath for 10 min to lyse cells. Do not add Laemmli buffer to the prestained protein standard or subject it to heat.
2. If needed, one volume of proteinase K solution (3.3 mg in 1 mL Laemmli sample buffer) can be added to the samples at this point. In this case, the sample must be incubated for 1 h at 60 °C (*see* **Note 6**).
3. Dilute SDS-PAGE Running Buffer (10×) with water to prepare 1 L of 1× solution.
4. With a razor or spatula, cut the well separators on a 10–20 % gel or 10 % gel to create one large well, leaving behind one of the small wells for the protein standard.
5. Load cell lysate (250 µL) into the large well. Load the protein standard (10 µL/well).
6. Electrophorese at 150 V until the dye front has reached the bottom of the gel.

7. Following electrophoresis, separate the gel from the plates with the help of a spatula. The gel will remain on one of the plates. Rinse the gel with water and transfer it carefully to a container with Western blot transfer buffer.

3.3 1D Western Blot

1. Cut a PVDF membrane to the size of the gel and immerse in methanol for 15 s, then transfer the membrane to a container with water and leave it for 2 min. Finally, transfer the membrane to a container with Western blot transfer buffer and let it soak for 5 min.
2. Blot the gel onto PVDF following a standard blotting protocol such as provided with the Trans-Blot Turbo (Bio-Rad).
3. Once transfer is completed, disassemble the sandwich and block the membrane in blocking solution at 4 °C with gentle shaking overnight.
4. Dilute the InMAD immune serum and pre-bleeds from each mouse 1:30 with TBST containing 5 % NFDM (*see* **Note 7**) and vortex.
5. Place the membrane in the Miniblotter 20SL (*see* **Note 8**) per manufacturer's protocol. Load the samples (pre-bleed adjacent to InMAD immune serum harvested from the same mouse) in the miniblotter (*see* Fig. 2 for an example from *B. pseudomallei* study) [1]. Remove any bubbles that form while loading the samples by pipetting up and down. Probe the blot at room temperature for 2 h.
6. Aspirate the serum from the miniblotter, and wash three times with 1× TBST, allowing 5 min between each wash.
7. Prepare the detecting antibody: Goat anti-mouse IgG-HRP at a 1:5,000 dilution in TBST containing 5 % NFDM.
8. Load each sample well on the miniblotter with the detecting antibody and incubate for 30 min at room temperature.
9. Aspirate the detecting antibody solution from the miniblotter, and wash three times with 1× TBST, allowing 5 min between each wash.
10. Disassemble the miniblotter, place the membrane in a plastic container with TBST, and allow for gentle rocking for about 5 min.
11. Detect IgG-HRP with Super Signal West Pico Chemiluminescent substrate, as per the manufacturer's instructions (*see* Figs. 2 and 3).

3.4 2-Dimensional Gel Electrophoresis and Western blot

Follow a standard two-dimensional gel protocol:

1. Lyse microbial cells within 200–300 µL of DeStreak Rehydration Solution. Enough cells should be lysed to yield 100–200 µg of microbial protein for each of two 2-D gels.

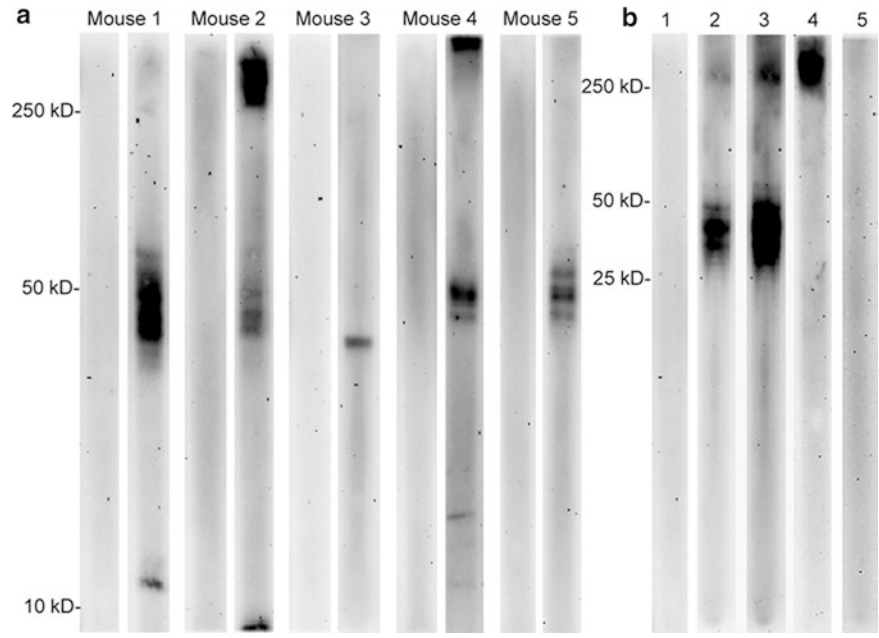


Fig. 2 *Burkholderia pseudomallei* proteins and the capsular polysaccharide are reactive with InMAD immune serum. (**a**) Five representative InMAD immune serum samples (*right lane* of each mouse) and pre-immune sera (*left lane* of each mouse) were used to probe a 1D Western blot of a *B. pseudomallei* whole cell lysate. (**b**) Serum from a representative mouse was used to determine if reactive antigens were polysaccharides or proteins. Pre-immune serum (*lane 1*) is not reactive with the *B. pseudomallei* whole cell lysate. InMAD immune serum from the same mouse is reactive with multiple antigens (*lane 2*; *lane 3* is a longer exposure). Only the high molecular weight antigen is reactive when InMAD immune serum is used to probe a blot of a proteinase K-treated lysate *(lane 4*). The same InMAD immune serum is not reactive with a proteinase K-treated lysate from a *B. pseudomallei* capsular polysaccharide (CPS) negative mutant strain SR1015 (*lane 5*)

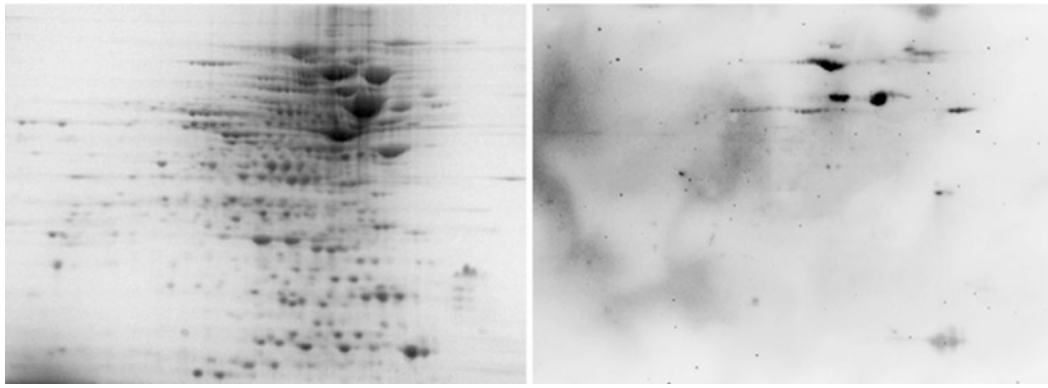


Fig. 3 *Francisella tularensis* proteins are reactive with InMAD immune serum. *F. tularensis* total protein (*left panel*) was separated by 2D-gel electrophoresis and blotted onto nitrocellulose. InMAD immune serum that was reactive with *F. tularensis* proteins by 1D-Western blot was pooled and used to probe a 2D-Western blot containing *F. tularensis* whole cell lystate *(right panel*). The 2D-Western blot was aligned with a duplicate total protein gel and reactive spots were excised and identified by mass spectrometry

Vortex the mixture frequently and sonicate for 10 min in a water bath sonicator over a period of about 1–1½ h. Spin the sample at full speed in a microcentrifuge at room temperature for 10 min.

2. Precipitate protein in the resulting supernatant overnight at –20 °C with four volumes of very cold acetone.
3. Pellet the precipitate, and wash twice with acetone/water (4:1, –20 °C).
4. Dry the final pellet and solubilize in at least 420 µL of DeStreak Rehydration Solution.
5. Determine protein concentration by EZQ analysis. Protein concentration should be between 0.5 and 1.0 mg/mL. If necessary, dilute sample with DeStreak Rehydration solution. To 420 µL of the appropriately diluted mixture, add 2.1 µL pH 3–10 ampholytes. Centrifuge the mixture for 10 min at top speed in a microfuge.
6. Rehydrate two 11 cm IPG strips by pipetting 200 µL of the supernatant into each of two wells in a rehydration plate and placing a strip face down on top of each sample (*see* **Note 9**).
7. Add enough mineral oil to cover the entire surface of the rehydration plate wells containing the strips. Allow the strips to rehydrate overnight.
8. Transfer strips from rehydration tray to IEF tray. Perform isoelectric focusing using a Bio-Rad Protean IEF cell. The following settings are used: 250 V, linear ramp for 20 min; 8,000 V, linear ramp for 2 h and 30 min; and 8,000 V for a total of 20,000 Vh (all steps with a maximum current of 50 µA per gel) (*see* **Note 10**).
9. When IEF is completed, drain the oil from each strip onto filter paper and store the strips face up at –80 °C if not being used immediately after the IEF cycles are completed.
10. Thaw IPG strips if they are frozen.
11. Incubate strips in DTT solution with gentle rocking for 10 min.
12. Discard the DTT containing solution and incubate in iodoacetamide solution with gentle rocking for 10 min.
13. Separate proteins by molecular weight with 8–16 % TGX Criterion Precast gels. Wash each strip in SDS-PAGE running buffer and load on top of a gel (*see* **Note 11**).
14. Spot 10 µL of prestained marker onto a Bio-Rad electrode wick (part no. BR 35327). Cut the wick in half vertically and load one of the halves in the narrow well on each gel.

15. Add roughly 2 mL of hot agarose solution to the top of each large well and allow it to solidify.
16. Run the gels until the dye front reaches the bottom of the gel at 200 V (*see* **Note 12**).
17. Immediately stain one of the gels with Sypro Ruby gel stain, as per the manufacturer's instructions.
18. Blot the other gel onto a nitrocellulose membrane using a Bio-Rad Criterion Blotter. Perform transfer at 100 V for 45 min to 1 h 15 min using 2-D gel transfer buffer.
19. Stain the nitrocellulose membrane with Sypro Ruby blot stain, as per the manufacturer's instructions, and image on a Bio-Rad ChemiDoc XRS+ molecular imager.
20. Probe the nitrocellulose membrane with pooled InMAD immune serum from mice that showed reactive bands in the 1-D Western blots. Dilute the InMAD immune serum 1:30 in 5–10 mL TBST containing 5 % NFDM. Incubate at room temperature with gentle shaking for 1 h.
21. Wash the membrane 3× with TBST, allowing 5 min between each wash.
22. Dilute goat anti-mouse IgG-HRP to 1:5,000 with TBST containing 5 % NFDM. Add the secondary antibody to the membrane and incubate for 30 min at room temperature with gentle shaking.
23. Wash the membrane 3× with TBST, allowing 5 min between each wash.
24. Reimage the membrane on a Bio-Rad ChemiDoc XRS+ molecular imager. Superimpose the images of the Sypro-stained membrane and the Western blot to locate the proteins spots that correspond to the antigens detected on the Western.
25. Excise the reactive spots/proteins of interest using a Bio-Rad ExQuest Spot Cutter (Fig. 3) from the non-blotted gel that was stained with Sypro Ruby gel stain. Proteins in the excised spots are identified by mass spectrometry [1].

4 Notes

1. The original InMAD studies were performed with blood harvested from a murine model of melioidosis [1, 6] and tularemia [1]. Filtered serum generated from any animal model of infection (InMAD serum) can be used for immunization. It is recommended that blood be harvested by cardiac puncture from infected animals that are moribund [7]. This will increase the likelihood that the serum contains microbial antigens.

2. We recommend immunizing at least ten mice with the pooled InMAD serum sample.
3. Do not fill the Microtainer tubes for blood collection to the top. Fill to half-capacity to allow for proper blood clotting. Collect between 100 and 200 μL.
4. Make sure the emulsification has a thick viscosity. Nonviscous emulsifications should not be used, since it means that the serum did not emulsify with the adjuvant.
5. Inject the emulsifications slowly.
6. The proteinase K-treated samples must be heated for 1 h at 60 °C following the boiling step. Proteinase K treatment will digest all proteins in the sample and will leave polysaccharides intact. If an antigen is still reactive on a second Western blot following proteinase K treatment of the microbial lysate, it is most likely a polysaccharide antigen.
7. Before diluting the InMAD immune serum, check the total volume that the miniblotter channel will hold in order to make sure the entire surface of the PVDF membrane will be covered by diluted InMAD immune serum.
8. When clamped together, the Miniblotter 20 SL creates 20 individual chambers so different serum samples from separate mice can be probed on the same membrane. It also has the advantage of accepting small volumes within each chamber; this is important since small volumes of InMAD immune serum are harvested from each mouse and the serum is used at a low dilution.
9. A 2D-Western blot will be prepared from one gel to detect microbial proteins reactive with InMAD immune serum. The other gel will be stained for total protein and then aligned with the Western blot. Reactive proteins will be excised and identified by mass spectrometry.
10. IEF will normally take roughly 8 h to complete all three cycles; however, it depends on the protein sample being analyzed. Individual strips may be removed when 20,000 Vh have been accumulated at the end of the third step or may be left on the IEF cell at a holding voltage of 500 V until isoelectric focusing is complete for all strips.
11. For two identical samples, label one gel “for total protein stain”, and the other “for 2D-Western blot”.
12. Some proteins are heavily glycosylated or are very abundant. In these cases, to get better separation, decrease the amount of protein being loaded and run the gels at a low voltage to prevent streaking.

Acknowledgments

This project was supported by award U54AI065359 and by grant number AI085548 from the National Institute of Health and the National Institute of Allergy and Infectious Diseases. Additional support was provided by award P20RR016464.

References

1. Nuti DE, Crump RB, Dwi Handayani F, Chantratita N, Peacock SJ et al (2011) Identification of circulating bacterial antigens by in vivo microbial antigen discovery. mBio 2(4)pii:e00136-11
2. AuCoin DP (2012) In vivo microbial antigen discovery: finding the 'needle in the haystack'. Expert Rev Mol Diagn 12:219–221
3. Wuthiekanun V, Limmathurotsakul D, Wongsuvan G, Chierakul W, Teerawattanasook N et al (2007) Quantitation of *B. pseudomallei* in clinical samples. Am J Trop Med Hyg 77: 812–813
4. Eyles JE, Unal B, Hartley MG, Newstead SL, Flick-Smith H et al (2007) Immunodominant *Francisella tularensis* antigens identified using proteome microarray. Proteomics 7: 2172–2183
5. Sundaresh S, Randall A, Unal B, Petersen JM, Belisle JT et al (2007) From protein microarrays to diagnostic antigen discovery: a study of the pathogen *Francisella tularensis*. Bioinformatics 23:i508–i518
6. Whitlock GC, Valbuena GA, Popov VL, Judy BM, Estes DM et al (2009) *Burkholderia mallei* cellular interactions in a respiratory cell model. J Med Microbiol 58:554–562
7. Hoff J (2000) Methods of blood collection in the mouse. Lab Anim 29:47–53

Chapter 10

Chemo-Enzymatic Production of *O*-Glycopeptides for the Detection of Serum Glycopeptide Antibodies

Alexander Nøstdal and Hans H. Wandall

Abstract

Protein microarray is a highly sensitive tool for antibody detection in serum. Monitoring of patients' antibody titers to specific antigens is increasingly employed in the diagnosis of several conditions, ranging from infectious diseases, allergies, autoimmune diseases, and cancer. In this protocol we present a detailed method for enzymatic generation of disease-specific *O*-glycopeptides and how to monitor the antibody response to these in serum using microarray technology.

Key words Glycopeptide microarray, GalNAc, Sialylation, Serum biomarkers, Autoantibodies, Posttranslational modification, Glycosylation

1 Introduction

Protein microarray enables simultaneous measurement of several biomarkers, such as antibodies against foreign antigens or autoantibodies [1, 2]. Antibody profiling has been shown to be useful in differential diagnosis of infectious diseases [3, 4] and allergies [5, 6]. Furthermore, elevated levels of specific autoantibodies are able to identify autoimmune disorders prior to onset of clinical symptoms [7, 8]. In recent years, the autoantibody response to tumor-associated antigens has received much interest for its potential capabilities as a biomarker for early detection of cancer [9]. Compared to the measurement of peptide and protein levels in bodily fluids, antibodies are produced in high, relatively stable titers, and have longer half-lives [10]. Besides their use as biomarkers, the identification of antibody targets could reveal specific vaccine candidates, for example viral and cancer vaccines [11–14].

The discovery of disease-specific antibodies has been approached through several different proteome-wide screening techniques, including expressed cDNA libraries (SEREX) [15], protein and peptide arrays [16, 17], both random and designed phage displays [18], and self-assembling protein arrays [19, 20].

Kelly M. Fulton and Susan M. Twine (eds.), *Immunoproteomics: Methods and Protocols*, Methods in Molecular Biology, vol. 1061, DOI 10.1007/978-1-62703-589-7_10,

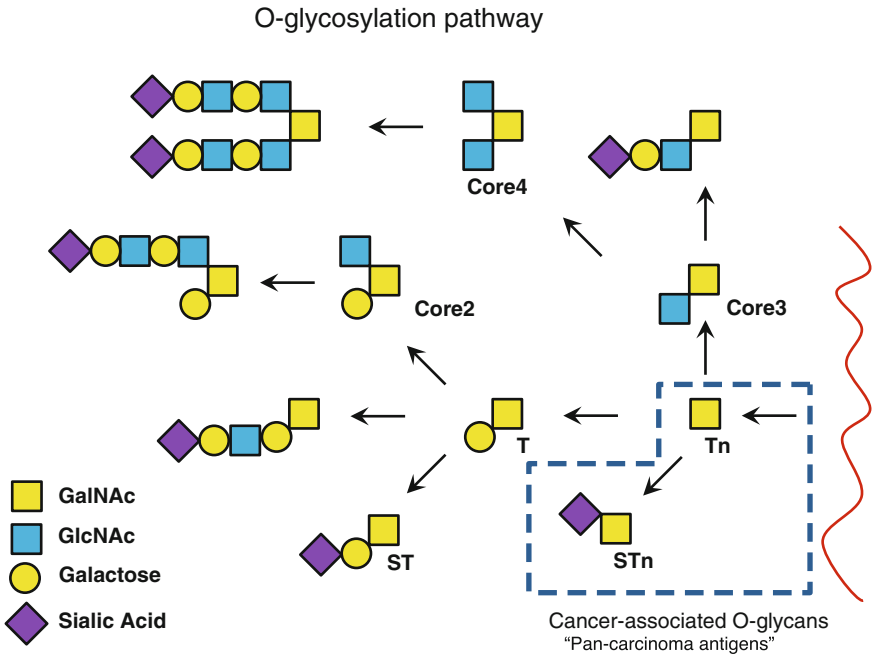


Fig. 1 The *O*-linked GalNAc glycosylation pathway. Glycosylation is initiated by the addition of GalNAc to Serine or Threonine residues in the protein core, and are elongated through subsequent enzyme reactions. *Symbols* for monosaccharides GalNAc, Gal, GlcNAc and Neu5Ac (sialic acid) are indicated

These methods are, however, designed to examine the proteome in the absence of posttranslational modification, which might limit the potential output of antibody screening [13].

Aberrant patterns of *O*-linked glycosylation (*see* Fig. 1) represent one of the most important cancer associated posttranslational changes [11, 21]. Changes in *O*-linked glycosylation are also associated with other pathological conditions, such as autoimmune diseases [22–25] and viral infection [26, 27]. Up regulation of unique short aberrant *O*-glyco-signatures on proteins may introduce novel glycopeptide epitopes that can elicit autoantibodies because of lack of tolerance.

By employing purified glycosyltransferases, we have previously described a method for *in vitro* *O*-glycosylation of synthetic MUC1 peptides [12, 13]. MUC1 is a heavily *O*-glycosylated mucin, with the glycans located predominately in a large 20 amino acid tandem repeat region [28]. While immunological tolerance is experienced to the tandem repeat protein core and its normal glycosylated forms [28, 29], aberrant truncation of the sugar chains results in the induction of autoantibodies to immunodominant *O*-glycopeptide epitopes. Importantly, the elicited immune response is specifically directed to combined glycopeptide epitopes, with little or no antibody specificity for the Tn carbohydrate hapten.

Through strict control of reactive enzymes we are able to recreate known disease-associated glycosylation patterns on our chosen antigens. In combination with microarray hydrogel slides, which provides remarkably low background levels [13], this offers a high-throughput method for screening patient sera for antibody reactivity.

In this protocol we are presenting a chemo-enzymatic approach using a synthetic 60-mer tandem repeat MUC1 peptide as an example to produce cancer-associated *O*-glycopeptides. The concept and methods described here could easily be transferred to other kinds of posttranslational modifications. The purpose of this chapter is to provide a detailed protocol for microarray detection of antibodies to glycosylated peptides and to offer aid with troubleshooting based on our experience and studies.

2 Materials

2.1 Generation of O-Glycopeptides

2.1.1 Enzymatic GalNAc-Glycosylation

1. 5× glycosylation buffer: 125 mM Cacodylate, 50 mM $MnCl_2$ (pH 7.4).

   Dissolve 802.7 mg $C_2H_6AsO_2Na{\cdot}3H_2O$ and 296.9 mg $MnCl_2{\cdot}4H_2O$ in ~20 mL MilliQ (MQ)-H_2O. Titer the solution to pH 7.4 using 6 M HCl, and once this is achieved add MQ-H_2O until total volume reach 30 mL. Store at 4 °C, up to 2 months.
2. 100 mM UDP-GalNAc.
3. Relevant GalNAc-Transferase(s): in this protocol we use GalNAc-T1, GalNAc-T2 [30], and GalNAc-T4 [31] to glycosylate our MUC1-peptide. Several trials using different enzymes, with subsequent Mass Spectrometry (MS) analysis, might be necessary to find a GalNAc-T with a suitable low K_m for the specific glycosylation site(s) (*see* Subheading 3.1 and **Note 1**). Subsequent elongation of GalNAc (Tn)-glycans, as shown in Fig. 1, into NeuAcα2,6GalNAc-S/T (STn) and GlcNAcβ3GalNAc-S/T (Core3) structures requires ST6GalNAc-I [32] and β3GnT6 [33], respectively.
4. Heating incubator (set to 37 °C).

2.1.2 Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry

1. Buffer A: 0.1 % trifluoroacetic acid (TFA), CF_3CO_2H.
   Mix 1 L MQ-H_2O with 1 mL TFA.
2. Matrix: 10 mg/mL 2,5-dihydroxybenzoic acid (DHB) dissolved in MQ-H_2O.
3. Matrix-assisted laser desorption/ionization (MALDI) sample plate.
4. Matrix-assisted laser desorption/ionization (MALDI) time-of-flight (TOF) instrument.

2.1.3 Peptide Purification by High-Pressure Liquid Chromatography

1. HPLC grade H_2O.
2. Buffer A: 0.1 % TFA (Subheading 2.1.2, **item 1**).
3. Buffer B: 90 % acetonitrile (ACN), 0.08 % TFA.
4. HPLC 1100 Hewlett Packard system (Avondale, PA).
5. Zorbax 300SB-C18 column, 100 mm × 4.6 mm (Agilent Technologies, Waldbronn, Germany).
6. Labconco Lyph-lock 1 L Lyophilizer (Labconco Corporation, Kansas City, MO).

2.2 Printing

1. Print buffer: 150 mM sodium phosphate (pH 8.5) with 0.005 % CHAPS and 0.03 % NaN_3

   Weigh 1.17 g $NaH_2PO_4 \cdot H_2O$, 11.8 g $Na_2HPO_4 \cdot 2H_2O$, and 25 mg CHAPS, and dissolve in MQ-H_2O. Add 5 mL 3 % NaN_3 and titer the solution to pH 8.5 using 6 M NaOH, and add MQ-H_2O until total volume reaches 500 mL.
2. Nexterion® Slide H MPX-48 (Schott AG, Mainz, Germany).
3. Light Microscope.
4. BioRobotics MicroGrid II spotter (Genomics Solution).
5. Stealth 3B Micro Spotting Pins (Telechem International ArrayIt Division).
6. Humidity chamber (*see* **Note 2**).

2.3 Scanning

1. Blocking buffer: 50 mM ethanolamine in 50 mM sodium borate (pH 8.5).

   Weigh 3.05 g ethanolamine and 10.06 g $Na_2B_4O_7$ (alternatively 19.6 g $Na_2B_4O_7 \cdot H_2O$) and dissolve in MQ-H_2O. Titer the solution to pH 8.5 using 6 M HCl, and add MQ-H_2O until total volume reaches 1 L. Store in amber glass bottle, at room temperature.
2. Phosphate-buffered saline (PBS) pH 7.4.
3. PBS 0.05 % Tween (PBS-T) pH 7.4.
4. Cy3-conjugated goat anti-human IgG (Fc-specific) antibodies.
5. Microarray Slide centrifuge.
6. ProScanArray HT Microarray Scanner (Perkin-Elmer).
7. ProScanArray Express 4.0 software (Perkin-Elmer), for image analysis.

3 Methods

3.1 Enzymatic In Vitro O-Glycosylation

In this protocol, we have glycosylated a 60mer MUC1 peptide $(VTSAPDTRPAPGSTAPPAHG)_{n=3}$ representing three tandem repeats, using purified GalNAc transferases. It is advised that you test

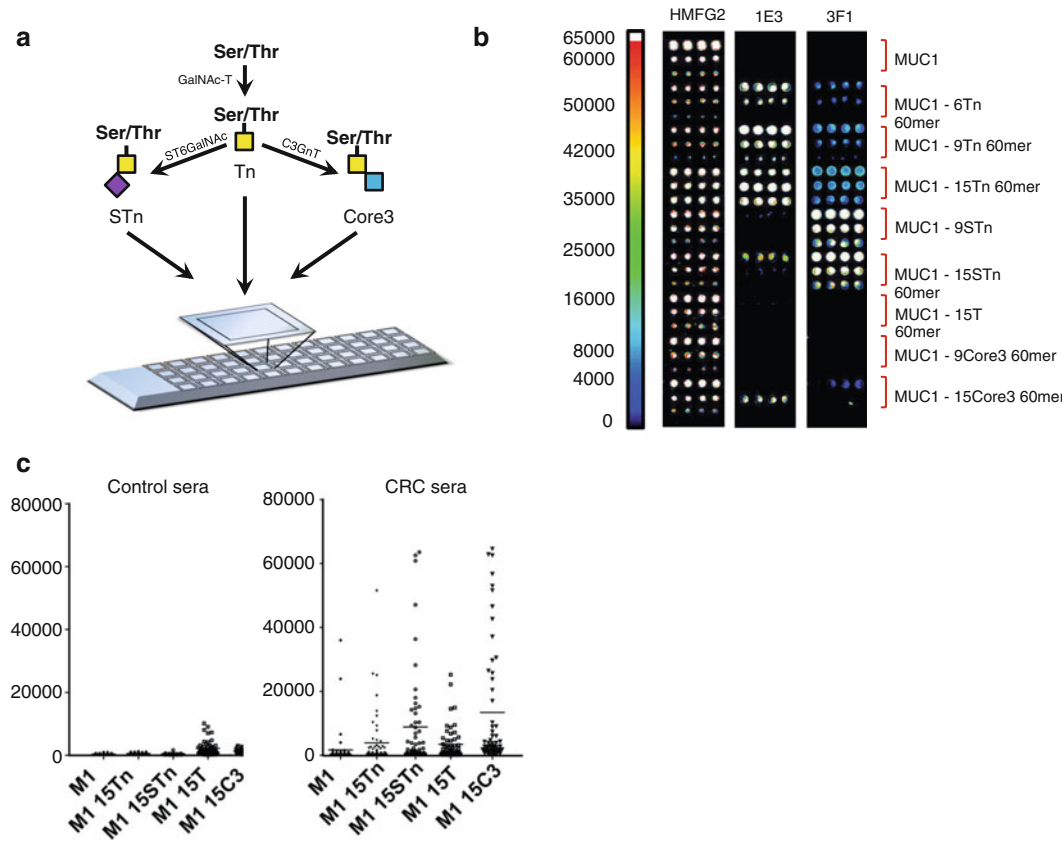


Fig. 2 MUC1 glycopeptides displayed on Micro Array. (**a**) GalNAc is added to Threonine and Serine residues on the peptide chain, directed by the specific activity of GalNAc-Ts, creating the Tn-antigen. The Tn structure can be elongated by subsequent enzyme reactions, e.g., into STn or Core3 structures. The Sialyl-Tn antigen is synthesized by the addition of a sialyl residue to GalNAc, whereas addition of a GlcNAc residue to GalNAc results in the formation of the Core3 structure. After purification the glycopeptide are displayed onto a micro array. (**b**) Each peptide is printed in quadruplicates, in three dilutions, and probed with the following antibodies (*left* to *right*): MUC1 specific mAb HMFG2; Tn specific mAb 1E3; STn specific mAb 3 F1. (**c**) *Dot-plot* diagram presenting IgG autoantibody reactivity against 60mer MUC1, MUC1 Tn, MUC1 STn, and MUC1 Core3. Each *dot* represents one individual of colorectal cancer patients (serum at time of diagnosis, pretreatment) or healthy controls. Figure elements previously published in Pedersen et al. [12]

your reaction on a small quantity of substrate with several different kinds of GalNAc-Transferases, to determine the most suitable transferase(s) (*see* **Note 1**).

More complex glycan structures (e.g., Core3 or STn) can be generated through elongating preexisting GalNAc residues (Tn antigens), as illustrated in Fig. 2a. The procedures generally follow the same steps as described here, after substituting enzymes and donor sugars. In some cases adjustments to the reaction buffers are necessary (*see* **Notes 3** and **4**).

The quantities used in the protocol are optimized for the glycosylation of 10 μg of the aforementioned peptide (60mer MUC1).

Adjustments to the concentrations of UDP-GalNAc, substrate, and enzyme might be necessary for other targets. Likewise the suggested timeframe might need adjustments based on the enzyme's potency and affinity for other targets.

1. Mix 5 µL 5× glycosylation buffer (Subheading 2.1.1) and 1 µL 100 mM UDP-GalNAc in an eppendorf tube.
2. Add 1 µL substrate (10 µg/µL).
3. Add 0.1 mU GalNAc-T1, 0.1 mU GalNAc-T2, and 0.1 mU GalNAc-T4 (*see* **Note 5**).
4. Add MQ-H_2O to a total volume of 25 µL.
5. Vortex-mix briefly and place the Eppendorf tube in the heating incubator at 37 °C, for 12–16 h.
6. Use MALDI-TOF to monitor glycosylation of peptide (Subheading 3.2).
7. STn-MUC1 and Core3-MUC1 is enzymatically synthesized based on purified GalNAc-MUC1 and is characterized and purified as described for GalNAc-MUC1 below (*see* **Notes 3** and **4**).

Note that no detergent (e.g., Triton) is added to the reaction (*see* **Note 6**).

3.2 MALDI-TOF Mass Spectrometer

1. Place a 0.5 µL droplet of DHB on the MALDI sample plate.
2. Dilute 0.5 µL of your sample in 10 µL buffer A (Subheading 2.2, **item 1**), and transfer 0.5 µL of this solution to the DHB on the sample plate.
3. Repeat **steps 1** and **2** of Subheading 3.2 for unglycosylated version of your peptide, for reference.
4. Let the matrix-sample-mix dry on the plate; then place the plate in the MALDI-TOF instrument.

Ionize the sample. Use high laser power until a peak correlating to the expected molecular weight appears, then lower laser power to minimize background noise.

After verification of the correct number of GalNAc additions by MALDI-TOF it is advisable, if possible, to confirm the position of the GalNAc in the intended site(s) [34, 35].

3.3 Peptide Purification by HPLC

Following verification of glycosylation, purify the peptide by high-performance liquid chromatography on a C18 column (Subheading 2.1.3, **item 4**), eluting with a 40 min gradient from 0 to 90 % ACN in 0.08 % TFA (buffer B, Subheading 2.1.3, **item 2**), at 0.800 mL/min flow. Lyophilize the sample and redissolve in MQ-H_2O the following day. Verify successful elution of your peptide through MALDI (Subheading 3.2).

3.4 Printing

We employed a BioRobotics MicroGrid II spotter (Genomics Solution) for printing our microarrays. When designing your own microarray, be sure that you make yourself familiar with the hardware and software by reading through the relevant chapters in the user manual. It offers detailed directions by the manufacturer on how to adjust settings to fit your needs.

We recommend printing each compound in duplicates or triplicates if space allows it on the array, and use the mean value of these in the analysis.

3.4.1 Preparing the Sample Source Plate (Biobank)

1. Dilute your peptides in Print buffer (Subheading 2.1.1). For the first run with new peptides, we recommend running a test-print with numerous dilutions, to determine the ideal concentration for each compound. Other methods of optimization might be necessary; for example, highly hydrophobic peptides might benefit from addition of dimethyl sulfoxide (DMSO) in the Print buffer, to be fully dissolved (*see* **Note 7**).
2. Seal the biobank with Parafilm to prevent evaporation and store at –20 °C between runs (*see* **Notes 8** and **9**).
3. Before printing it is important to remove any air bubbles from the wells of the biobank, as they may interfere with the pins' correct acquisition of compounds when dipping into the source plate. Air bubbles are easily removed by centrifugation (e.g., 200 × *g* for 1–2 min).

3.4.2 Printing Glycopeptides

1. We store microarray slides at –20 °C prior to use. Let them equilibrate at room temperature for approximately 30 min before printing; keep them in their sealed envelope for as long as possible. This will prevent condensation on the slides.
2. Fill the 6 L water reservoir. It might be necessary to refill during the print-run, depending on the number of slides, spots per slide, wash-settings and total run time.
3. Examine spotting pins in light microscope. Make sure the pins you choose for your print are clean and unbent. If needed, follow the manufacturer's recommendations for cleaning the pins.

   Place pins in tool, and load it onto the printer.
4. Prime the main wash station. Make sure the water flow is continuous and without air bubbles. To remove air, drain water from the 6 L water reservoir until no air bubbles are visible in the waste-tube. Prime the main wash station again, and perform the program "Regular wash cycle (without move to bath)" three consecutive times. Repeat if air bubbles persist.
5. Perform one "Regular wash cycle (with move to bath)"; make sure that the pins are lowered into the wash station accurately.

6. Load the source plate (biobank).
7. Place slides on the tray. Fill out empty slots with spare glass slides. After switching on the vacuum, make sure all slides are aligned correctly.
8. Start the run. The steps for selecting and adjusting run preferences will not be covered here. We advise that PreSpotting and Soft Touch (see user manual) are used, for the sake of consistent spot morphology. Make sure the pins are lowered correctly into the source plate wells during their route through first source visit to spotting; and that both PreSpotting and Soft Touch are working as intended.
9. After printing, place the slides in a humidity chamber for 2 h, before you either freeze the slides (−20 °C) for later use or proceed directly with serum placement and scanning.

3.5 Serum Placement and Scanning

1. Block slides in blocking buffer (Subheading 2.3, **item 1**) for 1 h. While the slides are being blocked, thaw serum samples on ice. Vortex mix briefly before use when thawed.
2. Wash slides in PBS for 2× 5 min.
3. Rinse slides in MQ-H_2O and dry through centrifugation (200 × *g*). This ensures that the teflon-coating separating each well will be capable of upholding its function.
4. Place slides in a petri dish.
5. Dilute the sera 1:5 in PBS-T, and add 9 μL to each well of the slide. This can also be done on slide, by placing a 7.5 μL droplet of PBS-T in each well and then adding 1.5 μL undiluted sera (*see* **Note 10**).
6. The addition of Tween to the diluted sera prevents nonspecific binding of antibodies.

   When doing large batches, we recommend devoting one to four wells per slide to the same sera or antibodies to estimate slide-to-slide variation. This also ensures that in the case of scanner malfunction or service, scanning parameters can be calibrated using the control wells.
7. Let serum incubate on slide for 1 h, under gentle agitation (~50 motions per minute).
8. Wash off the sera: wash with PBS-T for 2× 5 min; then a third 5-min wash cycle where PBS-T is swapped for PBS after 2.5 min.
9. While the slide(s) are washed, prepare a clean petri dish for the next steps, inside a light-blocking medium (e.g., a box or a sheet of aluminum foil).
10. Dry the edges of the slide off with a tissue, creating a dry teflon-frame around the wells.

11. Dilute Cy3-conjugated Goat anti-Human IgG (Fc-specific) antibodies 1:4,000 in MQ-H_2O.

    Place 1 mL per slide inside the dried teflon frame and incubate for 1 h.

    The slides should henceforth be shielded from light, by closing the box/aluminum foil around the petri dish.

12. Wash off the goat anti human antibodies: wash with PBS-T for 2× 5 min; then a third 5-min wash cycle with PBS. The slides should be shielded from light during these washes.

13. Rinse slides in MQ-H_2O, dry by centrifugation (200×*g*) and place slide in scanner-cassette.

    Scan the slides using ProScanArray HT Microarray Scanner (Perkin-Elmer). We recommend scanning with 100 % laser power at several different PMT gain (parameter for the activity of the Photomultiplier tubes of the scanner; typically ranging from 70 to 90 %) to find a suitable standard for your samples, with reasonably low background levels and where the reactive spots' intensity is easily readable. Resolution of the images should not surpass 10 μm (*see* **Note 11**).

14. After scanning, seal slides in a container by Parafilm, and freeze at –20 °C. If needed, slides can be thawed and rescanned (e.g., at other scanning parameters).

3.6 Quantification and Data Analysis

After scanning, set up a quantification protocol as per the Scan Array Express user manual. To simplify the data analysis, we highly recommend the creation and use of a .GAL-file for the quantification template, so that the Name- and Id-tag for each printed compound is imported automatically to their corresponding spots. By following the printing specifications used in this protocol, each spot should average 150 μm in diameter. When the quantification template is correctly aligned on top of the scanned image of the MicroArray, export the result spreadsheet to excel, and convert the text to columns.

Each spot is now easily identifiable with correct name; and the Array Column and Row information enables sorting of the data after serum sample. The final data output is gained from the mean spot Relative Fluorescence Units (RFU) after subtracting the surrounding background (found in column Y in the excel spreadsheet; labeled as "Ch1 Mean – B") (*see* **Note 12**).

4 Notes

1. There are over 200 glycosyltransferases in human cells, divided into 90 distinct families sharing mutual donor-sugar substrates. Of the GalNAc-Transferase (GalNAc-Ts) family, there are 20 known members, with varying affinity and specific activity for

each potential glycosylation site. Although some effort has been put into mapping glycosylation sites in different proteins for specific GalNAc-Ts, some trial and error employing different enzymes is usually needed to get the desired sites glycosylated. Also, if a peptide holds more than one potential site, it is not uncommon that several different transferases are needed to glycosylate them all.

2. Commercial humidity chambers are available; however, we made our own using a sealable plastic container with a perforated tray for slides. The lower compartment was filled with NaCl-saturated water, stabilizing the relative humidity at ~75 %, room temperature.
3. The purified Tn-bearing peptide can be enzymatically sialylated with ST6GalNAc-I (Sialyl Transferase), in the following reaction mixture: 50 mM MES (pH 6.5), 20 mM EDTA, and 2 mM DTT. The donor sugar used for this reaction is CMP-NeuAc, at 2 mM.
4. Enzymatic elongation of the purified Tn-bearing peptide can be performed by β3Gn-T6 (Core3 Synthase), in the following reaction mixture: 50 mM Cacodylate (pH 7.4) and 10 mM $MnCl_2$. The donor sugar used for this reaction is UDP-GlcNAc, at 2 mM.
5. The amounts of enzyme (0.1 mU GalNAc-T1, 0.1 mU GalNAc-T2 and 0.1 mU GalNAc-T4) was determined by scintillation counting after Dowex-1 formic acid chromatography of 25 μL reaction assays. The procedure is as described in Subheading 3.1, with UDP-GalNAc substituted for UDP-[$_{14}$C]GalNAc (2,000 cpm/nmol) (Amersham Biosciences).

   For GalNAc-T1 and -T2, IgA-Hinge peptide (VPSTPPTPS PSTPPPTSPSK) was used for the activity assay; for GalNAc-T4, a MUC7 peptide (APPTPSATTPAPPSSSAPPETTAA) was used.
6. Avoid using detergent during enzymatic *O*-glycosylation with soluble glycosyltransferases when possible. It is difficult to completely remove detergent during purification, and this could later cause problems when printing. Compounds mixed with detergent will often create non-homogenous spot morphologies and may in worst case cause spots to merge, rendering slide uninformative and unusable.
7. Very high concentrations of DMSO is not recommended for printing, as it may interfere with the hydrogel surface of the microarray slides; we have however, successfully printed with up to 60 % DMSO in our Print buffer.
8. During long runs, water evaporation of the biobank will lower the volume of each sample, and increase the concentration in each well. In these cases, it is possible to refill the wells with MQ-H_2O, before a new print.

9. Precipitation issues may arise after thawing a frozen biobank. A way to solve this problem is to incubate the thawed biobank at 37 °C for 20 min before printing. A more laborious method is to manually pipette up and down repeatedly in the well of each compound, thus homogenizing the content of each well.
10. Dilution of sera is application and source dependent, and should be optimized for each experimental design. In the current protocol we use a serum dilution of 1:5.
11. As with the MicroArray printer hardware and -software, it is important that you have made yourself familiar with the scanner and its software. The manufacturer's user manual offers a great overview on how to use the scanner, and after reading the relevant chapters in the manual, it is fairly easy to adjust the default settings to those specified in this protocol.
12. Presumably due to varying content composition between different serum samples, it is not unusual to observe varying levels of background signal between subarrays of a MicroArray slide, directly corresponding to the area covered by serum. To remove this variance from the results, the background level is automatically measured by the software, and subtracted from the mean spot RFU.

References

1. Kingsmore SF (2006) Multiplexed protein measurement: technologies and applications of protein and antibody arrays. Nat Rev Drug Discov 5:310–320
2. Robinson WH (2006) Antigen arrays for antibody profiling. Curr Opin Chem Biol 10:67–72
3. Chevaliez S, Pawlotsky JM (2008) Diagnosis and management of chronic viral hepatitis: antigens, antibodies and viral genomes. Best Pract Res Clin Gastroenterol 22:1031–1048
4. Murdoch DR, O'Brien KL, Driscoll AJ, Karron RA, Bhat N (2012) Laboratory methods for determining pneumonia etiology in children. Clin Infect Dis 54(Suppl 2):S146–S152
5. Fall BI, Eberlein-Konig B, Behrendt H, Niessner R, Ring J, Weller MG (2003) Microarrays for the screening of allergen-specific IgE in human serum. Anal Chem 75:556–562
6. Ferrer M, Sanz ML, Sastre J, Bartra J, del Cuvillo A, Montoro J, Jauregui I, Davila I, Mullol J, Valero A (2009) Molecular diagnosis in allergology: application of the microarray technique. J Investig Allergol Clin Immunol 19(Suppl 1):19–24
7. Leslie D, Lipsky P, Notkins AL (2001) Autoantibodies as predictors of disease. J Clin Invest 108:1417–1422
8. Routsias JG, Tzioufas AG, Moutsopoulos HM (2004) The clinical value of intracellular autoantigens B-cell epitopes in systemic rheumatic diseases. Clin Chim Acta 340:1–25
9. Tainsky MA (2009) Genomic and proteomic biomarkers for cancer: a multitude of opportunities. Biochim Biophys Acta 1796:176–193
10. Anderson KS, LaBaer J (2005) The sentinel within: exploiting the immune system for cancer biomarkers. J Proteome Res 4:1123–1133
11. Tarp MA, Clausen H (2008) Mucin-type O-glycosylation and its potential use in drug and vaccine development. Biochim Biophys Acta 1780:546–563
12. Pedersen JW, Blixt O, Bennett EP, Tarp MA, Dar I, Mandel U, Poulsen SS, Pedersen AE, Rasmussen S, Jess P, Clausen H, Wandall HH (2010) Seromic profiling of colorectal cancer patients with novel glycopeptide microarray. Int J Cancer 128(8):1860–1871
13. Wandall HH, Blixt O, Tarp MA, Pedersen JW, Bennett EP, Mandel U, Ragupathi G, Livingston PO, Hollingsworth MA, Taylor-Papadimitriou J, Burchell J, Clausen H (2010) Cancer biomarkers defined by autoantibody signatures to aberrant O-glycopeptide epitopes. Cancer Res 70:1306–1313

14. Clo E, Kracun SK, Nudelman AS, Jensen KJ, Liljeqvist JA, Olofsson S, Bergstrom T, Blixt O (2012) Characterization of the viral O-glycopeptidome: a novel tool of relevance for vaccine design and serodiagnosis. J Virol 86:6268–6278
15. Sahin U, Tureci O, Schmitt H, Cochlovius B, Johannes T, Schmits R, Stenner F, Luo G, Schobert I, Pfreundschuh M (1995) Human neoplasms elicit multiple specific immune responses in the autologous host. Proc Natl Acad Sci U S A 92:11810–11813
16. Pereira-Faca SR, Kuick R, Puravs E, Zhang Q, Krasnoselsky AL, Phanstiel D, Qiu J, Misek DE, Hinderer R, Tammemagi M, Landi MT, Caporaso N, Pfeiffer R, Edelstein C, Goodman G, Barnett M, Thornquist M, Brenner D, Hanash SM (2007) Identification of 14-3-3 theta as an antigen that induces a humoral response in lung cancer. Cancer Res 67:12000–12006
17. Stockert E, Jager E, Chen YT, Scanlan MJ, Gout I, Karbach J, Arand M, Knuth A, Old LJ (1998) A survey of the humoral immune response of cancer patients to a panel of human tumor antigens. J Exp Med 187:1349–1354
18. Mintz PJ, Kim J, Do KA, Wang X, Zinner RG, Cristofanilli M, Arap MA, Hong WK, Troncoso P, Logothetis CJ, Pasqualini R, Arap W (2003) Fingerprinting the circulating repertoire of antibodies from cancer patients. Nat Biotechnol 21:57–63
19. Anderson KS, Ramachandran N, Wong J, Raphael JV, Hainsworth E, Demirkan G, Cramer D, Aronzon D, Hodi FS, Harris L, Logvinenko T, LaBaer J (2008) Application of protein microarrays for multiplexed detection of antibodies to tumor antigens in breast cancer. J Proteome Res 7:1490–1499
20. Ramachandran N, Hainsworth E, Bhullar B, Eisenstein S, Rosen B, Lau AY, Walter JC, LaBaer J (2004) Self-assembling protein microarrays. Science 305:86–90
21. Schietinger A, Philip M, Yoshida BA, Azadi P, Liu H, Meredith SC, Schreiber H (2006) A mutant chaperone converts a wild-type protein into a tumor-specific antigen. Science 314:304–308
22. Chui D, Sellakumar G, Green R, Sutton-Smith M, McQuistan T, Marek K, Morris H, Dell A, Marth J (2001) Genetic remodeling of protein glycosylation in vivo induces autoimmune disease. Proc Natl Acad Sci U S A 98: 1142–1147
23. Opdenakker G, Dillen C, Fiten P, Martens E, Van Aelst I, Van den Steen PE, Nelissen I, Starckx S, Descamps FJ, Hu J, Piccard H, Van Damme J, Wormald MR, Rudd PM, Dwek RA (2006) Remnant epitopes, autoimmunity and glycosylation. Biochim Biophys Acta 1760: 610–615
24. Axford JS, Sumar N, Alavi A, Isenberg DA, Young A, Bodman KB, Roitt IM (1992) Changes in normal glycosylation mechanisms in autoimmune rheumatic disease. J Clin Invest 89:1021–1031
25. Delves PJ (1998) The role of glycosylation in autoimmune disease. Autoimmunity 27: 239–253
26. Olofsson S, Blomberg J (1977) Studies on glycopeptides of Herpes simplex virus infected cells. Arch Virol 55:293–304
27. Brennan PJ, Steiner SM, Courtney RJ, Skelly J (1976) Metabolism of galactose in herpes simplex virus-infected cells. Virology 69: 216–228
28. Taylor-Papadimitriou J, Burchell JM, Plunkett T, Graham R, Correa I, Miles D, Smith M (2002) MUC1 and the immunobiology of cancer. J Mammary Gland Biol Neoplasia 7:209–221
29. Rughetti A, Pellicciotta I, Biffoni M, Backstrom M, Link T, Bennet EP, Clausen H, Noll T, Hansson GC, Burchell JM, Frati L, Taylor-Papadimitriou J, Nuti M (2005) Recombinant tumor-associated MUC1 glycoprotein impairs the differentiation and function of dendritic cells. J Immunol 174:7764–7772
30. Wandall HH, Hassan H, Mirgorodskaya E, Kristensen AK, Roepstorff P, Bennett EP, Nielsen PA, Hollingsworth MA, Burchell J, Taylor-Papadimitriou J, Clausen H (1997) Substrate specificities of three members of the human UDP-N-acetyl-alpha-D-galactosamine:Polypeptide N-acetylgalactosaminyltransferase family, GalNAc-T1, -T2, and -T3. J Biol Chem 272:23503–23514
31. Bennett EP, Hassan H, Mandel U, Mirgorodskaya E, Roepstorff P, Burchell J, Taylor-Papadimitriou J, Hollingsworth MA, Merkx G, van Kessel AG, Eiberg H, Steffensen R, Clausen H (1998) Cloning of a human UDP-N-acetyl-alpha-D-Galactosamine:polypeptide N-acetylgalactosaminyltransferase that complements other GalNAc-transferases in complete O-glycosylation of the MUC1 tandem repeat. J Biol Chem 273:30472–30481
32. Ikehara Y, Kojima N, Kurosawa N, Kudo T, Kono M, Nishihara S, Issiki S, Morozumi K, Itzkowitz S, Tsuda T, Nishimura SI, Tsuji S, Narimatsu H (1999) Cloning and expression of a human gene encoding an N-acetylgalactosamine-alpha2,6-sialyltransferase (ST6GalNAc I): a candidate for synthesis of cancer-associated sialyl-Tn antigens. Glycobiology 9:1213–1224

33. Iwai T, Inaba N, Naundorf A, Zhang Y, Gotoh M, Iwasaki H, Kudo T, Togayachi A, Ishizuka Y, Nakanishi H, Narimatsu H (2002) Molecular cloning and characterization of a novel UDP-GlcNAc:GalNAc-peptide beta1,3-N-acetylglucosaminyltransferase (beta 3Gn-T6), an enzyme synthesizing the core 3 structure of O-glycans. J Biol Chem 277:12802–12809
34. Mirgorodskaya E, Hassan H, Wandall HH, Clausen H, Roepstorff P (1999) Partial vapor-phase hydrolysis of peptide bonds: a method for mass spectrometric determination of O-glycosylated sites in glycopeptides. Anal Biochem 269:54–65
35. Pedersen JW, Bennett EP, Schjoldager KT, Meldal M, Holmer AP, Blixt O, Clo E, Levery SB, Clausen H, Wandall HH (2011) Lectin domains of polypeptide GalNAc transferases exhibit glycopeptide binding specificity. J Biol Chem 286:32684–32696

Chapter 11

Enrichment and Characterization of Glycopeptide Epitopes from Complex Mixtures

Luc Tessier, Kelly M. Fulton, and Susan M. Twine

Abstract

Antigen posttranslational modifications, including glycosylation, are recognized by the innate and adaptive arms of the immune system. Analytical approaches, including mass spectrometry and allied techniques, have allowed advances in the enrichment and identification of glyco-antigens, particularly T-cell epitopes. Similarly, major advances have been made in the identification, isolation, and detailed characterization of prokaryotic and eukaryotic glycoproteins and glycopeptides. In particular, peptide centric approaches are now capable of enriching low level glycopeptides from highly complex peptide mixtures. Similarly, advanced mass spectrometry methods allow identification of glycopeptides, characterization of glycans, and mapping of modification sites. Herein, we describe methods developed in our laboratory for the broad study of glycopeptides and illustrate how these approaches can be exploited to further our understanding of the identity and nature of glycopeptide epitopes in various diseases or auto immune disorders.

Key words Glycosylation, Glycoprotein, Glycopeptide, Glycopeptide epitope, Mass spectrometry, Peptide identification, Electron transfer dissociation, Soft collision induced dissociation, Posttranslational modification, Modification site

1 Introduction

Over the past decade there have been advances in approaches to study protein posttranslational modifications (PTM) [1]. Glycosylation is one such PTM, and has a role in many diseases, including cancers, autoimmune diseases, viral infections, and bacterial pathogenesis. Changes in the frequency or pattern of protein glycosylation can be caused by numerous factors including infections, cellular transformation, or cell death [2–5]. There is increasing evidence that antigen posttranslational modifications are recognized by the humoral and cellular immune systems, and may have an important role in various diseases [6–8]. There is also evidence that protein glycans modulate T-cell recognition [6, 8]. How glycan antigens influence T-cell recognition is important in the generation of an immune response to pathogen or tumor

Kelly M. Fulton and Susan M. Twine (eds.), *Immunoproteomics: Methods and Protocols*, Methods in Molecular Biology, vol. 1061, DOI 10.1007/978-1-62703-589-7_11,

associated antigens, and also to the induction of tolerance to self-antigens. For example, in many cancers, aberrant patterns of *O*-linked protein glycosylation produce variants of proteins not covered by immune tolerance [9]. Either aberrant expression of proteins, mutation, or changes in glycosylation can introduce novel glycopeptide epitopes which have the potential to induce autoantibodies. Circulating autoantibodies are emerging as promising biomarkers for the detection of cancer, particularly in early diagnosis [10–13]. Studies have demonstrated cancer associated autoantibodies to aberrantly *O*-glycosylate MUC1 mucin in patients with prostate, breast or ovarian cancers [14, 15]. However, discovery of glycopeptide autoantibodies has been hampered by a lack of assays or analytical techniques that readily detect and characterize the truncated cancer associated glycans. Some novel technologies are emerging, including a recently reported glycopeptide microarray for high throughput discovery of these glycopeptide antigens [16].

Pathogen derived glycosylated proteins are also being increasingly found to have a role in infectious diseases, and this area too suffers from a paucity of analytical techniques for rapid glycoprotein detection and characterization. Bacterial glycoprotein characterization is challenging due to the wide diversity of glycan moieties elaborated by bacteria. Information regarding the biological role of these bacterial glycoproteins is emerging, with roles in interference with host inflammatory immune responses, adherence, motility and bacterial virulence [17, 18]. For example, the bacterial flagellum is a virulence factor in many motile bacteria and is recognized by the innate immune system. The flagellin monomer is recognized by the Toll-like receptor 5 (TLR5), a member of a family of pattern recognition receptors. These receptors play a front-line role in host defense, in particular inducing innate immune responses. Flagellin, the protein monomer, is a pathogen-associated molecular pattern (PAMP), recognized by both plants [19] and animals [20].

The identification of carbohydrate epitopes is a limiting factor in understanding their role and impact upon immune function. Analytical and mass spectrometry technologies have provided breakthroughs in enrichment and identification of the immunopeptidome [21]. However, there has been little focus upon glycosylated immunopeptides. In comparison, the study of eukaryotic and prokaryotic glycoproteins is a rapidly growing field, with increasingly sophisticated analytical approaches being developed to determine the identity, nature and dynamic changes in the glycoproteome. A number of these approaches could be applied to the enrichment and characterization of glycopeptide epitopes. In this work, we provide details of approaches we have developed and applied for the enrichment of glycopeptides from highly complex peptide mixtures. We then detail how our advanced mass spectrometry methods can be applied to identify and characterize enriched glycopeptides. These approaches have high utility for the discovery and characterization of glycopeptide epitopes from a variety of diseases.

2 Materials

2.1 Protein Samples

1. Lysates or sub-proteome fractions from cell line or bacterial strain under investigation.

   Typically 100–1,000 μg of total proteome extract or subcellular lysate is required.

2.2 Materials and Reagents

All solvents used are HPLC grade, and are degassed before use. Except where indicated otherwise, water used in solvent preparations should be Milli-Q® or higher purity.

1. Acetone (ice cold).
2. 4 mM Dithiothreitol (DTT).
3. 20 mM Iodoacetamide.
4. Sequencing grade trypsin.
5. 50 mM ammonium bicarbonate (ABC) buffer.
6. Polyhydroxyethyl aspartamide column: 1 mm × 50 mm; 5 μm; 300 Å (The Nest Group, Southborough, MA).
7. 90 % (v/v) acetonitrile (ACN) in water.
8. 10 % (v/v) trifluoroacetic acid (TFA) in water.
9. IP-NPLC Buffer A: 0.1 % (v/v) formic acid (FA) in ACN.
10. IP-NPLC Buffer B: 0.1 % (v/v) FA in high-performance liquid chromatography (HPLC)-grade H_2O.
11. Solvent A: 0.1 % (v/v) FA in HPLC grade H_2O.
12. Solvent B: 0.1 % FA in ACN.
13. Trapping column: 300 μm × 50 mm PepMap300 C18 μPre-Column (Dionex/ThermoFisher).
14. Nano HPLC analytical column: 100 μm internal diameter (ID) × 10 cm; 1.7 μm BEH130 C18 column (Waters) [22].
15. Clear Glass 12 mm × 32 mm Screw Neck Max Recovery Vials (Waters).

2.3 Instrumentation

1. A mass spectrometer (MS) equipped with an electrospray ionization source (ESI) and capable of tandem MS (MS/MS) analysis for online ion pairing normal phase liquid chromatography (IP-NPLC) monitoring and glycopeptide identification, for example a quadrupole time of flight instrument such as a QTOF2 or QTOF Ultima from Waters.
2. An MS equipped with an electrospray ionization source (ESI), capable of tandem MS (MS/MS) analysis, and equipped with electron transfer dissociation (ETD) for glycosylation site identification, for example an ion trap instrument such as the LTQ XL or LTQ XL Orbitrap from Thermo Scientific.

3. A capillary high performance liquid chromatography (HPLC) system capable of flow rates between 10 and 20 μL/min, such as an Agilent HP1100 with a pre-column flow splitter (0.5–1 mL/min split to 10–100 μL/min).
4. An online nanoflow liquid chromatography system, such as nanoAcquityUPLC (Waters) [22].
5. SpeedVac centrifuge.
6. Benchtop centrifuge.
7. Vortex.

2.4 Software for the Analysis of Mass Spectrometry Data

1. Typically, for *de novo* peptide sequencing, use the data visualization software that was provided by the mass spectrometry manufacturer (i.e., MassLynx for Waters instrumentation and Xcalibur for Thermo Fisher instruments).
2. Assisted *de novo* sequencing software is available, such as PEAKS Studio (Bioinformatics Solutions Inc).

3 Methods

Our approach begins with a proteome, subproteome, or affinity enriched sample. For protein containing samples, tryptic digestion is first required. If the sample contains affinity enriched MHC binding peptides in a mass spectrometry compatible buffer then proceed directly to Subheading 3.2. First, glycopeptides are enriched, and then glycopeptide containing fractions are analyzed using several tandem mass spectrometry approaches in order to determine the glycopeptide sequence, mass of the glycan and site of glycan modification. Figure 1 illustrates an overview of this workflow.

Fig. 1 (continued) by nLC-MS/MS or an aliquot tested for immune activation and only those fractions with immune activity investigated further. (*c*) Either peptides with activity or interest, or the total glycopeptide containing fractions are then analyzed by nano liquid chromatography tandem mass spectrometry to determine peptide sequence and glycan mass. It is not only of importance to map the peptide sequence, but also the site of amino acid modification. The presence of the mass of the glycan moiety means that traditional algorithms for protein identification from tandem mass spectrometry data are not always successful. Therefore in most cases *de novo* sequencing of glycopeptides, either using *de novo* sequencing software or by manual interpretation, is required for their identification. (*i*) Traditional collision induced dissociation (CID), breaks apart the peptide and generally permits mapping of some or all of the peptide sequence. The glycans, especially those that are *O*-linked, are typically labile under these conditions and are observed in the low *m*/*z* region of the spectrum as an intense glycan oxonium ion. (*ii*) Electron transfer dissociation (ETD) tandem mass spectrometry has gained popularity and has been exploited for mapping of amino acid modifications. There is a greater chance that the glycan moiety will be retained on the amino acid. However, the peptide fragmentation is highly dependent upon the nature of the amino acid sequence, and benefits from the presence of multiple charged amino acids. (*iii*) Recently, we have used soft collision induced dissociation (sCID), which gently fragments peptides and retains the glycan on the modified amino acid. Using a combination of these mass spectrometry peptide fragmentation approaches will provide peptide sequence information, the mass of the modifying glycan moiety and the site of modification

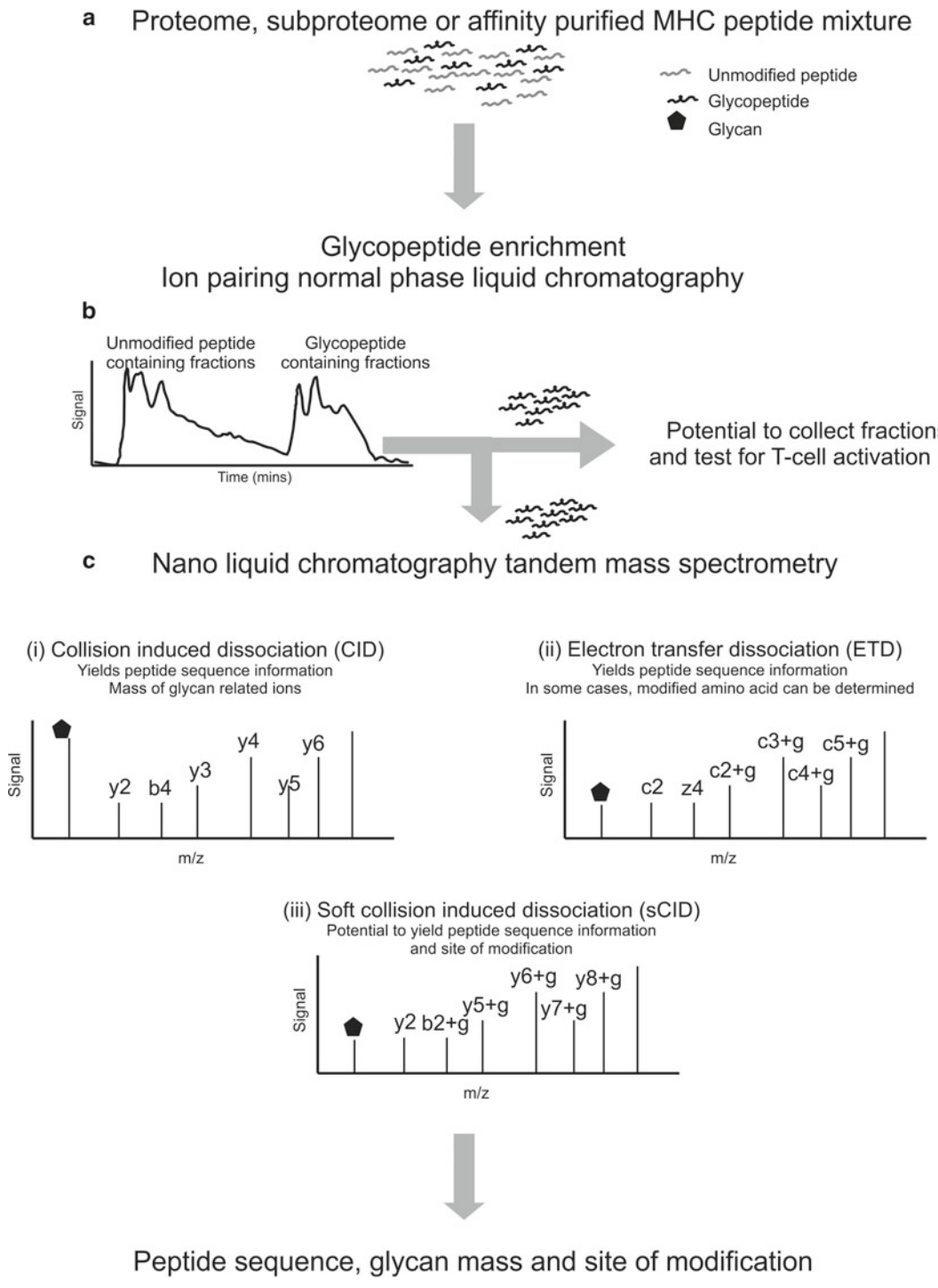


Fig. 1 Workflow for identification of glycopeptides and their associated glycan moiety. The starting material can be any proteome sample, or affinity purified peptide mixture, for example MHC bound peptides, purified using standard methods (for example ref. 25). The proteome sample is usually a complex mixture of proteins or peptides. Protein samples are typically digested with trypsin to yield mass spectrometry compatible peptides. Glycopeptides are usually the minority in this mixture and are not always readily identified. (*a*) Glycopeptides can be enriched from complex peptide mixtures using ion pairing normal phase liquid chromatography (IP-NPLC) [26], in which an ion pairing reagent is added prior to normal phase separations. Non-modified peptides elute first, followed by glycopeptide containing fractions. (*b*) The collected fractions may then immediately be analyzed

3.1 Preparation of Proteome Sample

1. Precipitate 100 or more µg of proteome sample (cell lysate) by incubation with four volumes of ice-cold acetone at −20 °C for 1 h (*see* **Note 1**).
2. Spin sample at maximum speed in a benchtop centrifuge and carefully remove acetone (*see* **Note 2**).
3. Air-dry the pellet in a laminar flow hood (*see* **Note 3**).
4. Reduce protein disulfide bonds by resuspending the pellet in 4 mM DTT solution. Incubate at 56 °C for 1 h.
5. Add iodoacetamide to a final concentration of 20 mM and incubate in the dark at room temperature for 1 h to alkylate the cysteine residues.
6. Repeat **steps 1**–**3**.
7. Add 10 µL of 50 mM ABC containing 2.5–5 µg of sequencing grade trypsin (30:1 protein to enzyme ratio (w/w)) and incubate at 37 °C for 14 h (*see* **Notes 4** and **5**).
8. Unless proceeding immediately, store digested proteome sample at −20 °C until required (*see* **Note 6**).

3.2 Enrichment of Glycopeptides Using Ion Pairing Normal-Phase Liquid Chromatography (IP-NPLC)

1. Assemble the capillary HPLC system with a polyhydroxyethyl aspartamide column (1 mm × 50 mm; 5 µm; 300 Å) such that the flow rate through the column is 12 µL/min. Arrange a post column splitter to allow for simultaneous fraction collection and real time monitoring of glycopeptide elution (*see* **Note 7**), with 1 µL/min of flow directed to the nano electrospray ionization (ESI) source of a mass spectrometer such as the QTOF 2 hybrid quadrupole time of flight (QTOF) instrument. Direct the remaining flow to fraction collection.
2. Dry peptide containing samples produced in Subheading 3.1 to a volume of 1 µL using a SpeedVac (*see* **Note 8**).
3. Dilute the sample with 8 µL of 90 % ACN and 1 µL of 10 % trifluoroacetic acid (TFA) for a final concentration of 1 % TFA (*see* **Note 9**).
4. Equilibrate the polyhydroxyethyl aspartamide column with 10 % IP-NPLC Buffer B using an equilibration time of 7 min.
5. Load diluted peptide sample with 10 % IP-NPLC Buffer B for 4 min.
6. Elute peptides using the following gradient: 10–30 % IP-NPLC Buffer B over 5 min, 30–60 % IP-NPLC buffer B over 5 min, 60–98 % IP-NPLC buffer B over 2 min,
7. Re-equilibrate the column with 10 % IP-NPLC Buffer B for 5 min.
8. Collect the first 10 min of the gradient as a single "non-glycopeptide" fraction. From 10 min onward, collect glycopeptide fractions at 30–60 s intervals depending on the degree of fractionation desired (*see* **Notes 10** and **11**).

9. Unless proceeding immediately, store fractions at −20 °C until required (*see* **Note 6**).

3.3 Tandem Mass Spectrometry for Identification of Enriched Glycopeptides

1. Using a centrifugal evaporator, reduce the volume of the glycopeptide containing target fractions obtained from Subheading 3.2, **step 8** to 1 μL (*see* **Note 8**) and then resuspend in 100 μL of 0.1 % formic acid (aqueous).
2. Using a nanoAcquity UPLC system coupled to a QTOF or an LTQ linear ion trap mass spectrometer with an ESI source, inject 10 μL of glycopeptide enriched fraction onto a 180 μm ID × 20 mm, 5 μm symmetry C18 trap column in trapping mode diverting flow through to waste.
3. Elute by reversed phase liquid chromatography (RPLC) to a 100 μm ID × 10 cm, 1.7 μm BEH130 C18 column in analytical mode using a linear gradient from 1 to 45 % solvent B in 18 min, 45–85 % solvent B for 2 min, 85–1 % solvent B over 1 min, and hold for 8 min at 1 % solvent B. For more complex samples, use a linear gradient from 1 to 45 % solvent B in 36 min, 45–95 % solvent B for 2 min, 95–1 % over 1 min, and hold for 9 min at 1 % solvent B (*see* **Note 12**).
4. Perform MS/MS in data dependent acquisition (DDA) mode (*see* **Note 13**).
5. Manually interpret MS/MS spectra by *de novo* sequencing to identify target glycopeptide ions and determine both the peptide and glycan sequences. Manual interpretation is often necessary because the additional mass of the attached glycan prevents most software algorithms from accurately identifying peptides (*see* **Note 14**). *De novo* sequencing involves manually determining sequential mass differences between peaks in an MS/MS spectrum that correspond to known amino acid and/or monosaccharide masses. For a detailed explanation of *de novo* sequencing, the reader is referred to [23].
6. Record glycopeptide ion m/z, charge state, and retention time for use in further targeted analyses (Subheading 3.4).

3.4 Mapping Glycan Linkage Sites Using ETD or Soft CID

1. Trap-load 10 μL of the glycopeptide containing fraction of interest as determined in Subheading 3.3, **step 6** onto a 180 μm ID × 20 mm, 5 μm symmetry C18 trap column using a nanoAcquity UPLC system coupled to an LTQ XL or LTQ Orbitrap mass spectrometer with a nanoESI source and capacity for ETD.
2. For low complexity peptide mixtures, elute sample through a 100 μm ID × 100 mm, 1.7 μm BEH130 C18 column with the following gradient: 1–45 % solvent B in 18 min, 45–85 % solvent B in 3 min, 85–1 % solvent B in 1 min. Re-equilibrate for 8 min at 1 % solvent B. Solvent A is composed of 0.1 % FA in water.

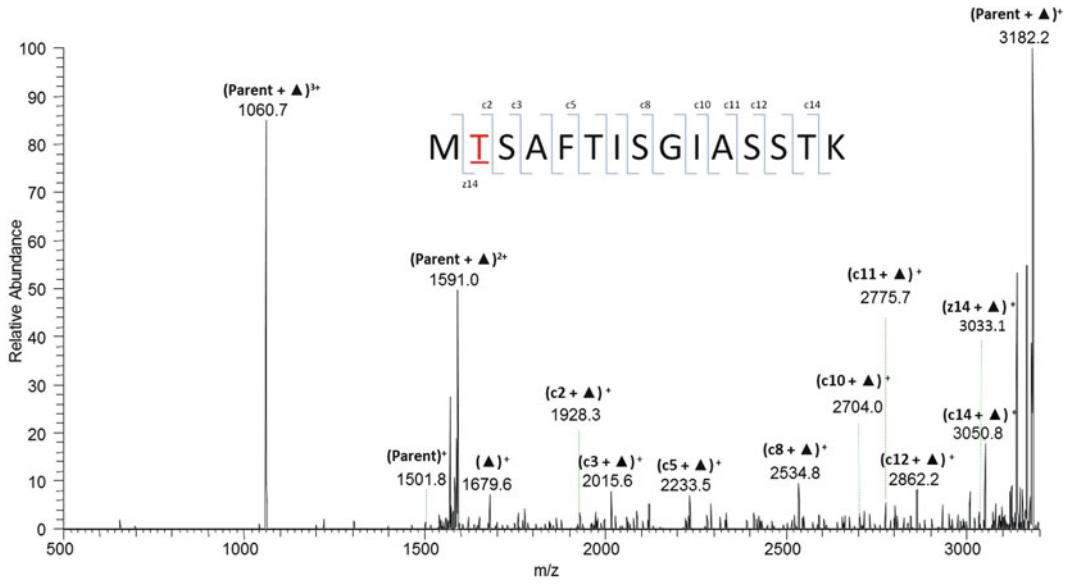


Fig. 2 Electron transfer dissociation fragmentation of flagellin glycopeptides from *Aeromonas hydrophila* AH3. ETD fragmentation is often the method of choice for the sequencing of a peptide with a labile modification, since the anion gas-phase ion-ion chemistry induced fragmentation leaves the weakly bound modifications linked to the peptide backbone. This produces not only peptide sequence information, but identifies the site of modification. This technology was used to identify the site of glycan modification on a tryptic peptide from the polar flagellin protein of *A. hydrophila* AH3. A targeted ETD MS/MS analysis was carried out on a triply charged ion at *m/z* 1,060.7, an ion which had been identified by traditional CID MS/MS as the ^{160}MTSAFTISGIASSTK^{174} glycopeptide. Due to its charge state (triply protonated), two analyses were carried out, one with a reaction time of 300 ms, the other 500 ms. Both runs were done with supplemental activation, a feature of the LTQ XL which adds a small amount of CAD energy to improve overall fragmentation. In this case, the 300 ms reaction between analyte and the gas anion with supplemental activation produced the best results indicating that T^2 is the site of modification. This research was originally published in The Journal of Biological Chemistry. Whilhelms, M., Fulton, K.M., Twine, S.M., Tomás, J.M., and Merino, S. Differential Glycosylation of Polar and Lateral Flagellins in *Aeromonas hydrophila* AH-3. *Journal of Biological Chemistry.* 2012; 287(33):27851–62. © The American Society for Biochemistry and Molecular Biology

3. For high complexity peptide mixtures, elute sample in analytical mode through a 100 μm ID × 100 mm, 1.7 μm BEH130 C18 column with the following gradient: 1–45 % solvent B in 36 min, 45–95 % solvent B in 2 min, 95–1 % solvent B in 1 min. Re-equilibrate for 10 min at 1 % solvent B. Solvent A is composed of 0.1 % FA in water.
4. For targeted ETD sequencing on an LTQ-XL, set the mass range over which the data should be acquired to encompass the mass of the parent glycopeptide (*see* **Note 15**). The isolation width should be set at 3.0. The ETD reaction time should be set between 100 and 800 ms depending on the size, composition, and charge state of the parent ion (*see* **Note 16**). Figure 2 shows a representative ETD MS/MS spectrum in which the glycosylation site was successfully determined.

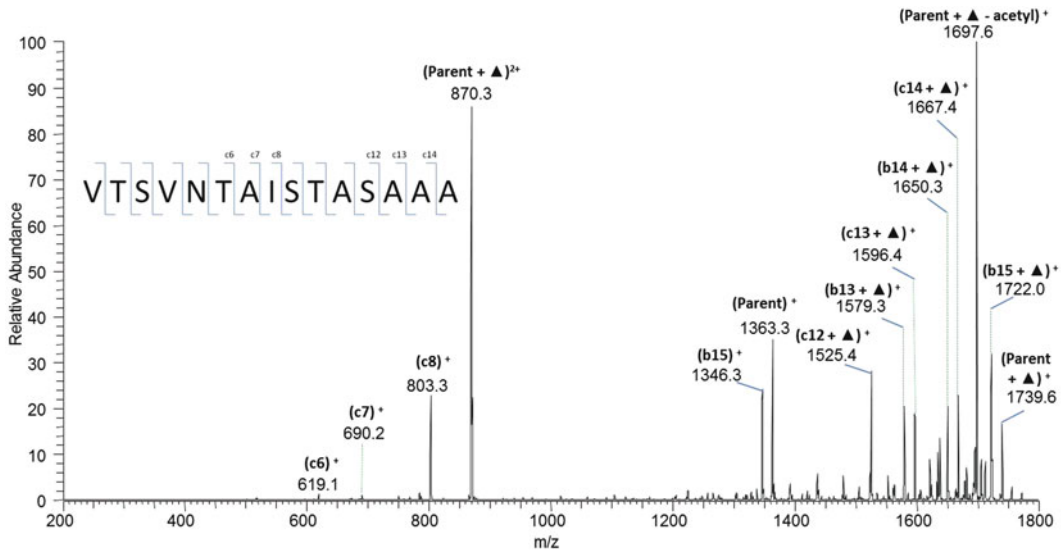


Fig. 3 Unsuccessful mapping of glycan modification sites of *Aeromonas hydrophila* AH3 lateral flagellin using electron transfer dissociation. Spectral quality drops with reduced peptide/glycopeptide charge, as seen in the targeted ETD MS/MS scan of a doubly charged tryptic glycopeptide from the lateral flagellin protein of *Aeromonas hydrophila* AH3. The 870.4^{2+} *m/z* ion, previously identified as the glycopeptide ^{170}VTSVNTAISTASAAA^{184} by CID MS/MS, produced poor fragmentation even with a long reaction time (500 ms) and supplemental activation. The spectrum even shows evidence of fragmentation of the glycan modification. While some useful sequence information can be obtained through a careful study of the ETD spectrum, it is not possible to confirm whether the glycan modification is linked to S^9 or T^{10}

5. In some cases, ETD MS/MS is not successful (*see* Fig. 3), and soft-CID (sCID) MS/MS offers an alternative (*see* **Note 17**). For targeted sCID sequencing on the LTQ-XL, as with the ETD experiment, set mass range to correspond with the size of the parent ion and set the isolation width to 3.0. The activation Q and reaction time should remain on the defaults 0.25 and 30, respectively. Set the collision energy such that it will fragment the parent, but not to the extent of removing glycan modification(s). If the parent ion is between 800 and 2,000 Da, start by setting "Normalized Collision Energy" to 25 V and evaluate the resultant fragmentation (*see* Fig. 4 and **Note 18**).
6. Manually interpret the MS/MS spectra of the ETD and/or sCID spectra using *de novo* sequencing techniques to determine the site of modification. This will require calculation of the predicted fragment ions for the known peptide sequence of interest (*see* **Notes 14** and **19**). The site of modification will be at the first amino acid where the mass difference between peaks corresponds to the mass of the amino acid plus the glycan mass.

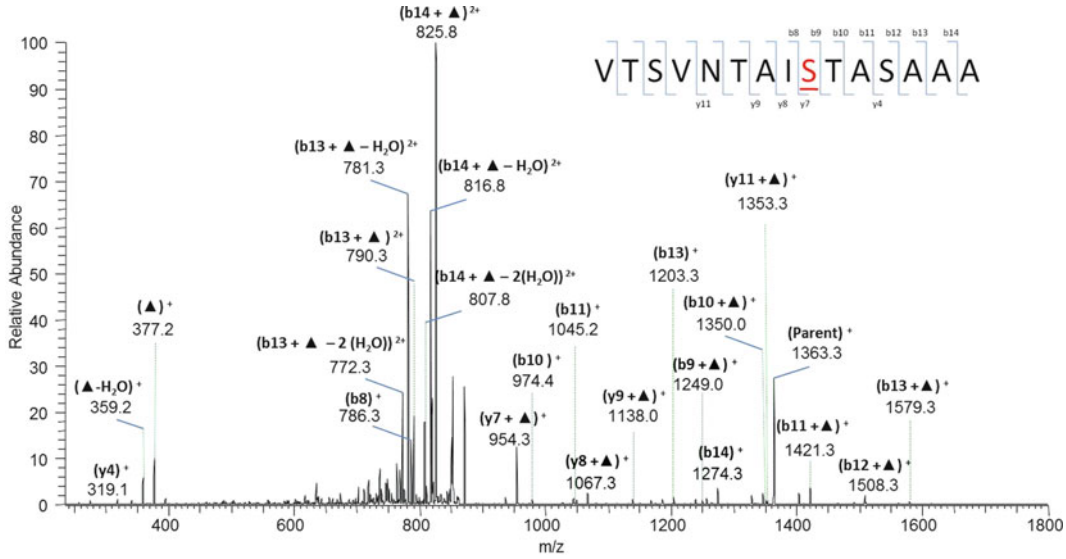


Fig. 4 Successful mapping of glycan modification sites of *Aeromonas hydrophila* AH3 lateral flagellin using soft CID. Two targeted High-Low CID experiments were carried out to determine the position of the glycan modification for the ^{170}TVTSVNTAISTASAAA^{184} glycopeptide that could not be determined in Fig. 2. Since CID spectra on the LTQ-XL are acquired at a much faster rate than ETD (<50 ms), two scan events were acquired over the same LC run. On the first LC run, we set the normalized collision energy (CE) at 18 V for the first scan event, and 25 V for the second. The resulting spectra from the scan event with a CE of 18 V did not produce sufficient fragmentation, while the scan with a CE of 25 V produced a spectrum lacking peptide + glycan fragments as the labile modification was disrupted. For the second LC run, the CEs were set at 20 and 24 V. 20 V was the optimal setting, providing a spectrum with a number of daughter ions where the glycan modification remained on the peptide chain, clearly showing that the glycan modification was on the S^9. This research was originally published in The Journal of Biological Chemistry. Whilhelms, M., Fulton, K.M., Twine, S.M., Tomás, J.M., and Merino, S. Differential Glycosylation of Polar and Lateral Flagellins in *Aeromonas hydrophila* AH-3. *Journal of Biological Chemistry.* 2012; 287(33):27851–62. © The American Society for Biochemistry and Molecular Biology

4 Notes

1. The acetone must be ice cold; typically leave acetone in the freezer or on ice for >1 h. After addition of the ice cold acetone to the protein sample, vortex vigorously and you will observe white flocculate. This can then be incubated on ice for 1–12 h to obtain maximal protein precipitation. This step is key to removing all detergents and chaotropic agents that may have been used in proteome preparation and subsequently interfere with trypsin digestion. This step is not required if the target cells are gently lysed or sonicated in low salt, low or no detergent, and low urea containing solutions.
2. Take care not to disturb the protein pellet. In most cases, a visible white protein containing pellet will be seen.

3. This step removes residual acetone that may denature trypsin used in the next steps. This can also be performed using a SpeedVac, although care must be taken not to overdry the pellet. This can make the protein pellet hard to resuspend in the trypsin containing buffer.
4. Even if the protein pellet was not overdried, complete protein resuspension in ammonium bicarbonate without the aid of detergents can be challenging. Addition of trypsin (30:1 protein–enzyme ratio) and overnight incubation is sufficient to fully digest the soluble and insoluble protein. However, extended incubation of trypsin containing solution with protein samples or addition of excess trypsin can result in large amounts of trypsin autolysis products that may interfere with subsequent mass spectrometry steps.
5. Trypsin is a common proteolytic enzyme, with specific and predictable cleavage sites on the C-terminal side of arginine and lysine residues. Trypsin cleavage sites can be predicted using *in silico* digestion algorithms, for example MSDigest, from the University of California San Francisco (http://prospector.ucsf.edu/prospector/cgi-bin/msform.cgi?form=msdigest). Should trypsin be insufficient for generating glycopeptides of appropriate length and charge for nLC-MS/MS analysis, alternate enzymes with varying cleavage sites are available. AspN, GluC, and LysC, for example have specific and predictable cleavage sites, while proteinase K, pronase, and thermolysin are nonspecific. Follow manufacturer's instructions for appropriate protein–enzyme ratios and buffers to be used in each case.
6. Storing samples at −20 °C between stages is acceptable, but freeze-thawing of samples throughout the procedure should be minimized as much as possible.
7. A real time MS chromatogram can be used to trace precursor glycopeptide ions of interest and potentially track highly labile glycan associated ions that may be visible in MS scans.
8. Do not completely dry the sample, as this will make resuspension of the pellet difficult.
9. Add the TFA only just be performing IP-NPLC analysis.
10. Fractions may be collected manually or using a fraction collector as desired.
11. It is advised to try a small scale separation to determine when glycopeptides elute, then scale up once parameters are optimized.
12. The HPLC system and gradient described here are specific for the setup that we routinely used. However, online separation of glycopeptides can be accomplished using a variety of HPLC

systems. Therefore conditions and gradients should be optimized according to the experiment when possible. We recommend using a nanoHPLC or nanoUPLC system, with C8 and/or C18 trapping and desalting prior to analytical separation using a 1.7 μm C18 column. These HPLC/UPLC systems, traps, and columns are available through companies such as Agilent, Waters, and Dionex/Thermo Fisher Scientific. Also note that the vials used for sample injection must be appropriate for the HPLC system chosen.

The proposed gradients for simple and complex samples are usually adequate for most uncharacterized samples. However, when targeting a specific glycopeptide, it may be necessary to shorten or lengthen the gradient, making the grade steeper or shallower respectively. It may also be beneficial in some instances to use a step-wise rather than linear gradient. Again, optimization is recommended.

13. For discovery and identification of glycopeptide ions in previously uncharacterized samples/fractions, a DDA approach should be taken to allow for an unbiased screening of the largest number of ions. Once glycopeptide ions are known, targeted MS/MS approaches can be used to confirm their presence in a particular fraction before proceeding to further characterization methods. The risk associated with targeted methods is that additional unknown glycopeptides may go undetected.
14. A few computational tools have been developed to help automate the sequencing of glycopeptides (i.e., PEAKS, SimGlycan, etc.). This has been more successful for eukaryotic glycopeptides that harbor a defined set of monosaccharides in various combinations. Across all prokaryotic organisms, however, there is a much larger repertoire of monosaccharides that can modify proteins, making the use of automated software more challenging. An excellent tutorial is available, that guides the reader step by step through the *de novo* sequencing of peptides [23]. For glycopeptide spectra, in many cases, glycan related fragment ions, particularly glycan oxonium ions, are observed as intense ions in the low m/z region of the spectrum. For example, an intense fragment ion at m/z 204.1 can indicate the presence of *N*-acetylhexosamine residues. In other cases, glycan masses may be observed as neutral losses from the parent ion in the MS/MS spectrum. In many spectra, the intensity and sometimes complexity of the glycan fragmentation make determination of the peptide sequence challenging when using traditional CID MS/MS.
15. Ion trap mass spectrometers acquire data within a dynamic range window, dependent on the precursor ion m/z. As a

result, the lower *m/z* limit should usually be set as low as this range will allow. The upper mass limit should be set to just exceed the *m/z* of the singly charged glycopeptide ion.

16. The transfer of a single electron from a donor anion in gas phase (in this case, fluoranthene) to a positively charged species (the peptide) initiates ETD fragmentation through an exothermic chain reaction that ultimately fragments the peptide backbone into *c* and *z* type ions with minimal disruption of posttranslational modifications. Therefore, ETD is most effective for highly protonated peptide ions [24]. Large, poorly charged ions with many neutral/hydrophobic amino acids will therefore require longer reaction times and the "supplemental activation" feature on the LTQ-XL can sometimes improve ETD fragmentation. However, longer reaction times run the risk of disrupting the glycan. Optimization of these setting for the specific glycopeptide of interest is advised.
17. Glycopeptides are often large and minimally charged species. Despite optimization of reaction time, ETD is sometime insufficient for peptide sequencing and modification site identification. The most frequent problem is the production of charge reduced precursor ions. In these cases, the donated electron fails to initiate the chain reaction necessary for peptide fragmentation and instead only displaces a single hydrogen atom. For example, a doubly charged ion becomes a singly charged ion, without fragmenting the peptide backbone. Soft CID is an alternative MS/MS method that can sometimes fragment the glycopeptide without disrupting the labile modification, allowing for site identification. The method used for glycosylation site identification will be glycopeptide dependent, and usually requires testing and optimization.
18. Multiple scan events within a single analysis can be created on the LTQ-XL with different "Normalized Collision Energy" values. This will provide different fragmentation profiles for evaluation, one for each collision energy value. The optimal setting will be the one that provides sufficient peptide fragmentation for sequencing without disrupting the labile modification
19. Peptide fragment ions can be predicted from the known amino acid sequence using software such as the Protein/Peptide Editor of BioLynx within the Waters MassLynx program or one of several peptide fragment ion calculators available online (i.e., http://db.systemsbiology.net/proteomicsToolkit/FragIonServlet.html). For ETD experiments, *c* and *z* type fragmentation ions will be required. For soft-CID experiments, *y* and *b* type fragmentation ions will be required.

References

1. Himmler A, Hauptmann R, Adolf GR, Swetly P (1987) Structure and expression in *Escherichia coli* of canine interferon-alpha genes. J Interferon Res 7:173–183
2. Krueger KE, Srivastava S (2006) Posttranslational protein modifications: current implications for cancer detection, prevention, and therapeutics. Mol Cell Proteomics 5:1799–1810
3. Atassi MZ, Casali P (2008) Molecular mechanisms of autoimmunity. Autoimmunity 41: 123–132
4. Hetzer C, Dormeyer W, Schnolzer M, Ott M (2005) Decoding Tat: the biology of HIV Tat posttranslational modifications. Microbes Infect 7:1364–1369
5. Anderton SM (2004) Post-translational modifications of self antigens: implications for autoimmunity. Curr Opin Immunol 16: 753–758
6. Engelhard VH, Altrich-Vanlith M, Ostankovitch M, Zarling AL (2006) Post-translational modifications of naturally processed MHC-binding epitopes. Curr Opin Immunol 18:92–97
7. Petersen J, Purcell AW, Rossjohn J (2009) Post-translationally modified T cell epitopes: immune recognition and immunotherapy. J Mol Med (Berl) 87:1045–1051
8. Purcell AW, van Driel IR, Gleeson PA (2008) Impact of glycans on T-cell tolerance to glycosylated self-antigens. Immunol Cell Biol 86:574–579
9. Schietinger A, Philip M, Yoshida BA, Azadi P, Liu H, Meredith SC, Schreiber H (2006) A mutant chaperone converts a wild-type protein into a tumor-specific antigen. Science 314: 304–308
10. Chapman C, Murray A, Chakrabarti J, Thorpe A, Woolston C, Sahin U, Barnes A, Robertson J (2007) Autoantibodies in breast cancer: their use as an aid to early diagnosis. Ann Oncol 18:868–873
11. Chapman CJ, Murray A, McElveen JE, Sahin U, Luxemburger U, Tureci O, Wiewrodt R, Barnes AC, Robertson JF (2008) Autoantibodies in lung cancer: possibilities for early detection and subsequent cure. Thorax 63:228–233
12. Desmetz C, Bascoul-Mollevi C, Rochaix P, Lamy PJ, Kramar A, Rouanet P, Maudelonde T, Mange A, Solassol J (2009) Identification of a new panel of serum autoantibodies associated with the presence of in situ carcinoma of the breast in younger women. Clin Cancer Res 15:4733–4741
13. Tan HT, Low J, Lim SG, Chung MC (2009) Serum autoantibodies as biomarkers for early cancer detection. FEBS J 276:6880–6904
14. Lavrsen K, Madsen CB, Rasch MG, Woetmann A, Odum N, Mandel U, Clausen H, Pedersen AE, Wandall HH (2013) Aberrantly glycosylated MUC1 is expressed on the surface of breast cancer cells and a target for antibody-dependent cell-mediated cytotoxicity. Glycoconj J 30(3): 227–236
15. Wandall HH, Blixt O, Tarp MA, Pedersen JW, Bennett EP, Mandel U, Ragupathi G, Livingston PO, Hollingsworth MA, Taylor-Papadimitriou J, Burchell J, Clausen H (2010) Cancer biomarkers defined by autoantibody signatures to aberrant O-glycopeptide epitopes. Cancer Res 70:1306–1313
16. Pedersen JW, Blixt O, Bennett EP, Tarp MA, Dar I, Mandel U, Poulsen SS, Pedersen AE, Rasmussen S, Jess P, Clausen H, Wandall HH (2011) Seromic profiling of colorectal cancer patients with novel glycopeptide microarray. Int J Cancer 128:1860–1871
17. Guerry P, Ewing CP, Schirm M, Lorenzo M, Kelly J, Pattarini D, Majam G, Thibault P, Logan S (2006) Changes in flagellin glycosylation affect Campylobacter autoagglutination and virulence. Mol Microbiol 60: 299–311
18. Logan SM (2006) Flagellar glycosylation—a new component of the motility repertoire? Microbiology 152:1249–1262
19. Felix G, Duran JD, Volko S, Boller T (1999) Plants have a sensitive perception system for the most conserved domain of bacterial flagellin. Plant J 18:265–276
20. Gewirtz AT, Navas TA, Lyons S, Godowski PJ, Madara JL (2001) Cutting edge: bacterial flagellin activates basolaterally expressed TLR5 to induce epithelial proinflammatory gene expression. J Immunol 167:1882–1885
21. Hunt DF, Michel H, Dickinson TA, Shabanowitz J, Cox AL, Sakaguchi K, Appella E, Grey HM, Sette A (1992) Peptides presented to the immune system by the murine class II major histocompatibility complex molecule I-Ad. Science 256:1817–1820
22. Brown SL, Stockdale VJ, Pettolino F, Pocock KF, de Barros Lopes M, Williams PJ, Bacic A, Fincher GB, Hoj PB, Waters EJ (2007) Reducing haziness in white wine by overexpression of Saccharomyces cerevisiae genes YOL155c and YDR055w. Appl Microbiol Biotechnol 73:1363–1376
23. Addona T, Clauser K (2002) De novo peptide sequencing via manual interpretation of MS/MS spectra. Curr Protoc Protein Sci Chapter 16:Unit 16 11
24. Syka JE, Coon JJ, Schroeder MJ, Shabanowitz J, Hunt DF (2004) Peptide and protein sequence analysis by electron transfer dissocia-

tion mass spectrometry. Proc Natl Acad Sci U S A 101:9528–9533
25. Purcell AW (2004) Isolation and characterization of naturally processed MHC-bound peptides from the surface of antigen-presenting cells. Methods Mol Biol 251:291–306
26. Ding W, Nothaft H, Szymanski CM, Kelly J (2009) Identification and quantification of glycoproteins using ion-pairing normal-phase liquid chromatography and mass spectrometry. Mol Cell Proteomics 8: 2170–2185

Chapter 12

Whole-Cell MALDI-TOF Mass Spectrometry: A Tool for Immune Cell Analysis and Characterization

Richard Ouedraogo, Julien Textoris, Aurélie Daumas, Christian Capo, and Jean-Louis Mege

Abstract

Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) is widely used in proteomics. It has been recently demonstrated that MALDI-TOF MS can be used to identify and classify numerous bacterial species or subspecies. We applied MALDI-TOF MS directly to intact mammalian cells, and we found that this method is valuable to identify human circulating cells and cells involved in the immune response including macrophages. As macrophages are characterized by a high degree of plasticity in response to their microenvironment, we stimulated human macrophages with cytokines, bacterial products, and a variety of bacteria. We found that MALDI-TOF MS discriminated unstimulated and stimulated macrophages, and also detected multifaceted activation of macrophages. We conclude that whole-cell MALDI-TOF MS is an accurate method to identify various cell types and to detect subtle modifications in cell activity.

Key words Mass spectrometry, Matrix-assisted laser desorption/ionization time-of-flight, Intact cell, Macrophage, Cell activation

1 Introduction

Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) is a powerful tool for the analysis of ionized molecules (i.e., proteins) by measuring their mass/charge (m/z) ratio. This technique is currently used in biochemistry to identify peptides, proteins, posttranslational modifications of proteins [1, 2], and nucleic acids [2, 3]. Typically, in cell biology, using MALDI-TOF MS to analyze cellular protein composition requires a critical cell lysis step, as well as a variety of fractionation and separation steps, including affinity separation methods, gel electrophoresis, chromatographic separations [4]. Combined with gel electrophoresis, MALDI-TOF MS allows for the study of the proteome [5] and the identification of a large number of proteins

Kelly M. Fulton and Susan M. Twine (eds.), *Immunoproteomics: Methods and Protocols*, Methods in Molecular Biology, vol. 1061, DOI 10.1007/978-1-62703-589-7_12,

in the proteome, the secretome, and membranes from activated macrophages [6]. Seemingly, MALDI-TOF MS can identify M1 responses of macrophages, such as the response to interferon (IFN)-γ and lipopolysaccharide (LPS) [5]. Using pulse stable isotope labeling of amino acids in cell culture, Kraft-Terry et al. identified a bioprofile in *macrophage colony-stimulating factor*-differentiated monocytes that is consistent with an M2 profile [6]. However, these approaches based on cell component separation require large sample quantities and cannot be used to analyze clinical specimens. Various attempts have been made to study single mammalian cells by MALDI-TOF MS, but, to date, these methods have been proven fastidious, and the biological information extracted is limited [7].

New applications of the MALDI-TOF MS method have been introduced recently in bacteriology laboratories. The fingerprints of intact bacteria allow rapid identification and taxonomic classification of numerous bacterial species and subspecies [8–10]. Using databases established from isolated bacterial species, the identification of many bacterial species in clinical samples [11] is fast, easy to perform, and inexpensive.

The MALDI-TOF MS procedure for the identification of bacterial species was expanded to three mammalian cell lines in 2006 [4]. More recently, 66 cell lines, representing 34 species from insects to primates, have been identified by MALDI-TOF MS [12]; but this method involves ethanol inactivation and formic acid–acetonitrile extraction. We applied MALDI-TOF MS directly to intact cells, and found this method highly valuable to the identification of human circulating cells, and cells involved in the immune response, including macrophages [13]. In addition, macrophages have a high degree of plasticity, and adapt quickly in response to their microenvironment. They sense microorganisms through receptors that bind conserved and ubiquitous microbial motifs, such as LPS. Macrophages stimulated with type I cytokines, i.e., IFN-γ, Tumor Necrosis Factor (TNF), or bacterial products, i.e., LPS, adopt an M1 macrophage phenotype which is inflammatory, tumoricidal, and microbicidal. Macrophages stimulated with interleukin (IL)-4 or IL-10, adopt an M2 macrophage phenotype, which regulates inflammatory and immune responses, and is only weakly tumoricidal and microbicidal [14]. Hence, macrophages constitute a model of choice to assess the accuracy of whole-cell MALDI-TOF MS to detect subtle modifications in cell activity [15].

We describe here the effective use of MALDI-TOF MS to identify many intact eukaryotic cell populations by creating a database of known samples. The experimental protocol, bioinformatics analysis of whole-cell MALDI-TOF MS spectra, and the comparison of unknown samples to the database allow the identification of various cell types within heterogeneous samples [13] or multifaceted activation of macrophages [15].

2 Materials

The different cell types or stimulated cells were prepared separately to establish databases. Experiments with mixed cell types (circulating cells) are indicated. Prepare sterile solutions for cell isolation and culture. Prepare and store all reagents at 4 °C. Prepare MALDI-TOF matrix just before use.

2.1 Human Circulating Cells

1. Isolate peripheral blood mononuclear cells (PBMCs) from leukopacks (or blood donors) by Ficoll gradient (MSL, Eurobio), as previously described [16].
2. Prepare $CD14^+$ monocytes using CD14 MicroBeads and the MACS separation system (Miltenyi Biotec), according to the manufacturer's protocol.
3. Prepare T $CD3^+$ lymphocytes using CD3 MicroBeads and the MACS separation system (Miltenyi Biotec), according to the manufacturer's protocol.
4. Obtain monocyte-derived macrophages (MDMs) by incubating monocytes (10^6 cells in 6-well plates) in 3 mL of RPMI 1640 containing 20 mM HEPES, 10 % human serum AB^+, 2 mM L-glutamine, 100 IU/mL penicillin, and 100 μg/mL streptomycin for 4 days. Replace human serum with fetal calf serum (FCS) for 3 additional days. The obtained cell population was identified as macrophages (more than 95 %) by acquisition of membrane CD68 expression and CD14 down-modulation.
5. Obtain dendritic cells (DCs) by incubating monocytes with 1,000 U/mL human recombinant *granulocyte macrophage-colony stimulating factor* (GM-CSF) and 500 U/mL of human recombinant IL-4 in RPMI 1640 containing 10 % FCS, 2 mM L-glutamine, 100 IU/mL penicillin, and 50 μg/mL streptomycin for 7 days. Change the medium every 3 days and add again 1,000 U/mL GM-CSF and 500 U/mL IL-4. The obtained cells expressed high levels of CD11c and CD1a, and low levels of CD14 and CD68.
6. Obtain polymorphonuclear cells (PMNs) after Ficoll centrifugation, by sedimentation of red blood cells (RBCs) within dextran T500 (1.5 % (w/v), Pharmacosmos) (*see* **Note 1**).
7. Obtain red blood cells (RBCs) by 1/1,000 dilution of blood in Phosphate-buffered saline (PBS).

2.2 Noncirculating Cells

The preparation of several cell populations including mammalian primary cells and cell lines, one *Xenopus laevis* cell line (XTC-2 cells), and four types of amoebae (*Acanthamoeba polyphaga*, *Acanthamoeba castellanii*, *Hartmannella vermiformis*, *Poteriochromonas melhamensis*) is described in ref. 13.

2.3 Macrophage Stimulation

1. Stimulate MDMs with 20 ng/mL of human recombinant IFN-γ, IL-4, IL-10, or TNF (purchased from R&D Systems or other suppliers) for different time points. Also stimulate MDMs with 1 μg/mL LPS from *Escherichia coli* or heat-killed bacteria (50 bacteria per cell) including *Mycobacterium tuberculosis*, *Mycobacterium bovis*, *Mycobacterium avium*, *Rickettsia prowazekii*, and *Orientia tsutsugamushi* (*see* **Note 2**)
2. *M. tuberculosis* (CIP H37Rv strain 103471), *M. bovis* (Bacillus Calmette-Guérin, BCG CIP strain 671203), and a clinical isolate of *M. avium* subsp. *hominissuis* were obtained from the Laboratory of Microbiology of the Hospital La Timone (Marseille) [17].
3. *R. prowazekii* strain Breinl (ATCC VR-142), the agent of epidemic typhus [18], and *O. tsutsugamushi* strain Kato (CSUR R163) [19], the agent of scrub typhus, were propagated in L929 cells (*see* **Note 3**).
4. Heat-killed bacteria were obtained after heating microorganisms at 95 °C for 1 h.

2.4 Reagents and Materials for MALDI-TOF MS

1. The matrix solution consists of a 10 mg/mL solution of acid-α-cyano-4-hydroxy cinnamic (HCCA) diluted in 500 μL acetonitrile, 250 μL Milli-Q grade water, and 250 μL trifluoroacetic acid at 10 %. Mix and sonicate for at least 20 min. Centrifuge at 13,000 × *g* for 5 min. Discard the pellet and keep the supernatant. The matrix solution is ready for use (*see* **Note 4**).
2. AutoFlex II mass spectrometer (Bruker Daltonics).
3. Bruker MSP 384 software polished steel target (Bruker Daltonics).
4. FlexControl 3.0 software (Bruker Daltonics).
5. FlexAnalysis 3.3 software (Bruker Daltonics).
6. Biotyper 3.0 software (Bruker Daltonics).
7. ClinProTools 2.2 software (Bruker Daltonics).

3 Methods

3.1 MALDI-TOF Target Preparation

1. Moisten the Bruker MSP 384 polished steel target with hot tap water. Rub with KIMTECH paper. Rub with 70 % ethanol. Rinse with hot tap water by rubbing with KIMTECH paper. Rub finally with 70 % ethanol.
2. Immerse the target in 70 % ethanol and sonicate for at least 15 min.
3. Cover the target with 500 μL to 1 mL of trifluoroacetic acid. Rub with KIMTECH paper. Rinse with high performance liquid chromatography (HPLC) grade water without rubbing. Dry the target at room temperature (*see* **Note 5**).

3.2 Preparation of Deposits

1. Centrifuge cells (2×10^6 cells per assay) at $300 \times g$ for 5 min and wash them in PBS without Ca^{2+} or Mg^{2+} (*see* **Note 6**). Centrifuge to remove traces of culture medium. Collect cell pellets in 20 µL of PBS without Ca^{2+} or Mg^{2+}. Freeze cells for 2–3 days before analysis at −80 °C.
2. Thaw gently samples on ice (4 °C) (*see* **Note 7**).
3. Put the Bruker MSP 384 polished steel target on a horizontal support to obtain uniform deposits throughout the spot (*see* **Note 8**).
4. Homogenize cells in Eppendorf tubes before deposition of 1 µL on the MALDI target. Add 1 µL of the HCCA matrix to the sample on the target. Avoid mixing the spot with the pipette (*see* **Note 9**).
5. Drop 12–16 different spots of the same sample.
6. The evaporation takes place gradually at room temperature, and the formed HCCA crystals contain dispersed sample molecules.
7. Samples may be immediately analyzed or stored in the dark for several days before analysis.

3.3 Acquisition of Data

1. Insert the Bruker MSP 384 polished steel target containing samples in the Autoflex II mass spectrometer (Bruker Daltonics) outfitted with the Compass 1.2 software suite (consisting of FlexControl 3.0 and FlexAnalysis 3.3 from Bruker Daltonics). Run samples in positive mode, with 240 laser satisfactory shots in 40 shot steps intervals and 40 % laser power, performed in different regions of the analyzed sample spot. A signal-to-noise ratio of 3.0 was selected to define peaks, with a maximum of 100 peaks per spectrum. After the target plate calibration was complete, the AutoExecute command was used to analyze the samples. The processes described below are manufacturer or software defaults, and do not require adjusting.
2. Laser settings. Fuzzy Control, On; Weight, 2.00; Laser power, between 30 and 45 %; Matrix Blaster, 5.
3. Data Evaluation. Peak Selection Masses, 4,000–10,000; Mass Control List, Off; Peak Exclusion, ignore the largest peaks in the defined mass range; Peak Evaluation Processing Method, Default; Smoothing: On; Base-line Subtraction, On; Peak, Resolution higher than 400; Protein maximal resolution, ten times above the threshold.
4. Accumulation. Parent Mode, On; Sum up to 240 satisfactory shots in 40 shot steps; Dynamic Termination, On; Criteria Intensity, Early termination if reaching intensity value of 20,000 for ten peaks.

5. Movement. Spiral large, Maximal allowed number at one raster position; Parent Mode, 80; Quit sample after 25 subsequently failed judgments.
6. Processing. Flex analysis. Method, none; BioTools MS method, none.
7. Sample Carrier. Manual fine control speed, *x* (10,000 μm/s) *y* (20,000 μm/s); Relative *x* (5 μm) *y* (5 μm); Absolute *x* (−2,000 μm) *y* (−2,000 μm); State *x* (0) *y* (0); Random walk, 50 shots at raster spot; Mode, off.
8. Spectrometer. High voltage, switched On; Ion Source 1, 19,99 kV; Ion Source 2, 18,74 kV; Lens, 7 kV; Pulsed Ion Extraction, 330 ns; Polarity, Positive; Matrix Suppression, Mode Gaiting; Gaiting strength, height; Suppress up to, *m/z* 1,500 Da.
9. Detection. Mass range, 2,000–20,137; Mode, Medium Range; Detector Gain, Linear (18×); Sample Rate, 1.00 GS/s; Electronic Gain, Enhanced (100 mV); Real-time Smooth, High; Spectrum, Size (63,463 pts), Delay (29,412 pts).
10. Processing Method. MBT process.
11. Setup. Range, Medium; Laser Frequency, 25 Hz; Autoteaching, On; Instrument-specific Settings: Digitizer Trigger Level (2,000 mV), Digital off Linear (127 cnt), Digital off Reflector (127 cnt); Detector Gain Voltage Offset, Linear (1,300 V), Reflector, 1,400 V; Laser Attenuator, Offset (75 %), Range (15 %); Electronic Gain Button Definitions: Gain, regular (Offset Lin, 100 mV; Offset Ref 100 mV; 200 mV/full scale); enh. (Offset Lin, 51 mV; Offset Ref 51 mV; 100 mV/full scale); highest (Offset Lin, 25 mV; Offset Ref 25 mV; 50 mV/full scale).
12. Calibration. Calibration strategy, Interactive; Mass Control List, Bacterial test standard; Zoom Range, ±5 %; Peak Assignment Tolerance, User Defined (1,000 ppm); Mode, linear.

3.4 Data Analysis

3.4.1 Spectrum Analysis

1. The FlexAnalysis software 3.3 allows raw spectrum processing, baseline subtraction, smoothing, peak list editing, and displays several spectra into one window or superimposes spectra (for the comparison of different types of circulating cells, *see* Fig. 1).
2. The ClinProTools 2.2 software from Bruker Daltonics is used to analyze the variability between different samples. Load spectra of each cell category to create according classes.

   2D representation generated by ClinProTools 2.2. The software selects automatically two peaks that are present in each cell type but have different intensities (here, the peaks 2 and 30). This representation highlights the reproducibility of spectra between spectra of each class (here, ten spectra obtained from ten different blood donors), and the differences between

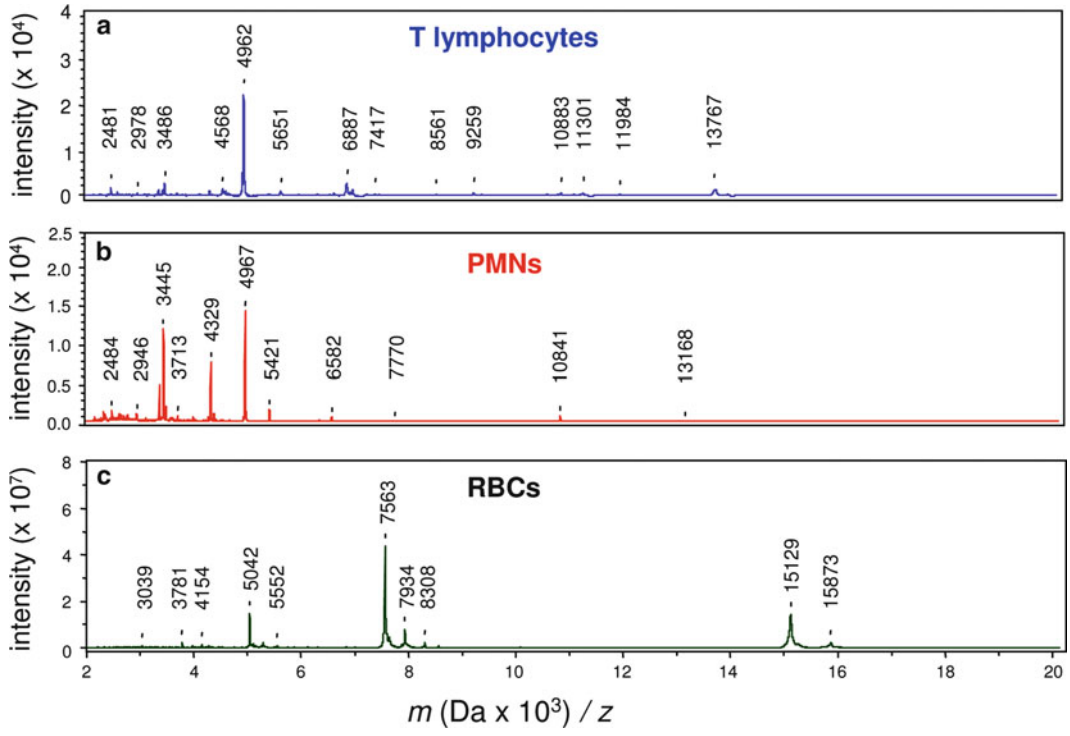


Fig. 1 MALDI-TOF MS spectra of circulating cells. T lymphocytes (**a**), PMNs (**b**), and RBCs (**c**) were isolated from a healthy blood donor. Representative MALDI-TOF MS spectra are shown. The figure is extracted from the ref. 13

the different classes (here, monocytes, T lymphocytes and PMNs) (*see* Fig. 2).

3. Gel-view representation generated by ClinProTools 2.2. This representation compares the reproducibility of spectra within the same class, and the differences between different classes. The different bands represent different peaks of each class. The intensity of bands corresponds to the intensity of detected peaks. This representation shows the reproducibility of spectra within each class (here, four spectra obtained from four different blood donors) and the peaks that are differentially expressed in different classes (here, monocytes and T lymphocytes isolated from each blood donor) (*see* Fig. 3).

3.4.2 Database Creation

1. The Biotyper 3.0 software from Bruker Daltonics is used to create and manage databases. An averaged spectrum for each cell category corresponds to at least ten individual spectra. Here, we created a database that includes 17 mammalian cell types, one *X. laevis* cell line (XTC-2 cells), and four types of amoebae (*A. polyphaga*, *A. castellanii*, *H. vermiformis*, *P. melhamensis*) (*see* **Note 10**).

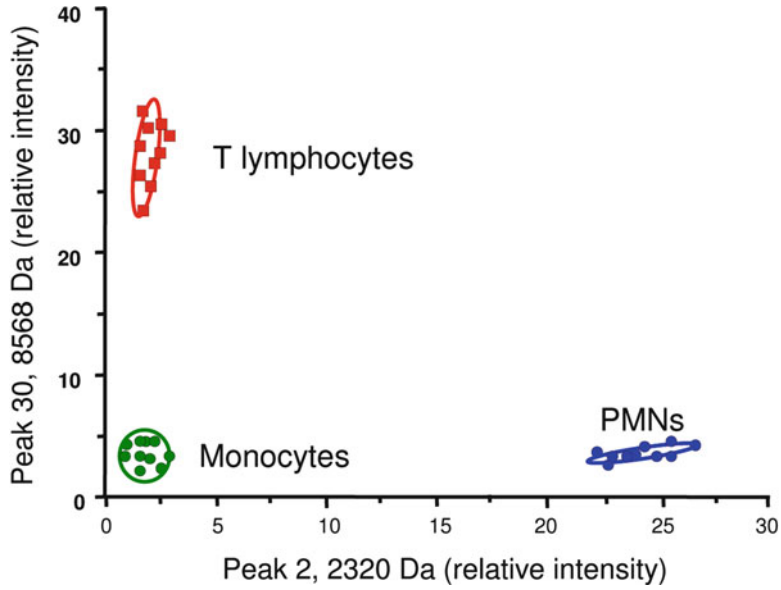


Fig. 2 Reproducibility of MALDI-TOF MS signatures. Monocytes, T lymphocytes, and PMNs were isolated from ten healthy blood donors. MALDI-TOF MS spectra were analyzed using 2D Peak Distribution View. The relative intensities of the two peaks automatically selected were homogenous among blood donors, and the *ellipses* represent the standard deviation within each cell population (monocytes, T lymphocytes, and PMNs, respectively). *See* ref. 13

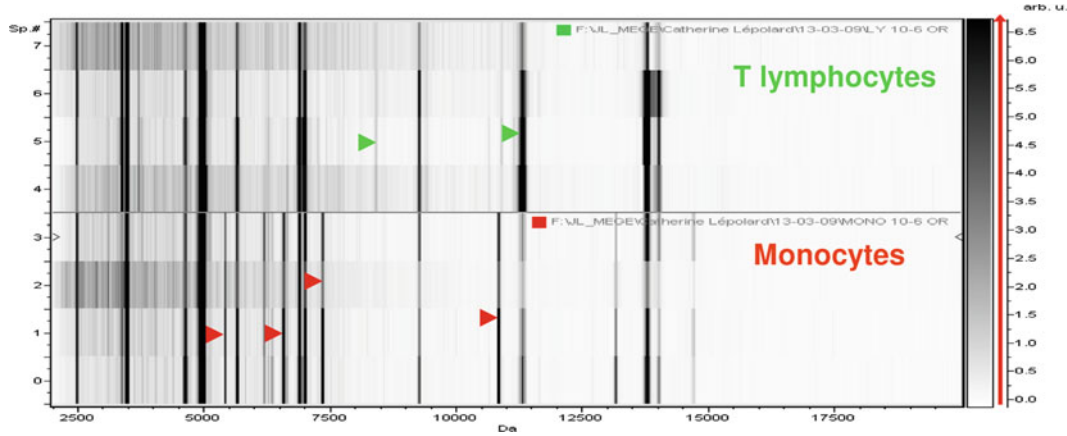


Fig. 3 Gel view representation of monocytes and T lymphocytes. Circulating cells were isolated from four different healthy blood donors. MALDI-TOF MS spectra are presented in Gel View representation. Spectra are shown with *m*/*z* values on the *x*-axis and the peak intensity (in arbitrary units) is coded with the grey scale presented on the *right*. Major differences between monocytes and T lymphocytes are indicated by *arrowheads*. *See* ref. 13

2. Baselines are automatically subtracted from spectra, and the background noise smoothed. An average spectrum is automatically created using default Biotyper method settings provided by the manufacturer. The sensitivity (the maximum tolerated

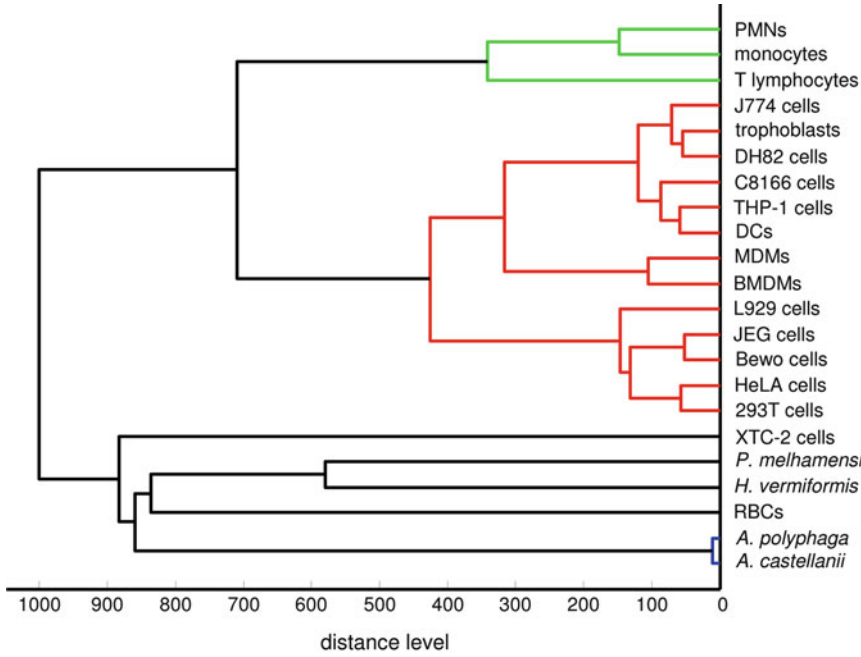


Fig. 4 Dendrogram of 22 eukaryotic cell types. MALDI-TOF MS was performed on 22 cell types with at least 20 spectra per cell type. A mean spectrum for each cell type was added to the database using the BioTyper 3.0 software and the dendrogram creation method. *See* ref. 13

error) of mass spectrum values and spectrum shifts was 200 particles per million. The minimum frequency to benchmark selection of peaks was 25 %, and only peaks with signal/noise intensity above background are automatically selected by the software. An average virtual spectrum consisting of 70 peaks is added to the database as a new reference.

3. The Biotyper 3.0 software is used to generate a dendrogram representation of cell categories according to their protein fingerprint (*see* Fig. 4). This figure shows that mammalian cells (in red and green) and nonmammalian cells (in black and blue) were in two distinct branches of the dendrogram. Circulating nucleated cells (in green) clustered within a subbranch distinct of primary cells (trophoblasts, DCs, MDMs, murine bone marrow-derived macrophages (BMDMs)) and cell lines. Note that human RBCs, which are unnucleated cells, clustered with nonmammalian cells.
4. The Biotyper 3.0 software is used to identify unknown spectra by comparison with database references. The sensitivity and the maximum error tolerated to determine the mean m/z values are 1 Da. Score values between 0.000 and 1.699 indicate that the unknown spectra did not match with known references. Values between 1.700 and 1.899 indicate probable cell identification.

Table 1
Identification of subpopulations by MALDI-TOF MS

Multiple cell population	Identification	Scores
Monocytes + T lymphocytes (equal concentration)	Monocytes	2,250
	T lymphocytes	2,247
PBMCs	Monocytes	2,078
	T lymphocytes	2,024
Whole blood (after hypotonic shock)	PMNs	2,049
	Monocytes	1,585
	T lymphocytes	1,654

The fingerprints of monocytes and T lymphocytes were identified in a mixed population (50 % monocytes, 50 % T lymphocytes) and in PBMCs. In blood, it was possible to identify PMNs (that represent about 70 % of total leukocytes), but not monocytes or T lymphocytes

Scores between 2.000 and 3.000 are considered statistically significant, and allow effective identification of the unknown spectra [13]. This procedure is currently used to identify bacterial species in clinical samples [11]. Here, we extend this method and the scores provided by Bruker Daltonics to identify the different cell populations present in a complex tissue. First, isolated monocytes and T lymphocytes were mixed, and the resulting fingerprint shows that monocytes and T lymphocytes are identified by MALDI-TOF MS (Table 1) The Table 1 also shows that the fingerprints of monocytes and lymphocytes T, respectively, are identified in peripheral blood mononuclear cells. In whole blood that contains leukocytes essentially composed of PMNs, the fingerprint of PMNs is identified, but not those of monocytes and T lymphocytes.

3.5 Macrophage Activation Analysis with R Software

The analysis of MALDI-TOF MS spectra described above is performed using Bruker Daltonics software. We present here a similar analysis performed with an open-source software (R), and specific algorithms that are presented as supplementary material in a recently published manuscript [15].

1. Load raw spectra in R (version 2.14) using the readBrukerFlexData library.
2. Analyze spectra using the MALDIquant library and specific algorithms. The square root of the intensities is used to enhance graphical visualization of the spectra.
3. Correct background using Statistics-sensitive Nonlinear Peak-clipping algorithm for baseline estimation [20]. Peaks are detected using a Signal-to-Noise Ratio of 6.0. The detected

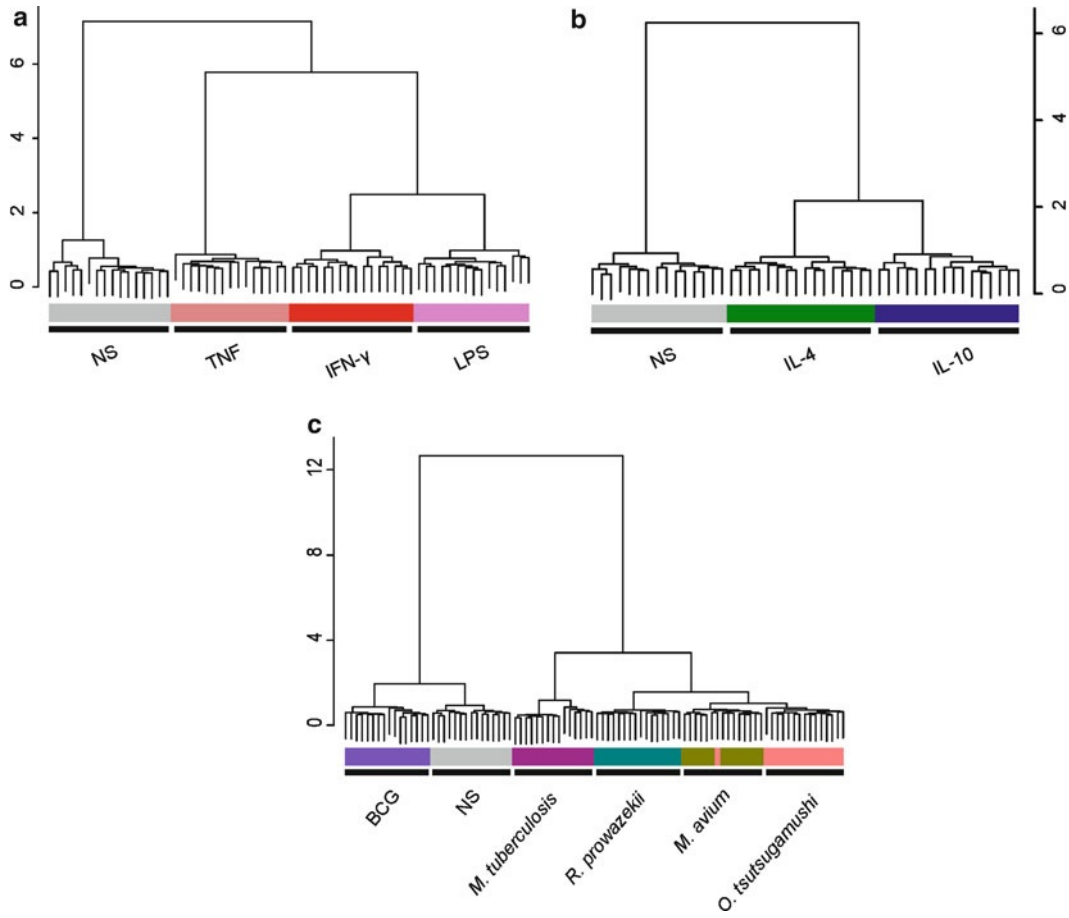


Fig. 5 Hierarchical clustering of activated macrophages. Monocyte-derived macrophages were stimulated with M1-related agonists (**a**), M2-related agonists (**b**), and intracellular bacteria (**c**) for 18 h. Unstimulated macrophages are presented in *grey* (NS)

peaks are considered similar across spectra when the m/z values are within a 2,000 ppm window.

4. Use hierarchical clustering to classify the spectra, ward algorithm for agglomeration, and a dissimilarity matrix based on the Jaccard distance. The Jaccard index measures similarity between boolean sample sets. The Jaccard distance, which measures dissimilarity between sample sets, is complementary to the Jaccard index and is obtained by subtracting the Jaccard coefficient from 1 or by dividing the difference of the sizes of the union and the intersection of two sets by the size of the union. This procedure was used to discriminate unstimulated (in grey) and stimulated macrophages (in colors) (*see* Fig. 5). Note that the responses of macrophages to different M1 agonists clustered but were not superimposable (*see* Fig. 5a). Similarly, the responses of macrophages to IL-4 and IL-10, two M2 agonists,

clustered compared to unstimulated macrophages but were distinct (*see* Fig. 5b). Different intracellular bacteria induced specific signatures (*see* Fig. 5c). Taken together, these results show that MALDI-TOF MS detected the multifaceted activation of macrophages.

4 Notes

1. PMNs must be isolated from remaining RBCs after dextran T500 sedimentation. Lyse RBCs by a 30 s hypotonic shock to obtain pure PMNs. In the absence of lysis, RBC signals were detected in MALDI-TOF MS and masked the detection of PMNs.
2. The stimulation of human MDMs is usually performed in RPMI 1640 supplemented with 10 % FCS, 100 UI/mL penicillin and 50 µg/mL streptomycin.
3. Wash bacteria with PBS to remove the components contained in growth media (such as serum proteins) that may interfere with MALDI-TOF MS spectra.
4. A matrix solution containing crystals does not allow a good ionization of sample molecules, and may affect the quality of spectra.
5. A target improperly cleaned may bias the results. It is therefore very important to take the time to carefully clean targets.
6. Cells may agglutinate in the presence of Ca^{2+} and Mg^{2+}. In addition, salts may interfere with MALDI-TOF MS.
7. Rapid and vigorous thawing alters samples, thus affecting the MALDI-TOF MS analysis.
8. Homogeneous deposits are necessary to obtain reproducible and high-quality spectra.
9. Mixing spots with pipettes alters spectrum quality. It is therefore important to respect the proportions and indicated details.
10. Each laboratory needs to construct its own databases before comparing cell populations and looking at unknown samples. We may export our databases to other laboratories.

Acknowledgments

We thank Nicolas Armstrong, Carine Couderc, Philippe Decloquement, and Christophe Flaudrops for their technical assistance.

References

1. Johnson RP, El-Yazbi AF, Hughes MF, Schriemer DC, Walsh EJ, Walsh MP et al (2009) Identification and functional characterization of protein kinase A-catalyzed phosphorylation of potassium channel Kv1.2 at serine 449. J Biol Chem 284:16562–16574
2. Li X, Cowles EA, Cowles RS, Gaugler R, Cox-Foster DL (2009) Characterization of immunosuppressive surface coat proteins from *Steinernema glaseri* that selectively kill blood cells in susceptible hosts. Mol Biochem Parasitol 165:162–169
3. Mellmann A, Cloud J, Maier T, Keckevoet U, Ramminger I, Iwen P et al (2008) Evaluation of matrix-assisted laser desorption ionization-time-of-flight mass spectrometry in comparison to 16S rRNA gene sequencing for species identification of nonfermenting bacteria. J Clin Microbiol 46:1946–1954
4. Zhang X, Scalf M, Berggren TW, Westphall MS, Smith LM (2006) Identification of mammalian cell lines using MALDI-TOF and LC-ESI-MS/MS mass spectrometry. J Am Soc Mass Spectrom 17:490–499
5. Brown J, Wallet MA, Krastins B, Sarracino D, Goodenow MM (2010) Proteome bioprofiles distinguish between M1 priming and activation states in human macrophages. J Leukoc Biol 87:655–662
6. Kraft-Terry SD, Gendelman HE (2011) Proteomic biosignatures for monocyte-macrophage differentiation. Cell Immunol 271: 239–255
7. Rubakhin SS, Churchill JD, Greenough WT, Sweedler JV (2006) Profiling signaling peptides in single mammalian cells using mass spectrometry. Anal Chem 78:7267–7272
8. Sauer S, Freiwald A, Maier T, Kube M, Reinhardt R, Kostrzewa M et al (2008) Classification and identification of bacteria by mass spectrometry and computational analysis. PLoS One 3:e2843
9. Dieckmann R, Helmuth R, Erhard M, Malorny B (2008) Rapid classification and identification of salmonellae at the species and subspecies levels by whole-cell matrix-assisted laser desorption ionization-time of flight mass spectrometry. Appl Environ Microbiol 74:7767–7778
10. Lay JO Jr, Holland RD (2000) Rapid identification of bacteria based on spectral patterns using MALDI-TOF MS. Methods Mol Biol 146:461–487
11. Seng P, Drancourt M, Gouriet F, La Scola B, Fournier PE, Rolain JM et al (2009) Ongoing revolution in bacteriology: routine identification of bacteria by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. Clin Infect Dis 49:543–551
12. Karger A, Bettin B, Lenk M, Mettenleiter TC (2010) Rapid characterisation of cell cultures by matrix-assisted laser desorption/ionisation mass spectrometric typing. J Virol Methods 164:116–121
13. Ouedraogo R, Flaudrops C, Ben Amara A, Capo C, Raoult D, Mege JL (2010) Global analysis of circulating immune cells by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. PLoS One 5:e13691
14. Benoit M, Desnues B, Mege JL (2008) Macrophage polarization in bacterial infections. J Immunol 181:3733–3739
15. Ouedraogo R, Daumas A, Ghigo E, Capo C, Mege JL, Textoris J (2012) Whole-cell MALDI-TOF MS: a new tool to assess the multifaceted activation of macrophages. J Proteomics 75:5523–5532
16. Delaby A, Gorvel L, Espinosa L, Lépolard C, Raoult D, Ghigo E et al (2012) Defective monocyte dynamics in Q fever granuloma deficiency. J Infect Dis 205:1086–1094
17. Mba Medie F, Ben Salah I, Henrissat B, Raoult D, Drancourt M (2011) *Mycobacterium tuberculosis* complex mycobacteria as amoeba-resistant organisms. PLoS One 6:e20499
18. Meghari S, Bechah Y, Capo C, Lepidi H, Raoult D, Murray PJ et al (2008) Persistent *Coxiella burnetii* infection in mice overexpressing IL-10: an efficient model for chronic Q fever pathogenesis. PLoS Pathog 4:e23
19. Tantibhedhyangkul W, Prachason T, Waywa D, El Filali A, Ghigo E, Thongnoppakhun W et al (2011) *Orientia tsutsugamushi* stimulates an original gene expression program in monocytes: relationship with gene expression in patients with scrub typhus. PLoS Negl Trop Dis 5:e1028
20. Ryan CG, Clayton E, Griffin WL, Sie SH, Cousens DR (1988) SNIP, a statistics-sensitive background treatment for the quantitative analysis of PIXE spectra in geoscience applications. Nucl Inst Meth Phys Res B 34:396–402

Chapter 13

Cell-Based Arrays for the Identification of Interacting Polypeptide Domains or Epitopes

Richard H. Maier, Christina J. Maier, and Kamil Önder

Abstract

The specific regions on proteins which are responsible for protein–protein interaction are called interacting domains, or epitopes in case of antigen–antibody binding. These domains are one feature to characterize proteins and are important in clinical diagnostics and research. For the mapping of such domains the use of protein/peptide arrays has become popular. Regardless of which kind of array, the major requirements are a high number of candidates arranged in the array, high quality, ease of use, and cost-effectiveness. Here, the authors describe a general protocol for mapping the interacting domains of proteins demonstrated by a high affinity protein interaction, the interaction of an antibody to an antigen. The chapter describes a stepwise protocol from library production to the verification of the domain by the use of an automated cell-based polypeptide array, which comprises the named requirements of a good array.

Key words Protein chip, Peptide array, Peptide library, Epitope mapping, Domain mapping

1 Introduction

The function of proteins is often mediated through their interactions with other proteins. Protein–Protein Interactions (PPIs) determine the outcome of most cellular processes [1] and the characterization of such identified PPIs would help to understand cellular mechanisms and protein networks. One type of characterization would be the identification of the interacting domains. Also, in the field of drug development, the identification of specific regions on virulence factors responsible for host invasion, antibiotic resistance, or immunosuppression would help to overcome the challenge against pathogens. In the area of molecular immunology, the characterization of antibody-antigen interactions, another PPI, has become an important field. The power of antibodies to interact with high affinity to specific parts of antigens (proteins or peptides), the epitope, has made antibodies interesting subjects in basic and clinical research [2]. To identify such PPIs,

Kelly M. Fulton and Susan M. Twine (eds.), *Immunoproteomics: Methods and Protocols*, Methods in Molecular Biology, vol. 1061,
DOI 10.1007/978-1-62703-589-7_13,

protein–peptide, but also peptide–peptide interactions in a high-throughput manner, the use of protein chips or arrays has become popular.

In the last 10 years many protein chips have been developed [3] and they can generally be classified in cell-free or cell-based arrays. For simplicity we use the term *protein array* to refer to a large number of different polypeptides displayed on a surface. In a cell-free protein array the proteins or peptides first have to be translated by *in vitro* systems or have to be manufactured synthetically. Later, the prepared samples have to be spotted onto suitable surfaces like glass slides or membranes. Because of the need for pure or purified proteins these kinds of protein arrays become costly and are time consuming when large numbers of peptides have to be screened or when large numbers of the same array are required [3–8]. Regarding cell-based arrays the proteins or peptides are translated in living cells on suitable surfaces [2]. In the case of human cell arrays, the array is often limited to a few hundred probes [9, 10]. Furthermore, in typical cell-based arrays the amount of candidates are limited through the minimum space on the surface. In general, the design of a protein array, cell-free or cell-based, becomes time consuming and expensive if large numbers of candidates have to be examined.

Here, we present a protocol for screening polypeptides against other proteins or peptides. This technique uses the simplicity of a robust bacterial expression system. This system is easy to use and provides no posttranslational modifications. The posttranslational systems of eukaryotic systems, like yeast, could complicate the efforts to map interacting domains of human proteins. The created array can examine up to 50,000 candidates onto one nitrocellulose membrane (22.2 cm × 22.2 cm) which has high affinity for protein binding. This high throughput array was demonstrated by Maier et al. in 2008 [2] by mapping the epitope of a specific monoclonal antibody. An epitope–antigen interaction is also an PPI and was used for the verification of the system, but apart from technical issues the mapping of epitopes for antibodies is important for clinical diagnostics and research. For example, 3 % of the world's population suffers from autoimmune diseases [11]. The method we describe here is applicable to profiling the antibodies in sera of patients with autoimmune diseases. Its application to other diseases, such as cancer or infectious diseases is also possible [2]. Beyond antibodies, in general, the described protein array is applicable for the detection of specific binding regions between two proteins. Last, because each found epitope/domain is expressed in a bacterial expression vector and fused with an affinity tag, the fusion protein can be used for affinity purification. Here, we describe the method stepwise from the creation of the needed libraries to the verification steps and demonstrate it by the mapping of a specific epitope to a monoclonal antibody.

2 Materials

All the named vendors and commercial kits are only examples and can be replaced by several other equal products.

2.1 Amplification and Sonication of the Antigen

1. Gene-specific forward and reverse primers. For the human vitamin D receptor (VDR; NCBI accession nr. NM_000376) the sequences are 5′-ATGGAGGCAATGGCGGCCAGCA-3′ (forward) and 5′-TCAGGAGATCTCATTGCCAAACA-3′ (reverse).
2. Template DNA containing the full-length coding sequence of your protein/gene of interest.
3. Taq polymerase such as BioTherm™ Polymerase (GenXpress; 5 U/μL) with the supplied 10× reaction buffer.
4. 100 mM dNTP set (Invitrogen) consists of dATP, dCTP, dTTP, and dGTP. Prepare a working solution with the concentration of 2 mM (each nucleotide) in dH_2O.
5. Thermal cycler, for example the Applied Biosystems 2720 was used in this protocol.
6. A horizontal gel electrophoresis system with a power supply for the use of agarose gels and DNA separation, e.g., the PerfectBlue Gelsystem Mini M used with the peqPOWER E300 Power Supply-230VAC (both Peqlab).
7. Agarose for gel electrophoresis e.g., SeaKem LE Agarose (Biozym).
8. TAE buffer pH 8: 40 mM Tris, 10 mM Sodium acetate, 1 mM EDTA.
9. Midori Green Advance (Biozym).
10. UV-transilluminator.
11. Scalpel.
12. 100 bp DNA ladder, such as GeneRuler™ DNA Ladder Mix SM0332 (Fermentas).
13. DNA blunt-end/repair kit, e.g., End-It™ DNA-Repair Kit (Epicentre® Biotechnologies).
14. Spectrophotometer.
15. Sonicator, such as Bandelin Sonopuls GM 70 (Bandelin).

2.2 Cell Culture, DNA and Plasmid Purification

1. Liquid Luria broth (LB); 5 g/L Yeast extract, 10 g/L Tryptone, and 10 g/L Sodium chloride containing 100 μg/mL Ampicillin.
2. LB plates; 5 g/L Yeast extract, 10 g/L Tryptone, 10 g/L Sodium chloride and 15 g/L Agar containing 100 μg/mL Ampicillin.

3. Low Salt Luria Broth plates and liquid; like normal liquid Luria-Broth medium but with 5 g/L sodium chloride with 50 µg/mL of the antibiotic Zeocin™ (Invitrogen) for pDONR/Zeo/ScaI plasmid selection.
4. Incubator with temperature control and shaking platform.
5. GenElute™ HP Plasmid Maxiprep/Miniprep Kit (Sigma-Aldrich).
6. DNA purification kit; for example Wizard® SV Gel and PCR Clean-Up System from Promega.
7. Spectrophotometer.
8. Vented Q Tray plates with cover (240 mm × 240 mm × 20 mm).
9. 384-Well microplates with cover suitable for your colony picking device (*see* Subheading 2.5).
10. Isopropyl-β-d-thiogalactopyranoside (IPTG). Prepare a working solution with the concentration of 100 mM in dH_2O.
11. Sterile flasks.
12. Microcentrifuge for 1.5 mL tubes and a centrifuge with rotor capable of accommodating tubes for the higher volumes. 15 and 50 mL tubes are needed.

2.3 Cloning by Restriction Endonucleases and Recombination

1. 8–12 U/µL Restriction Endonuclease ScaI with the supplied reaction buffer and the supplied acetylated bovine serum albumin.
2. pENTRY/Zeo/ScaI vector [16].
3. 1–3 U/µL T4 DNA Ligase with the supplied reaction buffer.
4. Gateway® LR Clonase II Enzyme Mix (Invitrogen).
5. pDest™15 vector (Invitrogen).
6. TE Buffer: 10 mM Tris, 1 mM EDTA, pH 8.0.
7. 2 µg/µL Proteinase K solution. The supplied Proteinase K solution (supplied with the Clonase) can be used, or prepare the solution by your own in TE buffer.

2.4 Transformation of Escherichia coli (E. coli) Competent Cells

1. One Shot® BL21 (DE3) chemically competent cells (Invitrogen).
2. One Shot® TOP10 Electrocomp™ cells (Invitrogen) as super competent cells for library transformation.
3. Water bath with temperature control.
4. Electroporator, e.g., the MicroPulser Electroporation System (Bio-Rad).
5. Electroporation GenPulser/MicroPulser Cuvettes 0.1 cm gap (Bio-Rad).

6. SOC Medium: 2 % Tryptone, 0.5 % Yeast extract, 10 mM Sodium chloride, 2.5 mM Potassium chloride, 10 mM Magnesium sulfate, and 20 mM Glucose.
7. Thermomixer—Shaking Heat-Block with temperature control.

2.5 Creation of the Peptide Array

1. Picking and spotting robot/system, such as the QPix2XT system (Genetix).
2. Nitrocellulose membrane (size 23 cm × 23 cm).
3. Methanol.
4. Whatman paper (Blotting paper)
5. QTray plates, media and microtiter plates (as mentioned in Subheading 2.2).

2.6 Array Detection

1. Phosphate-buffered saline (PBS); 8 g/L Sodium chloride, 0.2 g/L Potassium chloride, 1.44 g/L Sodium phosphate dibasic, 0.2 g/L Potassium phosphate monobasic. Prepare the buffer in dH_2O and adjust the pH to 7.2–7.4.
2. Blocking buffer: PBS containing 3 % bovine serum albumin (BSA) and 0.02 % Tween-20.
3. Wash buffer: PBS containing 0.02 % Tween-20 (PBST).
4. Detection and labelled antibodies/molecules. Here, the primary antibody was a commercially available mouse monoclonal IgG_{2A} anti-human VDR antibody (sc-13133; Santa Cruz Biotechnology, Santa Cruz, CA, USA). For detection, an Alexa Fluor 488-conjugated goat anti-mouse IgG antibody (Invitrogen) was used.
5. Fluorescence scanner, e.g., the Amersham Typhoon Scanner.
6. Vector specific forward primer. In case of pDest™15 use a primer with following sequence: 5′-GGCAAGCCACGTTTGGTG-3′.
7. Quantification Software (e.g., ImageQuant).
8. DNA Analyzer: For example the 3730xL from Applied Biosystems.

2.7 Protein Production and Enzyme-Linked Immunosorbent Assay (ELISA)

1. MagneGST™ Protein Purification System (Promega) with the provided MagneGST™ cell lysis reagent.
2. Quick Start™ Bradford Protein Assay (Bio-Rad).
3. 96-Well microtiter plate (e.g., NUNC F Polysorb™).
4. Coating buffer, e.g., ELISA/ELISPOT coating buffer from eBioscience.
5. Blocking/Wash buffer (*see* Subheading 2.6).

6. Antibodies. Here, the primary antibodies were a commercially available mouse monoclonal IgG_{2A} anti-human VDR antibody (sc-13133), a mouse IgG_{2A} monoclonal anti-RACK1 (receptor for activated C kinase 1) antibody (sc-17754; Santa-Cruz Biotechnology), and normal mouse-IgG (sc-2025; Santa-Cruz Biotechnology). For detection, a horseradish peroxidase (HRP)-conjugated goat polyclonal to mouse IgG (Abcam).
7. SIGMAFAST™ OPD (*o*-phenylenediamine dihydrochloride) tablets (Sigma-Aldrich), dissolved in dH_2O as substrate for the HRP.
8. Microplate reader, such as the Paradigm™ detection platform (Life Technologies).

3 Methods

An overview of all steps required for this method is given in Fig. 1.

3.1 Creation of a Comprehensive Antigen/Domain Presenting Polypeptide Library

To find the interacting region of a PPI one of the two partners has to be presented on the array in many short fragments covering all possible binding regions. The other partner should be available in a form, in which it can easily be detected, e.g., biotinylated. For the later analysis it is easier to cleave the DNA sequence of the protein in many short fragments and clone it into a vector for bacterial (*E. coli*) recombinant protein expression. The power of a protein array depends also on the number of candidates that can be arrayed. For that purpose the use of highly competent *E. coli* cells is necessary to provide enough material to exhaust the possibility to screen up to 50,000 candidates per one membrane. As example, this protocol describes the generation of an antigen presenting library. The antigen is the human vitamin D receptor (VDR) and a commercially available antibody (with known epitope) was used to map the epitope. For library construction the antigen was PCR amplified, cloned into an *E. coli* expression vector and transformed into a protein producing *E. coli* strain. Cloning by recombination [12, 13], or Gateway® cloning, was used here because of the known benefits of this technique. The reactions are highly efficient, quick, simple, and are especially useful for high-throughput projects involving thousands of candidates [14]. To obtain Gateway® flexibility the antigen representing DNA fragments have first to be cloned into a Gateway compatible DONR™ vector (creation of an Entry library).

3.1.1 Fragmentation of the Antigen

1. The DNA sequence of the antigen (or the protein of interest) has to be provided as template DNA for PCR amplification, e.g., cDNA of the gene, a vector containing the coding sequence of the gene, or, in case of bacterial proteins, genomic DNA.

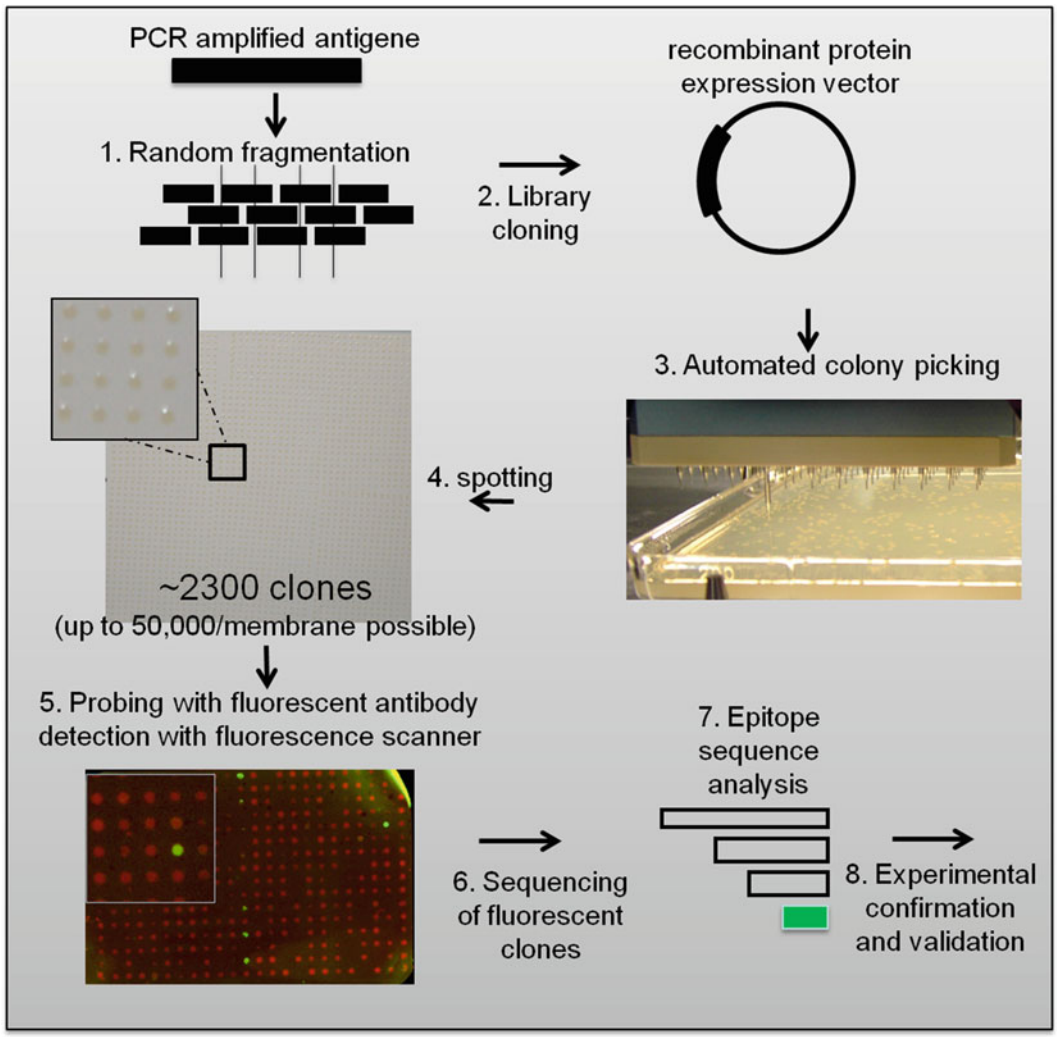


Fig. 1 Schematic of the epitope mapping procedure. The antigen-encoding gene is PCR-amplified and fragmented by sonication (*1*). The ends are repaired and cloned by ligation followed by recombination into an *E. coli* GST-fusion expression vector (*2*). An *E.coli* expression strain is transformed with the library and colonies are picked by a QPix2XT (Genetix) robot (*3*). After cultivation the cultures are spotted by a robot onto a 22.2 cm × 22.2 cm nitrocellulose membrane (*4*). A standard dot-blot protocol is used for fluorescence detection under a Typhoon scanner (*5*). Cells harboring a positive, here *bright spot*, are used for plasmid preparation and sequencing (6). The derived sequences are aligned to the antigen for identification of the epitope (*7*). The minimal overlapping sequence is separately cloned and subjected to epitope validation (*8*). (Reprinted from ref. 2. Copyright notice/credit © Journal of Biomolecular Screening/SAGE Publications 2010 used by Permission)

2. Design two gene specific forward and reverse primers. Take the first 20–30 nucleotides of the 5′ end of the coding and complementary strand of the gene of interest, and choose the length of the primers so that they have a similar annealing temperature, between 50 and 60 °C.

3. PCR amplify the antigen's sequence in a 50 µL assay consisting of 2.5 U BioTherm™ Polymerase, gene specific forward primer and reverse primer (0.3 µM each), dNTP mix (0.2 mM each), 10× BioTherm™ reaction buffer (5 µL), 10 ng of template DNA, and dH_2O up to 50 µL. Make 8 equal PCR samples (50 µL each). The 40 PCR cycles (94 °C for 30 s, 50 °C for 30 s, and 72 °C for 1 min/kb) were preceded by heating to 94 °C for 5 min and followed by a 7 min incubation at 72 °C in a standard thermal cycler (*see* **Note 1**).
4. Determine the correct size of the amplified material by agarose gel electrophoresis and DNA staining, e.g., ethidium-bromide or midori green. For midori green staining prepare an agarose solution (0.8–3 %) in TAE buffer, boil it in a microwave until the solution is clear, and cool it down (~60 °C). Use 5 µL Midori Green Advance for a 100 mL Agarose solution, and pour the gel into the chamber. Place the comb into the gel and wait until the gel becomes jellied. Load the sample into the slots of the gel and separate the samples using 10 V/cm for 45 min. You can visualize the bands under UV light.
5. Pool all eight PCR samples, and mix the resulting 400 µL with 1,600 µL sterile dH_2O in a 15 mL centrifuge tube.
6. Sonicate the 2 mL sample (*see* Fig. 1, **step 1**). For that, put the 15 mL tube in an ice water bath, set the sonicator-probe into the 2 mL sample, and sonicate the DNA for 20 min. The sonication settings are as follows: constant sonication, cycle 30 with ~90 % output. During sonication the sample becomes hot and you will maybe have some loss of material. Furthermore, please ensure that the sonicator-probe does not touch the bottom of the 15 mL tube (*see* **Note 2**).
7. Load the complete 2 mL onto one purification column (if the column's capacity is less than 2 mL load the sample twice) and purify the sample using a DNA purification kit following the manufacturer's instructions. Elute the DNA with 50 µL sterile dH_2O.
8. Load the complete 50 µL together with a 100 bp DNA ladder onto a 2 % agarose gel, separate the gel for a 45 min with 10 V/cm and visualize the DNA by midori green staining or equivalent using a UV-transilluminator. If the sonication procedure was successful the fragmented DNA will appear as a smear with fragments going down to 50 bp.
9. Use again the same DNA purification kit as already mentioned, excise the gel region where the fragments between 50 and 300 bp appear with a scalpel, and use this for DNA purification following the manufacturer's instructions. Elute the DNA with 50 µL sterile dH_2O.
10. Sonication of DNA will result in fragments with all kind of possible overhangs. For cloning, the best strategy is to blunt-end

all fragments. For that, use the complete purified fragmented DNA between 50 and 300 bp in a DNA blunt-end/repair kit by using the kit's instructions. The resulting sample should again be purified using the DNA purification system and the manufacturer's instructions. Elute the DNA with 50 µL sterile dH_2O and determine the concentration of the purified linearized and blunt-ended DNA by absorbance measurement at 260 nm. The typical concentration is between 50 and 100 ng/µL. The concentration may vary from each experiment. In total you need 200 ng of fragmented antigen DNA. In cases of low concentrated DNA, repeat this section and use a smaller volume for the elution steps, e.g., 25 µL.

3.1.2 Creation of the Entry Library

To use Gateway® terminology, the term DONR™ vector means a vector ready for a so-called attB × attP (BP) reaction. A cloning step where the to be cloned gene, fragment, etc. is cloned into this vector by replacing the existing ccdB [15] gene (negative selection) and the information for chloramphenicol resistance. For more information please read the Gateway® cloning instruction manual (Invitrogen). For terminology it is important that the vector is called Entry vector when recombination has occurred successfully. Here we describe the recombination of the fragmented antigen into an in house created Entry vector harboring a blunt-end restriction site for ScaI (pENTRY/Zeo/ScaI). For the details of the production of this vector please *see* Maier et al. [14, 16]. With this vector the flexibility of the Gateway® system was obtained but it is also possible to use standard restriction enzyme based expression vectors for the construction of the array.

1. Place 2 µg pENTRY/Zeo/ScaI vector in a 20 µL sample together with 10 units ScaI, 2 µL 10× reaction buffer, 2 µg acetylated bovine serum albumin and up to 20 µL with dH_2O. Mix the reaction sample well and centrifuge the sample briefly. Let the sample incubate at 37 °C for 2 h (*see* **Note 3**).
2. To stop the reaction purify the complete sample with a DNA purification kit following the kit's instructions. Elute with 20 µL sterile dH_2O and determine the concentration of the DNA by absorbance measurement at 260 nm.
3. Set up a ligation reaction with 200 ng repaired and fragmented antigen DNA, 200 ng linearized blunt-ended vector DNA, 1.5 µL 10× reaction buffer, 1–3 units T4 ligase and up to 15 µL with dH_2O. Mix the reaction and spin the sample briefly. Let the sample incubate at 4 °C for at least 3 days (*see* **Note 4**).
4. Stop the reaction again by purification with a DNA purification kit following the kits instructions. Elute with 20 µL sterile dH_2O.
5. Transform 5 µL of the purified ligation reaction into super competent *E. coli* TOP10 cells. The following transforming

instructions assume the use of electrocompetent cells. Thaw the cells on ice and pipet the 5 µL into 50 µL of the competent cells. Do not pipet up and down or vortex. Take the complete ~55 µL and pipet it between the two electrodes of the electroporation cuvette. Place the cuvette in the MicroPulser, select the bacteria setting and the Eco1 program and pulse. Rinse the cells out of the cuvette with 1 mL SOC by pipetting the media 3–5 times up and down in the cuvette and put them finally in a sterile tube for incubation at 37 °C for 1 h with gentle shaking.

6. To test for transformation efficiency plate 50 µL of the reaction (1:1,000 diluted) onto LB plates containing the appropriate antibiotic (in case of pENTRY/Zeo/ScaI use 50 µg/mL Zeocin™). Incubate the plate overnight at 37 °C. Use the complete rest of the reaction (undiluted) to inoculate 150 mL liquid LB medium with the respective antibiotic. Incubate the culture overnight at 37 °C with gentle shaking (*see* **Note 5**).
7. Determine the next day the amount of primary colonies by counting the colonies on the plate. A good library should have at least 10^6 primary clones (*see* **Note 6**).
8. Use the grown culture to purify the plasmid DNA using a plasmid purification kit following the manufacturer's instructions. Determine the concentration of the plasmid DNA by absorbance measurement at 260 nm.

3.1.3 Creation of the Expression Library

To use again Gateway® terminology the term "Destination" vector refers to a vector with attR attachment sites suitable for LR reactions. After successful recombination of an attL site flanked piece in an Entry vector in a LR reaction the Destination vector becomes an Expression vector with again attB-sites. The expression library represents the fragmented antigen fused to an affinity or purification tag (in this case with a N-terminal GST tag) which can be expressed by a specific *E. coli* strain (*see* Fig. 1, **step 2**). If you have used a standard restriction enzyme based cloning vector for fragment expression please proceed to **step 4** of this section.

1. Insert 200 ng of the created Entry library in pENTRY/Zeo/ScaI in a 10 µL assay with 200 ng pDest™15 vector, 2 µL LR Clonase™ II Enzyme Mix, and TE buffer up to 10 µL and let the sample incubate at 25 °C or room temperature over night. Stop the reactions the next day by adding 1 µL proteinase K and incubate at 37 °C for 30 min (*see* **Note 7**).
2. Transform 5 µL of the stopped LR reaction into super competent *E. coli* TOP10 cells as described in Subheading 3.1.1, **steps 5** and **6**. The appropriate antibiotic for pDest™15 selection is ampicillin (100 µg/mL final concentration).

3. Determine the amount of primary colonies, purify the plasmid DNA and measure the concentration of the purified DNA like it is described in Subheading 3.1.1, **steps 7** and **8**.
4. Transform 10–50 ng of the purified expression library into chemically competent *E. coli* BL21 (DE3) cells. For that, thaw the competent cells on ice and pipet the DNA directly into the cells. Mix the sample briefly and incubate it on ice for 20 min. Heat shock the cells for 30 s in a 42 °C water bath and cool it down on ice for 1 min. Add 1 mL of SOC media and incubate the sample at 37 °C for 1 h. After incubation plate 300 and 700 µL onto two big Q Tray plates containing solid LB medium with the respective antibiotic. Incubate the plates overnight at 37 °C (*see* **Note 8**).

3.2 Cell-Based Polypeptide Array

The protein array is structured out of a set of single colonies, each expressing a different polypeptide, picked and arrayed automatically in 384-well plates, printed and grown onto a nitrocellulose membrane, and analyzed with fluorescence labelled antibodies. In the described approach here 2,304 colonies are arranged in four 384-well plates and automatically spotted onto a 22.2 cm × 22.2 cm nitrocellulose membrane that is specific for protein and has high affinity for protein binding, giving a density of about 4.7 spots per cm^2. The array can be up-scaled up to 50,000 colonies on the same area to perform large scale studies.

3.2.1 Colony Picking, Spotting and Protein Production

1. Start with the transformation Q Tray plates from **step 4** in Subheading 3.1.3 and prepare some 384-well plates, each well prefilled with 50 µL liquid LB medium containing 100 µg/mL ampicillin. The number of 384-well plates depends on the amount of colonies which have to be screened and arranged in the array. For example, for 2,304 colonies four 384-well plates are needed.
2. Place the 384-well plates in the stacker of the picking robot, place the grown Q Tray plates in the desired place for colony picking and start the manufacturers picking software. For detailed instructions to use your robotic system please refer to the manufacturer's instruction manuals. Spotting the previously mentioned ~50,000 colonies onto one membrane is possible with the Genetix QPix2XT system (*see* **Note 9**).
3. Let the robot pick and inoculate the desired amount of colonies in the prepared 384-well plates (*see* Fig. 1, **step 3**). Incubate these plates overnight at 37 °C. As positive control, you can inoculate a desired well with a colony harboring the full-length molecule of your antigen/protein in the same vector/cells. As negative control the *E. coli* BL21 (DE3) strain with a just GST tag producing vector is possible (*see* **Note 10**).

4. The next day, prepare a ~23 cm × 23 cm piece of a nitrocellulose membrane and put it in pure methanol for 10 min and subsequently place for 1 min into liquid LB medium containing 100 μg/mL ampicillin.
5. Moisten a piece of Whatman paper (same size as the membrane) with liquid LB medium containing 100 μg/mL ampicillin.
6. For spotting place the grown 384-well plates into the stacker of the spotting system (for example we use the QPix2XT system here). Assemble in the spotting area of your system first the wet Whatman paper and onto it the equilibrated membrane. Pay attention to avoid air bubbles between the Whatman paper and the membrane. Set the pin height of the spotting head to an optimum so that they can touch the membrane but never force through the membrane. Start the spotting/gridding software of your device and let the robot spot your samples onto the membrane (*see* Fig. 1, **step 4**). For later documentation please note which plate is spotted in which direction onto the membrane, or use the documentation software of your device. Keep the 384-well plates at 4 °C.
7. Transfer the membrane carefully onto a Q Tray plate containing solid LB medium containing 100 μg/mL ampicillin. The spotted side should be on top and no air bubbles should be under the membrane to allow growth of every spot. Incubate the plate with the membrane overnight (but not more than 16 h) at 28–30 °C (*see* **Note 11**).
8. The next day visible colonies should be grown on the membrane. Transfer the membrane carefully onto a fresh Q Tray plate containing solid LB medium containing 100 μg/mL ampicillin and 0.5 mM IPTG, and induce protein expression by incubating the plate for 2 h at 37 °C.

3.2.2 Domain/Epitope Screening

For the detection of the possible epitopes to the specific antibody a refined dot-blot procedure was used, which was scaled up to accommodate a 22.2 cm × 22.2 cm large membrane.

1. The cells on the membrane from **step 8** in Subheading 3.2.1 were lysed by transferring the membrane into blocking buffer containing Tween-20 for 1 h with gentle shaking at room temperature. This little amount of detergent is enough to lyse the cells and the released recombinant polypeptides stay bound to the nitrocellulose (*see* **Note 12**).
2. Incubate the membrane in blocking buffer containing the specific antibody to the "unknown" epitope. In this approach we used the commercial mouse monoclonal IgG_{2A} anti-human VDR antibody (sc-13133; Santa Cruz Biotechnology, Santa Cruz, CA, USA) diluted 1:2,000 for 1 h at room temperature (*see* **Note 13**).

3. Pour off the antibody solution and wash the membrane two times in washing buffer. Each time for 5 min with gentle shaking at room temperature.
4. To detect the first antibody, incubate the membrane in blocking buffer containing an 1:1,000 diluted Alexa Fluor 488-conjugated goat anti-mouse IgG antibody at room temperature for 1 h (*see* **Note 14**).
5. Pour off the antibody solution and wash the membrane three times in washing buffer. Each time for 5 min with gentle shaking at room temperature.
6. Detect positive array spots with a scanner, such as the Amersham Typhoon Scanner (*see* Fig. 1, **step 5**). Alternatively you can use any kind of fluorescence scanner which is able to scan surfaces and your fluorescence label. Excitation occurred at 532 nm, and emission was measured with a 526 nm filter (*see* **Note 15**).
7. Only spots with a clear fluorescence signal were considered as hits and subjected to DNA sequencing (*see* Fig. 1, **step 6**). For that, determine the positions of the original positive spot producing culture in the 384-well plates used for spotting. Inoculate 1 μL of the culture into fresh liquid LB medium containing 100 μg/mL ampicillin. Incubate the culture at 37 °C with gentle shaking.
8. Use the grown culture to purify the plasmid DNA using a plasmid purification kit following the kit's instructions and then sequence the vector with a vector specific forward primer. For DNA sequencing many provider are available, or use your own device, e.g., the 3730xL DNA Analyzer and follow the protocols provided by the manufacturer.
9. Analyze the resulting DNA sequences. Convert the nucleotide sequences into amino acid sequences and look for an overlap (*see* Fig. 1, **step 7**). The shortest found overlap of the translated peptides represents the possible interacting domain or here, the epitope (*see* **Note 16**).
10. For validation of the epitope, clone the minimum overlapping sequence separately by your cloning strategy into the used expression vector. For a detailed description of how to use Gateway® cloning refer to the Gateway® cloning instruction manual.

3.3 Epitope/Domain Validation

The easiest way for validation (*see* Fig. 1, **step 8**) is to produce the minimum found domain/epitope in the used expression system and detect it with its interaction partner. For a detailed protocol of how to induce expression in the bacterial BL21 (DE3) cells, please see the well explained protocols by the manufacturer or other open source protocols. Purified peptides are better for usage in the validation protocols, but also crude lysates of the peptide producing

cells are possible. It depends on the validation method. For simple validation by gel separation and standard western blot analysis, a crude lysate can be used. For that, use the same conditions which are described in Subheading 3.2.2 and confirm the size and binding of the antibody to the epitope. This section will describe the validation by an Enzyme-linked Immunosorbent Assay (ELISA) with purified peptides/proteins.

1. Clone the minimal overlapping sequence and the full-length DNA sequence of the protein (positive control) in the desired expression vector (e.g., pDEST™15). The expression vector should harbor for later purification an n-terminal purification tag (in case of pDEST™15 a GST tag). An in-house pDEST™15 vector harboring a stop codon was used as negative control [2]. Purify all plasmids using a plasmid purification kit following the manufacturer's instructions. Determine the concentration of the plasmid DNA by absorbance measurement at 260 nm (*see* **Note 17**).
2. Transform 50 ng of each plasmid into *E. coli* BL21 (DE3) chemically competent cells (*see* Subheading 3.1.3, **step 4**) and inoculate 100 µL of each transformation into 7 mL of fresh liquid LB medium containing 100 µg/mL ampicillin. Incubate the samples overnight at 37 °C with gentle shaking.
3. Add the next day IPTG to the grown cultures to reach a final IPTG concentration of 1 mM and let the cultures further incubate for 2 h at 37 °C with shaking.
4. Pellet the cells at 3,000–4,000 × *g* for 20 min. Discard the supernatant and freeze the pellets once for 15 min at –80 °C.
5. Thaw the cell pellet at room temperature and lyse the pellets by resuspending in 400 µL MagneGST™ cell lysis reagent. And incubate the lysates for 30 min at room temperature with gentle shaking (*see* **Note 18**).
6. Use this lysates to purify the recombinant peptides/proteins according to the MagneGST™ Protein Purification System kits instructions. Elute the samples with 50 µL Elution buffer twice (=100 µL). Determine the protein concentration with the Quick Start™ Bradford Protein Assay kit.
7. Use a 96-well plate which is adaptable for protein binding and absorbance measurement (e.g., NUNC F PolySorb™) and coat with 1 µg/well of purified peptide and controls in 200 µL coating buffer overnight at 4 °C. To fulfill statistic requirements coat altogether 12 wells of each peptide and control (three technical replicates for every tested antibody or control).
8. Wash each well three times with 200 µL of dH_2O.

9. Block nonspecific binding by adding 200 µL/well of 3 % BSA in PBST. Incubate the plate at room temperature for 1 h.
10. Discard the blocking buffer and incubate each well with primary antibody at a concentration of 0.4 ng/µL in PBST containing 3 % BSA at room temperature for 3 h. In case of domain mapping use the binding partner instead of the primary antibody and use further suitable detection antibodies. For human vitamin d receptor (VDR) epitope mapping we used the epitope-mapped mouse IgG2A monoclonal anti-VDR antibody, a mouse IgG2A monoclonal anti-RAC K1 (receptor for activated C kinase 1) antibody, and normal mouse-IgG. Each antibody was tested in triplicate against the epitope and the controls. As a further control, just 3 % BSA in PBST was applied to wells to detect possible binding of the secondary antibody.
11. Wash each well thrice with 200 µL PBST.
12. Incubate each well with 100 µL of horseradish peroxidase (HRP)-conjugated goat polyclonal to mouse IgG at a concentration of 2 ng/µL diluted with 3 % BSA in PBST for 1 h at room temperature in the dark.
13. Wash each well five times with 200 µL PBST.
14. Incubate each well with 200 µL of dissolved SIGMAFAST™ OPD tablets for 60 min in the dark. For the detection of background activity incubate also five empty wells with the substrate.
15. Measure the absorbance at 450 nm in a plate reader like the Paradigm™ detection platform from Life Technologies. The absorbance (optical density, OD) of each well was determined by subtracting the mean of the 5 background test wells which contains just 200 µL of substrate.
16. The wells coated with the found epitope/domain and the positive control should deliver high absorbance values (OD between 0.6 and 0.8 or higher) in contrast to the negative control where no binding should result in low (below OD 0.15) OD values (*see* Fig. 2). This then reveals the epitope of interest.

4 Notes

1. The use of a standard taq polymerase with 40 cycles instead of a proofreading polymerase with low cycles is recommended because much material is required for sonication. In many cases a proofreading polymerase delivered less material (also with 40 cycles) than a standard taq polymerase (own experience). The possibility that mutations will be incorporated into

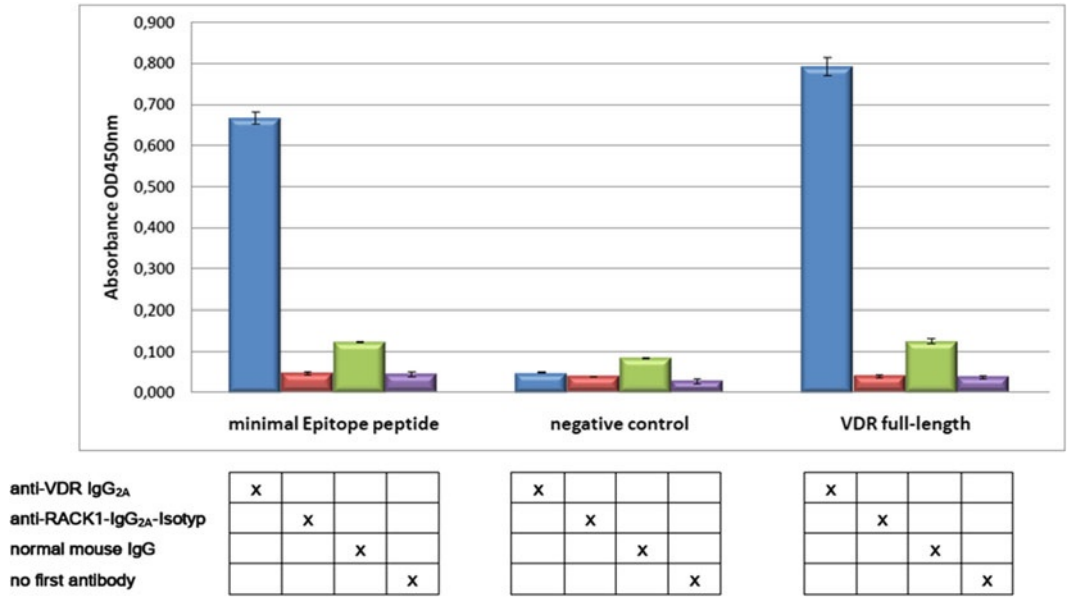


Fig. 2 Epitope validation. ELISA to verify binding of the monoclonal anti-VDR antibody only to the epitope. The minimal found epitope, full-length VDR (both fused with a GST-tag) and the negative control (GST-tag only) were coated onto a microtiter plate. The specificity of the VDR IgG antibody was shown in competition with other isotype-matched control IgG antibodies (normal mouse-IgG and anti-RACK1 IgG). As a control for the secondary anti-mouse IgG, one series was incubated with blocking buffer only instead of a primary antibody to exclude possible binding of the secondary antibody. Only with the anti-VDR antibody was significant absorbance detected in combination with the minimum epitope (37 amino acids) or the full-length VDR. Error bars represent the standard error of $n=3$ values. (Reprinted from ref. 2. Copyright notice/credit © Journal of Biomolecular Screening/SAGE Publications 2010 used by Permission)

DNA is insignificant because these molecules will likely not result in positive interactions.

2. These settings are optimized for the use with the named sonicator. With other devices a sonication time of 1–5 min may be enough depending of the sonication performance/power of your device. For each device the settings have to be optimized to sheer the DNA into small fragments of around 50 base pairs.
3. The use of other *E. coli* expression vectors is possible (pay attention to the different expression systems—some vectors are not inducible, etc.). If you will use standard restriction enzyme based cloning, we recommend cloning the sonicated DNA into the blunt-end site of the multiple cloning site as described in this section with the modified Entry vector, but skip the protocol for LR reaction and proceed with **step 4** in Subheading 3.1.3. After ligation into the final expression vector without Gateway cloning you do not have to make the step from the entry library into the expression vector.

4. Standard ligation reactions can be done in a few minutes to 1 h (at 16 °C, depending on your enzyme). In our own experiments we reached the best ligation efficiency with these conditions. More colonies harbor an insert (the cloning efficiency is increased) and the number of self-ligated vectors goes down.
5. When using systems with an inducible promoter, the transcription machinery is not completely switched off. Those clones with vectors harboring an insert that is toxic or regulates cell growth will not grow as fast in a liquid culture. The clones with nontoxic inserts (strong ones) can grow potentially more rapidly. On plates this discrimination is not given, because even if the some colonies grow slower than the others, they can still grow. The inoculation of liquid cultures will likely benefit clones with "strong" fragments or will eliminate those clones which have a disadvantage because of their cloned inserts. To circumvent such discrimination the transformation reaction can be plated on 5–10 larger plates (145 mm in diameter). After overnight incubation at 37 °C, use the grown cells for plasmid preparation.
6. You can prove the cloning efficiency by colony PCR or sequencing with suitable primers for the chosen vector. A good library should harbor at least 80 % of the clones with different inserts.
7. If there is a problem to get the same amount of primary colonies (compared to your Entry library) after LR reaction we suggest performing 3–5 of the same sample for LR reaction, pooling them after stopping them and purifying with a DNA purification kit. Use a smaller amount of elution buffer (~20 μL) so that your DNA is more concentrated in the final elution volume.
8. It is also possible to transform the LR reaction directly into the BL21 cells. This maybe result in just a few hundred colonies. If this amount is enough for your planned array you do not have to make the steps of propagation and purification in the other *E.coli* strain. Furthermore, the use of highly competent cells is not necessary here. Even if you use competent cells with a low competence, like 10^7 transformants per μg of DNA, you can reach the number of the desired primary candidates. But if you get fewer colonies for the array, use either more DNA for transformation or cells with a better competence. Here, we used chemically competent cells because the provider does not sell electrocompetent cells of this special *E. coli* strain. It is also possible to create electrocompetent cells of this cell type. Common protocols for this can be found on the internet. In general, you can reach a higher competence with the protocols to make electrocompetent cells (own experience).

9. If you plan to design an array with just a few hundred or less candidates this step can also be performed manually. For that, use a 96-pin head for manually use, dip the pins into the cultures (one pin into one well), and spot the pins with the adherent cultures onto the membrane. Just use little pressure. Sterilize the pins before each step with ethanol and flame treatment.
10. Due to evaporation some liquid in the outer plate positions may be lost. If there is a little remaining in the wells this will be enough for colony spotting, but please make sure that the culture is not completely desiccated. The cultures will also grow better if you incubate the plate with gentle shaking. The shaking of the culture aerates the bacteria sufficiently to allow for good growth. Although *E. coli* is a facultative anaerobic organism, it grows better under such conditions.
11. The size of the colonies is important for the discrimination of each spot. The bigger the clones grow, the less space there will be on the membrane between the colonies, so larger colony size is not preferable. If you plan an array with ~50,000 spots on this space an incubation time of 4–6 h will be enough.
12. In some cases, the cells will not detach easily from the membrane. If there are still cells after 30 min of incubation on the membrane (you will easily notice it, because you will still see the colony spots on the membrane) please shake the membrane a little to get all material off.
13. If you use an interacting protein use this protein as "primary antibody." For further detection this protein has to be equipped with a label, like Biotin, Fluorescein or similar for detection purposes. Also specific antibodies to this protein can be used. The conditions have to be tested individually.
14. If your first antibody or the interacting protein is already labelled with the fluorophore, proceed with the next step.
15. To get concrete values (Data), analyze the scanned blot with a quantification software (like ImageQuant), but with the named detection settings positive spots are clearly detectable. That means that you can clearly discriminate a positive spot from a negative one by analyzing the scanned array with your eyes. But if there is the need for quantification of the positive spots please use the named software.
16. It is also possible to get more than one overlapping region on the mapped protein. This is for example due to discontinuous epitopes which is also possible to detect.
17. The negative control should be of the same vector background expressing the n-terminal tag with a subsequent stop codon in frame with the tag. It is not possible to use an original

pDEST™15 vector as negative control because of the ccdB gene (negative selection).

18. The described method is just one example for purification. Other systems are also suitable. It is likely that all methods have to be optimized because of the usage of different proteins with different properties.

References

1. Shoemaker BA, Panchenko AR (2007) Deciphering protein–protein interactions. Part I. Experimental techniques and databases. PLoS Comput Biol 3:e42
2. Maier RH, Maier CJ, Rid R, Hintner H, Bauer JW, Onder K (2010) Epitope mapping of antibodies using a cell array-based polypeptide library. J Biomol Screen 15:418–426
3. Spisak S, Tulassay Z, Molnar B, Guttman A (2007) Protein microchips in biomedicine and biomarker discovery. Electrophoresis 28:4261–4273
4. Liu R, Barrick JE, Szostak JW, Roberts RW (2000) Optimized synthesis of RNA–protein fusions for in vitro protein selection. Methods Enzymol 318:268–293
5. Zhu H, Bilgin M, Snyder M (2003) Proteomics. Annu Rev Biochem 72:783–812
6. Ramachandran N, Hainsworth E, Bhullar B, Eisenstein S, Rosen B, Lau AY, Walter JC, LaBaer J (2004) Self-assembling protein microarrays. Science 305:86–90
7. Tao SC, Zhu H (2006) Protein chip fabrication by capture of nascent polypeptides. Nat Biotechnol 24:1253–1254
8. Beutling U, Stading K, Stradal T, Frank R (2008) Large-scale analysis of protein–protein interactions using cellulose-bound peptide arrays. Adv Biochem Eng Biotechnol 110:115–152
9. Hu YH, Warnatz HJ, Vanhecke D, Wagner F, Fiebitz A, Thamm S, Kahlem P, Lehrach H, Yaspo ML, Janitz M (2006) Cell array-based intracellular localization screening reveals novel functional features of human chromosome 21 proteins. BMC Genomics 7:155
10. Mannherz O, Mertens D, Hahn M, Lichter P (2006) Functional screening for proapoptotic genes by reverse transfection cell array technology. Genomics 87:665–672
11. Robinson WH, Steinman L, Utz PJ (2003) Protein arrays for autoantibody profiling and fine-specificity mapping. Proteomics 3:2077–2084
12. Hartley JL, Temple GF, Brasch MA (2000) DNA cloning using in vitro site-specific recombination. Genome Res 10:1788–1795
13. Walhout AJ, Temple GF, Brasch MA, Hartley JL, Lorson MA, van den Heuvel S, Vidal M (2000) GATEWAY recombinational cloning: application to the cloning of large numbers of open reading frames or ORFeomes. Methods Enzymol 328:575–592
14. Maier R, Brandner C, Hintner H, Bauer J, Onder K (2008) Construction of a reading frame-independent yeast two-hybrid vector system for site-specific recombinational cloning and protein interaction screening. Biotechniques 45:235–244
15. Bernard P, Kezdy KE, Van Melderen L, Steyaert J, Wyns L, Pato ML, Higgins PN, Couturier M (1993) The F plasmid CcdB protein induces efficient ATP-dependent DNA cleavage by gyrase. J Mol Biol 234: 534–541
16. Maier RH, Maier CJ, Onder K (2011) Construction of improved yeast two-hybrid libraries. Methods Mol Biol 729:71–84

Chapter 14

Identification and Quantitation of MHC Class II-Bound Peptides from Mouse Spleen Dendritic Cells by Immunoprecipitation and Mass Spectrometry Analysis

Leonia Bozzacco and Haiqiang Yu

Abstract

Advances in immunology and immune therapies require knowledge of antigenic peptide sequences that are presented on MHC class II and class I molecules of antigen presenting cells. The most specialized antigen presenting cells are dendritic cells (DCs). In the past, the small number of DCs that could be isolated from mouse spleen prevented direct analysis of the MHC II peptide repertoire presented by DCs. Here we describe a protocol that integrates immunological methods (*in vivo* enrichment of mouse spleen DCs by Flt3L treatment and immunoprecipitation of MHC II-peptide complexes), mass spectrometry analysis and peptide synthesis (LC-MS/MS and quantitation analysis for non tryptic peptides) to identify and quantitate the endogenous peptides that are bound to MHC II molecules on DCs. The described method produces quantitative data that are reproducible and reliable enough to cover a wide range of peptide copy numbers. We propose the application of this method in future studies to quantitatively investigate the MHC II repertoire on DCs presented during viral infections or different immunizations in vaccine development research.

Key words Endogenous peptides, MHC, DCs, Flt3L, LC-MS/MS, MASCOT search engine, Quantitation analysis, Isotope-labeled peptides

1 Introduction

Major histocompatibility complex class II (MHC II) molecules are expressed on the surface of antigen-presenting cells, such as dendritic cells (DCs) and display short bound peptide fragments derived from self- and nonself antigens. These peptide-MHC complexes function to maintain immunological tolerance in the case of self-antigens and to initiate $CD4^+$ T cell-mediated adaptive immunity in the case of foreign proteins [1, 2].

Here we report in great detail the application of mass spectrometry (MS) analysis to identify MHC II peptides derived from endogenous proteins expressed in freshly isolated murine splenic DCs [3].

Kelly M. Fulton and Susan M. Twine (eds.), *Immunoproteomics: Methods and Protocols*, Methods in Molecular Biology, vol. 1061, DOI 10.1007/978-1-62703-589-7_14,

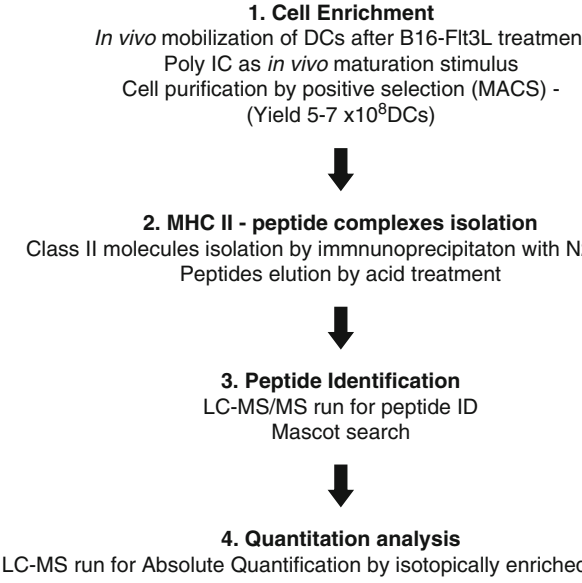


Fig. 1 A schematic diagram summarizing the steps performed to identify peptides presented by MHC II molecules on DCs, from sample preparation to peptide identification and quantitation by LC-MS/MS analysis

In the past, limited numbers of DCs *in vivo* have restricted their use for proteomic studies, but now the number of splenic DCs can be expanded *in vivo* by treatment with Flt3L [4–6], which is a regulator of hematopoietic cell development [7, 8]. In addition, rapid advances in mass spectrometry (MS) instrumentation and computational technologies have made it feasible to detect candidate peptides in a complex biological sample with high sensitivity [9]. MS-based strategies for the precise measurements of peptides have been exploited to quantitatively determine the absolute expression levels of a given protein or post-translationally modified protein [10]. These strategies may vary depending on the MS platforms, but they all rely on the introduction of isotope-labeled heavy peptides, which precisely mimics their endogenous counterparts. Synthetic standard peptides with incorporated ^{13}C, ^{14}N, or both are used as internal standards, which are introduced at a known concentration in order to quantify the corresponding endogenous counterparts [3, 9, 10]. With the method presented here, starting with about 5×10^8 splenic DCs, we were able to identify a repertoire of hundreds of MHC II peptides from endogenous proteins localized in all cell compartments [3]. Using synthetic isotope-labeled peptides, in a single experiment, we detected peptides in a wide range of concentration spanning from 2.5 fmol/μL to 12 pmol/μL [3].

This method, as outlined in Fig. 1, sets a foundation for future MS based peptide analysis to quantitatively investigate the MHC II repertoire on DCs generated under viral infections or, in vaccine development studies, during different immunization conditions.

2 Materials

We do not add sodium azide to any reagents. Carefully follow all waste disposal regulations when disposing waste materials.

2.1 Cell Culture

Prepare all solutions for cell culture with culture grade reagents and store media and other biological reagents at 4 °C.

1. Balb/c × C57BL/6 (I–$A^{b,d}$/E^{d}) F1 mice, maintained under specific pathogen-free conditions and used at 6–8 weeks of age.
2. Melanoma cells expressing Fms-like tyrosine kinase 3 ligand (Flt3L) [11] cultured with Dulbecco's modified Eagle medium (DMEM) containing 10 % heat-inactivated fetal bovine serum (FBS).
3. Anti-CD11c magnetic beads and MACS LS columns (Miltenyi Biotec).
4. Phosphate-buffered saline (PBS): 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 2 mM KH_2PO_4, pH 7.4.
5. Anti-MHC class II hybridoma cells (clone N22) [12], maintained in DMEM medium with 2 mM l-glutamine, 5 % heat-inactivated FBS.

2.2 Generation of Affinity Column

1. CNBr Sepharose beads (GE Healthcare).
2. Coupling buffer (CB): 0.5 M NaCl, 0.1 M $NaHCO_3$, pH 8.3.
3. Activating buffer: 1 mM HCl.
4. Glycine solution: 0.2 M glycine, pH 8.0.
5. Poly-prep columns (Bio-Rad).
6. Affinity-purified N22 antibody from culture supernatants using Protein G Sepharose beads as per manufacturer's instructions.

2.3 Affinity Purification of MHC II Molecules

1. Cell lysis buffer: 20 mM Tris–HCl pH 8, 1 % 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS), 5 mM ethylenediaminetetraacetic acid (EDTA), 10 μg/mL aprotinin, 10 μg/mL leupeptin, 10 μg/mL pepstatin, 0.1 mM iodoacetamide, 1 mM phenylmethanesulfonylfluoride (PMSF) (all kept at −20 °C) (*see* **Note 1**).
2. Washing buffer 1: 250 mM NaCl, 50 mM Tris–HCl pH 8.
3. Washing buffer 2: 50 mM Tris–HCl, pH 8.
4. 10 % glacial acetic acid.

2.4 Western Blot

1. Mini PROTEAN Tetra Cell with Mini Trans-Blot module (Bio-Rad).
2. PowerPac Basic power supply (Bio-Rad).

3. 10 % sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) gel, 10 mL per two 1 mm-gels: 4.8 mL H_2O, 2.5 mL 1.5 M Tris–HCl pH 8, 2.5 mL 40 % Acrylamide/Bis 37.5:1, 100 µL 10 % SDS, 50 µL 10 % APS, 7 µL TEMED.
4. 4× SDS protein sample buffer: 40 % glycerol, 240 mM Tris/HCl pH 6.8, 8 % SDS 0.04 % bromophenol blue, 5 % beta-mercaptoethanol.
5. Prestained protein standards (Bio-Rad).
6. 10× Running buffer: 0.25 M Tris base, 1.92 M glycine, 1 % SDS, pH 8.3.
7. Polyvinylidene difluoride (PVDF) transfer membrane.
8. Transfer *N*-cyclohexyl-3-aminopropanesulfonic acid (CAPS) buffer: 10 mM CAPS, 10 % Methanol, pH 11.
9. Tris buffered saline + tween (TBST) washing buffer: 20 mM Tris–HCl, 137 mM NaCl, 0.2 % $Tween_{20}$, pH 8.
10. Blocking solution: 5 % skim milk in TBST.
11. Anti-mouse IA antibody (clone KL295) [13].
12. Anti-mouse IgG_1 HRP-conjugated antibody (Southern-Biotech).
13. Enhanced chemiluminescence (ECL) Plus Western Blotting Detection kit.
14. BioMax XAR Film (Kodak).
15. SRX-101A processor and IS-199 X-ray developer/fixer Auto mixer (Konica).

2.5 Solid Phase Extraction (SPE)

Store all SPE buffers at 4 °C.

1. SPE Equilibration buffer: 0.1 % (V/V) trifluoroacetic Acid (TFA) in high performance liquid chromatography (HPLC) grade water up to 1 L. Store at 4 °C.
2. SPE Washing buffer: 50 % (V/V) HPLC grade acetonitrile in HPLC grade water up to 1 L. Store at 4 °C.
3. SPE Elution buffer: 0.1 % TFA in 50 % HPLC grade acetonitrile in HPLC grade water up to 1 L. Store at 4 °C.
4. 10 kDa cutoff membrane filter (Sartorius Stedim, Aubagne, France).
5. C-18 cartridge (Waters, Milford, MA).

2.6 LC-MS/MS Analysis

Prepare all solutions for MS analysis using HPLC grade water and other analytical reagents. Store all chemical reagents at room temperature.

1. Aqueous mobile phase A: 0.1 % formic acid in water up to 1 L (Solvent A).

2. Organic mobile phase B: 0.1 % formic acid in acetonitrile up to 1 L (Solvent B).
3. Ultimate 3000 nano-HPLC (ThermoFisher, Sunnyvale, CA).
4. LTQ-Orbitrap (ThermoFisher, San Jose, CA).
5. Bioworks Software (ThermoFisher, San Jose, CA).
6. PepMap trap column (100 Å, 5 μm, 800 μm × 2 mm, Dionex, Sunnyvale, CA).
7. Acclaim Pepmap 100 C18 analytical column (100 Å, 3 μm, 75 μm × 15 cm, Dionex, Sunnyvale, CA).
8. Mascot software (Matrix Science, Boston, MA).

3 Methods

3.1 In Vivo DC Expansion

1. Inject 5×10^6 Flt3L B16 cells subcutaneously into the abdomen region of mice. After 15–20 days, all major splenic DC subsets had expanded >10-fold in agreement with previous reports [4, 14].

3.2 DC Prep

1. Remove Flt3L-treated spleens, cut in small fragments, and digest into single cell-suspensions with 400 U/mL collagenase D (Roche Applied Science) for 25 min at 37 °C.
2. Purify $CD11c^+$ DCs by positive selection using anti-CD11c magnetic beads and MACS columns as per manufacturer's instructions.
3. Wash DCs three times with PBS by pelletting cells for 5 min at $500 \times g$.

3.3 Generation of Affinity Columns

1. To begin coupling of CNBr Sepharose beads (adapted from ref. 15), measure the OD of the antibody solution at 280 nm by using a quartz cuvette and a spectrophotometer with UV.
2. Use 40 mg of Sepharose per mg of antibody based on 1 $OD_{280} = 0.7$ mg/mL antibody.
3. Take up Sepharose in 1 mM HCl for 30 min (rotate).
4. Centrifuge and discard supernatant.
5. Add purified N22 monoclonal antibody, 15 mg (*see* **Note 2**) to coupling buffer (adjust concentration to give 2–3 mg of antibody/mL of coupling buffer; measure OD_{280} before start).
6. Incubate for 60 min on a rotator.
7. Measure the OD_{280} to determine coupling efficiency (*see* **Note 3**) and incubate longer if necessary.
8. Centrifuge and discard the supernatant.

9. Incubate for 30 min with glycine solution on rotator.
10. Centrifuge and discard the supernatant.
11. Wash the beads twice with PBS and transfer them to a poly-prep column (*see* **Note 4**).
12. Store the finished column at 4 °C in PBS and use it within 24–48 h.

3.4 Affinity Purification of MHC II Molecules

1. Lyse cells with cell lysis buffer (10^8/mL) for 45–60 min at 4 °C on a rotator.
2. Clear cell lysate by 20 min centrifugation at 21,000 × *g* and transfer cleared lysate into a fresh tube.
3. Incubate cleared lysate with N22 antibody-conjugated CNBr Sepharose beads (*see* **Note 5**), overnight at 4 °C on rotator.
4. Spin down beads at 50 × *g* for 2 min.
5. Remove and save the flow through (*see* **Note 6**).
6. Wash beads six times with washing buffer 1 (6 mL/wash).
7. Wash beads six times with washing buffer 2 (6 mL/wash).
8. Elute MHC II-peptides complexes from beads with 1 mL of 10 % glacial acetic acid, 3 min at room temperature (RT) on rotator. Then, spin down for 3 min at 50 × *g* and recover elution fraction E1.
9. Repeat acid elution an additional four times: collect elution fractions E2, E3, E4, E5.
10. Boil each elution fraction for 10 min at 70 °C, then store at −20 °C.

3.5 Western Blot to Evaluate Yield of Immunoprecipitated MHC II-Peptide Complexes

General: detailed instructions for SDS-PAGE are provided in ref. 16.

1. Mix 5 µL aliquot of each elution fraction from **step 10** of Subheading 3.4 with 7.5 µL 4× SDS sample buffer; add 1 M triethylammonium bicarbonate (TEAB), to neutralize very acidic pH (<3), up to 30 µL volume/lane (*see* **Note 7**).
2. Heat samples to 100 °C for 4–5 min. Do not boil too long, proteins get destroyed/degraded.
3. Run on 10 % SDS-PAGE, under reducing condition, ~2 h at 30 mA, constant current, until the dye front reaches the bottom of the gel.

General: detailed instructions for blotting under wet conditions are found in ref. 17.

1. Pre-wet PVDF membrane in methanol (5 min is sufficient) and then in transfer buffer. Pre-wet sponge and filter paper in transfer buffer.

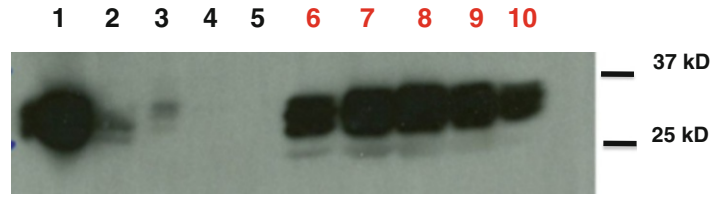


Fig. 2 Western blot analysis of immunoaffinity purified MHC II-peptide complexes. A cell lysate made of Flt3L DCs was first immunoprecipitated for class II molecules on N22 antibody-CNBr Sepharose beads. (*1*) Cell lysate, as 0.3 % input, and (*2*) flow through from beads, 15 μL/lane. (*3–5*) Serial washes with Tris–HCl pH 8/250 mM NaCl (*3*, *4*), Tris–HCl pH 8 (*5*), 15 μL/lane. (*6–10*) Acid eluted MHC II-complexes, E1–E5 fractions, 5 μL/lane. Membrane was immunoblotted with anti-IA antibody (KL295 sup, 1:1,000) and with an anti-mouse IgG_1 HRP-conjugated (1:10,000) as secondary antibody

2. Assemble transfer components included in the Mini Trans-Blot module (as per manufacturer's instructions): the gel and membrane are sandwiched between sponge and paper (sponge/paper/gel/membrane/paper/sponge) and all are clamped tightly together after ensuring no air bubbles have formed between the gel and membrane.
3. Submerge sandwich in transfer buffer: place the gel/filter sandwich in the holder such that the proteins will migrate from the gel to the membrane, in the direction of the positive (red) electrode (anode).
4. Apply electrophoresis to PVDF membrane for 2 h at 300 mA, constant current.
5. Block PVDF membrane with blocking solution for 1 h at RT.
6. Blot membrane with anti-mouse I-A primary antibody, clone KL295 (1:1,000 dilution in blocking solution) and incubate on rotator for 1 h 30 min at RT or overnight at 4 °C.
7. Wash six times for 10 min each in ~30 mL of TBST.
8. Add a secondary anti-mouse IgG_1 HRP-conjugated antibody (1:10,000 dilution in blocking solution) and incubate on rotator for 1 h 30 min at RT.
9. Wash as in **step 7**.
10. Detect immunoreactive protein bands with ECL Plus Western Blotting Detection Kit. Add 2 mL of ECL solution onto the membrane and wait for 5 min. Drain ECL solution, wrap in plastic.
11. Place the blot into an autoradiography cassette, and expose to film (in a dark room) for 5 s to 2 min depending on the signal. Insert film in film processor according to manufacturer's instructions (*see* Fig. 2).

3.6 MHC Peptides Purification and Concentration

1. Pipette acid-extracted MHC peptide mixture into a 10 kDa cutoff membrane filter.
2. Centrifuge filter at 3,000 × *g* for 15 min.
3. Collect the flow-through from the filter.
4. Wash C-18 cartridge (Waters, Milford, MA) with 1 mL 50 % acetonitrile in water and discard the flow-through. Repeat this three times.
5. Equilibrate the C-18 cartridge with 1 mL 0.1 % TFA in water, and discard the flow-through. Repeat this four times.
6. Load the peptides mixture from membrane filter to the C-18 cartridge and adjust the vacuum so the flow is very slow (*see* **Note 8**).
7. Wash the C-18 cartridge with 1 mL 0.1 % TFA in water and discard the flow-through. Repeat this three times.
8. Elute the peptides by adding 500 μL of 0.1 % TFA in 50 % acetonitrile and keep the flow-through. Repeat this two times, and combine all the flow-throughs.
9. Speed-vac the eluate to dryness.
10. Prior to MS analysis, reconstitute dried peptides by adding 20 μL 0.1 % TFA in water.

3.7 LC-MS/MS Analysis

1. Load the peptide mixture onto a Pepmap trap column at a flow rate of 30 μL/mL running with aqueous mobile phase A using a Dionex Ultimate 3000 nanoHPLC coupled with LTQ-Orbitrap for LC-MS/MS analysis. After 5 min of loading and washing, switch the trap column online with the analytical Acclaim Pepmap 100 C18 column running at 250 nL/min.
2. Separate the peptide mixture with a shallow gradient: 0–55 % organic solvent B in 120 min, followed by 25 min gradient from 55 to 80 % organic solvent B. Organic solvent B is maintained at 80 % for another 10 min and then decreased to 0 % in 10 min. Re-equilibrate the column at 0 % mobile phase B for another 10 min. The flow rate was maintained at 250 nL/min.
3. Operate the LTQ-Orbitrap in the data-dependent acquisition mode with a MS full scan (620–1,200 m/z; 30,000 resolution) followed by six data dependent MS/MS scans in the ion trap at 35 % normalized collision energy (*see* **Note 9**).
4. Use the following dynamic exclusion parameters, to maximize acquisition of data from more components in a complex peptide mixture: Repeat count = 1; Repeat duration = 30 s; Exclusion list = 100: and Exclusion time = 90 s.

3.8 Database Searching

1. In order to prepare the database search with MASCOT, the Raw data files from LTQ-Orbitrap need to be converted to a single combined DTA file. Open the RAW data file in Bioworks

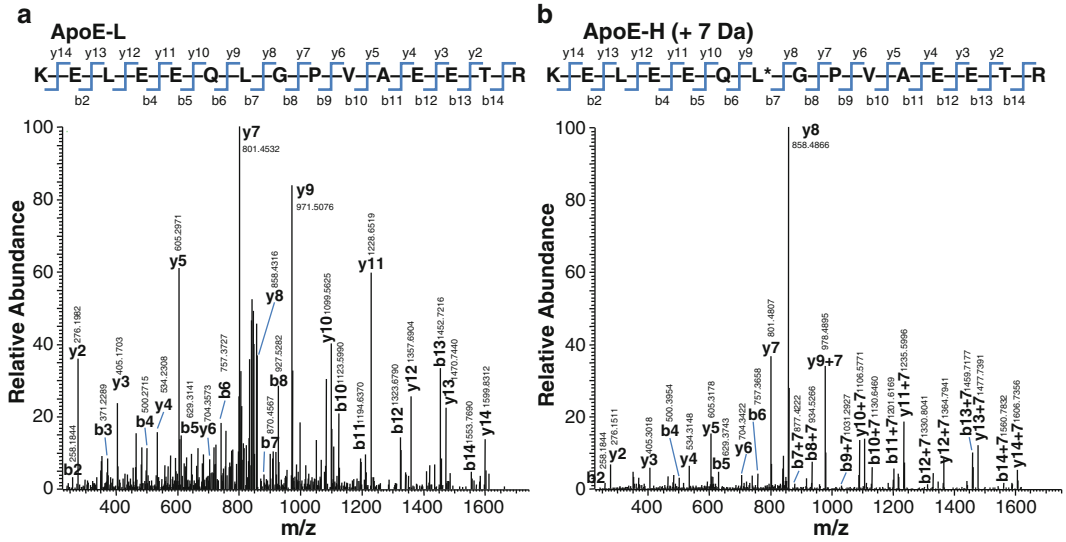


Fig. 3 Identification of a representative MHC class II-associated peptide derived from DCs. Comparison of MS/MS spectra of naturally eluted apolipoprotein E (APOE) peptide (**a**), identified using MASCOT software in a representative DC sample, with MS/MS spectra of the matching synthetic isotope-labeled APOE peptide (**b**). The corresponding *y* and *b* series are marked

software (ThermoFisher, San Jose, CA) and run the DTA extraction tool. Finally combine all the extracted DTA files to a single DTA file.

2. Search Peptides against IPI (International Protein Index) mouse database on the converted DTA file by using MASCOT software. The following search parameters are used in all MASCOT searches: the digestion enzyme is set as none and methionine oxidation as the variable modification. The maximum error tolerance for MS scans is 10 ppm for MS and 1.0 Da for MS/MS respectively.
3. Manually inspect MS/MS spectra to confirm that the major fragmented ions match the identified peptide sequences (*see* **Note 10**).
4. Compare MS/MS spectra of eluted natural peptides with the MS/MS spectra obtained from the corresponding synthetic (isotopically labeled) counterparts, as shown in Fig. 3 (*see* **Note 11**).

3.9 MHC II Peptide Quantitation

1. Mix 10 µL peptide mixture with 10 µL isotope labeled peptides dissolved in 50 % acetonitrile/0.1 % TFA in water.
2. Load the peptides mixture onto a nano LC-LTQ Orbitrap (the LTQ-Orbitrap is operated in full scan mode with a mass resolution of 30,000). The MS settings are identical to the ones indicated in the experimental **step 3** of Subheading 3.7.

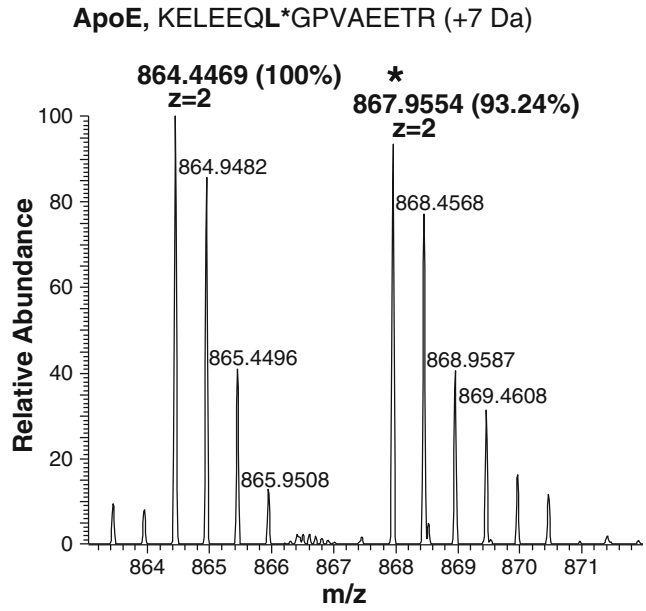


Fig. 4 Quantitation analysis by LC-MS of selected MHC II endogenous peptides. Synthetic isotope-labeled peptides (0.8 ng) were spiked into a representative DC sample. Quantitation of the endogenous counterpart was obtained comparing peaks' intensity of the selected peptide pair. The heavy isotope peak is indicated with *asterisk*. MS profile of the APO E peptide pair identified in the MHC II peptide mixture eluted from one representative DC sample. Isotopic patterns of the ions were consistent with the predicted patterns based on the isotopic ratios. Mass shift of the isotope-labeled peptide is consistent with the predicted value

3. Compare the peak intensity of the native peptides and isotope labeled peptides as a function of *m/z* values.
4. Determine the amount of the native peptides based on the ratio of peak intensities of the native peptides and isotope labeled peptides and the exact amount of isotope labeled peptides spiked in the peptides (*see* **Note 12**). As an example, Fig. 4 illustrates the intensity's comparison of an endogenous MHC II peptide and its synthetic isotope labeled counterpart at *m/z* 864.4469 and *m/z* 867.9954, respectively. Using the ratio of the intensity observed in the spectra (100 vs. 93.24 %), and the known quantity of isotopically labeled peptide spiked into the sample (0.4 ng/10 μL sample), the quantity of the endogenous peptide can be extrapolated. Thus, our calculation results in $100/93.24 \times 0.4$ ng of standard/10 μL sample $\times 2 =$ total 0.858 ng of endogenous peptide/20 μL sample.

4 Notes

1. Alternatively, replace protease inhibitors with Mini Protease Inhibitor Tablet (Roche Applied Science) and add iodoacetamide and PMSF as above.
2. For optimal preparation, the antibody is dialyzed overnight against 0.5–1 L of 0.1 M $NaHCO_3$, pH 8.3, 0.5 M NaCl buffer using dialysis tubing. Change the buffer once after 2–4 h. Recover antibody and store at 4 °C.
3. The efficiency of antibody coupling to CNBr Sepharose beads is determined by measuring protein concentration in the supernatant obtained from the coupling reaction (Subheading 3.3, **step 6**). The protein remaining in these supernatant fractions represents antibody that was not immobilized on beads. The efficiency of conjugation can be estimated by:
   (a) Percentage bound = 100 [(total mg antibody added) – (mg antibody in supernatant)]/(total mg antibody added).
   (b) Efficiencies less than 70 % indicate that very inefficient antibody immobilization has occurred. The cause of poor coupling is often the presence of traces of buffer components (free amines in Tris–HCl buffer) that competes with primary amines of the antibody for binding at the active sites of the CNBr activated Sepharose beads.
4. Rinse the beads with several bed volumes of PBS. Monitor the level of liquid in the reservoir and do not allow the beads to become dry. Discard rinse.
5. Alternatively, protein G-Sepharose beads (~1 mL volume), previously washed with PBS are saturated with 15 mg of N22 antibody, 2 h binding at RT or overnight at 4 °C on rotator. Then the lysate is incubated with N22-saturated protein G beads, overnight at 4 °C on rotator.
6. We recommend that cell lysates and flow through not be frozen or stored in order to not compromise peptide recovery. Optionally, the flow through from the anti-MHC II antibody column can be loaded on a serial MHC class I antibody column in order to immunoprecipitate MHC I-peptide complexes.
7. For Western blot sample preparation, 5 μL aliquots of acid extracted MHC-peptide proteins are diluted to the final volume of 30 μL/lane with 1 M triethylammonium bicarbonate (TEAB) to neutralize very acidic pH (<3). The rest of samples, which are constituted of Tris–HCl pH 8 buffer, can be diluted safely in PBS.

8. Alternately, let material flow through the C-18 cartridge by gravity.
9. This means that the most intense six masses from each full mass spectrum with doubly and triply charged states are selected for fragmentation by collision-induced dissociation in the linear ion-trap. The 35 % normalized collision energy ensures that the optimum MS/MS data are automatically collected, independently on the mass of the analyte. This is particularly relevant when analyzing mixture where peptides have masses that range from below 600 Da to over 2,000 Da.
10. For peptides with Mascot scores of 20–30, we recommend manually inspecting their MS/MS spectra to confirm that the major fragmented ions matched the identified peptide sequences.
11. To unambiguously confirm peptide identification, we recommend comparison of MS/MS spectra of natural eluted peptides, identified by MASCOT software in DC samples, with MS/MS spectra of the matching synthetic (isotopically labeled) peptides. The identical fragmentation patterns obtained in the MS/MS spectra confirm the identity of the peptide sequence between the native peptide and the synthetic isotopically labeled counterpart.
12. The absolute quantification is determined by comparing the peak intensity of the native peptide with the peak intensity of heavy peptide added at 0.8 ng. The actual copy numbers of MHC II-bound peptides from DCs were calculated as follows: (a) moles of native peptide determined by quantitation analysis = g/MW; (b) Molecules of native peptide = moles (from **step 1**) $\times 6.022 \times 10^{23}$/mol; (c) Molecules of native peptide per cell = molecules (from **step 2**)/number of cells used per quantitation analysis.

Acknowledgments

We thank Dr. Svetlana Mojsov, the Rockefeller University, New York and Dr. Jörn Dengjel, University of Freiburg, Freiburg, for critical discussion and analytical tools. We are grateful to Marguerite Nulty for help with the references. This work was supported by grants from National Institute of Allergy and Infectious Diseases (NIAID), AI13013, 081677, and AI051573 to the late Ralph M. Steinman.

References

1. Banchereau J, Steinman RM (1998) Dendritic cells and the control of immunity. Nature 392:245–252
2. Mellman I, Steinman RM (2001) Dendritic cells: specialized and regulated antigen processing machines. Cell 106:255–258
3. Bozzacco L, Yu H, Zebroski HA et al (2011) Mass spectrometry analysis and quantitation of peptides presented on the MHC II molecules of mouse spleen dendritic cells. J Proteome Res 10:5016–5030
4. Bozzacco L, Trumpfheller C, Huang Y et al (2010) HIV gag protein is efficiently cross-presented when targeted with an antibody towards the DEC-205 receptor in Flt3 ligand-mobilized murine DC. Eur J Immunol 40: 36–46
5. Maraskovsky E, Brasel K, Teepe M et al (1996) Dramatic increase in the numbers of functionally mature dendritic cells in Flt3 ligand-treated mice: multiple dendritic cell subpopulations identified. J Exp Med 184:1953–1962
6. O'Keeffe M, Hochrein H, Vremec D et al (2002) Effects of administration of progenipoietin 1, Flt-3 ligand, granulocyte colony-stimulating factor, and pegylated granulocyte–macrophage colony-stimulating factor on dendritic cell subsets in mice. Blood 99:2122–2130
7. Curti A, Fogli M, Ratta M et al (2001) Stem cell factor and FLT3-Ligand are strictly required to sustain the long-term expansion of primitive CD34(+)DR(−) dendritic cell precursors. J Immunol 166:848–854
8. D'Amico A, Wu L (2003) The early progenitors of mouse dendritic cells and plasmacytoid predendritic cells are within the bone marrow hemopoietic precursors expressing Flt3. J Exp Med 198:293–303
9. Bozzacco L, Yu H, Dengjel J et al (2012) Strategy for identifying dendritic cell-processed $CD4^+$ T cell epitopes from HIV gag p24 peptides. PLoS One. 7(7):e41897. doi: 10.1371/journal.pone.0041897. Epub 2012 Jul 30
10. Gerber SA, Rush J, Stemman O et al (2003) Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. Proc Natl Acad Sci U S A 100: 6940–6945
11. Dranoff G, Jaffee E, Lazenby A et al (1993) Vaccination with irradiated tumor cells engineered to secrete murine granulocyte–macrophage colony-stimulating factor stimulates potent, specific, and long-lasting anti-tumor immunity. Proc Natl Acad Sci U S A 90: 3539–3543
12. Metlay JP, Witmer-Pack MD, Agger R et al (1990) The distinct leukocyte integrins of mouse spleen dendritic cells as identified with new hamster monoclonal antibodies. J Exp Med 171:1753–1771
13. Kang Y-S, Yamazaki S, Iyoda T et al (2003) SIGN-R1, a novel C-type lectin expressed by marginal zone macrophages in spleen, mediates uptake of the polysaccharide dextran. Int Immunol 15:177–186
14. Pulendran B, Lingappa J, Kennedy MK et al (1997) Developmental pathways of dendritic cells in vivo: distinct function, phenotype, and localization of dendritic cell subsets in FLT3 ligand-treated mice. J Immunol 159: 2222–2231
15. Shelver WL (2005) Immunochemical protocols. In: Burns R (ed) Methods in molecular biology. Humana Press, Totowa, NJ
16. Hames BD, Rickwood D (1998) Gel electrophoresis of proteins: a practical approach. Oxford University Press, Oxford
17. Bolt MW, Mahoney PA (1997) High-efficiency blotting of proteins of diverse sizes following sodium dodecyl sulfate-polyacrylamide, gel elecrophoresis. Anal Biochem 247: 185–192

Chapter 15

Construction and Screening of an Antigen-Derived Peptide Library Displayed on Yeast Cell Surface for CD4+ T Cell Epitope Identification

Fei Wen and Huimin Zhao

Abstract

Identification of T cell epitopes is a critical, but often difficult step in studying T cell function and developing peptide-based vaccines and immunotherapies. Unlike antibodies that recognize free soluble antigens, T cell receptor (TCR) recognizes its epitope bound to major histocompatibility complex (MHC) expressed on antigen presenting cells (APCs). In addition, the examination of T cell epitope activity requires the use of professional APCs, which are difficult to isolate, expand, and maintain. To address these issues, we have developed a facile, accurate, and high-throughput method for T cell epitope mapping by screening antigen-derived peptide libraries in complex with MHC protein displayed on yeast cell surface. Here, we use hemagglutinin and influenza A virus X31/A/Aichi/68 as examples to describe the key steps in identification of CD4+ T cell epitopes from a single antigenic protein and the entire genome of a pathogen, respectively. Methods for single-chain peptide-MHC complex vector design, yeast surface display, peptide library generation in *Escherichia coli*, and functional screening in *Saccharomyces cerevisiae* are discussed.

Key words CD4+ T cell epitope mapping, Peptide library, Major histocompatibility complex, Single-chain peptide-MHC complex, HLA-DR1, Yeast display, Influenza A virus, Flow cytometry, High throughput screening

1 Introduction

T cells respond to selected peptides (termed T cell epitopes) in complex with major histocompatibility complex (MHC) molecules expressed on antigen presentation cells (APCs) through their unique surface receptors (T cell receptors, TCRs) [1]. With their critical role in T cell development and activation, identification of T cell epitopes is very important in studying T cell lineages and phenotypes [2, 3], elucidating self-tolerance mechanisms [4], vaccine design and assessment [5, 6], tracking T cell *in vivo* [7], etc. Much effort has been devoted to identification of these T cell

Kelly M. Fulton and Susan M. Twine (eds.), *Immunoproteomics: Methods and Protocols*, Methods in Molecular Biology, vol. 1061,
DOI 10.1007/978-1-62703-589-7_15,

epitopes in the last two decades. The most straightforward T cell epitope identification method is to use HPLC-MS to determine the sequence of the peptides extracted directly from the peptide-MHC (pMHC) complexes on the surface of professional APCs [8]. This method has met limited success due to the limited amount of peptides attainable and representation of antigenic peptides as nested sets with heterogeneous lengths.

When the antigenic protein sequence is known, the most commonly used T cell epitope identification method is to chemically synthesize overlapping peptides spanning the entire protein, which are then individually loaded onto professional APCs (usually irradiated peripheral blood mononuclear cells) and tested for their ability of activating T cells [9–11]. This method becomes impractical when the overall size of the antigenic protein increases (e.g., viruses with large genome size) and is not applicable when the antigen sequence is unknown. With the knowledge of MHC-binding motifs, computational algorithms, such as TEPITOPE [12], could be used first to pre-select MHC class II (MHCII)-binding peptides and reduce the number of peptides to be synthesized and tested, but may overlook some candidates due to the complexity of peptide-MHC interaction. As an alternative and more comprehensive approach, combinatorial synthetic peptide libraries with up to a trillion variants, such as positional scanning synthetic combinatorial libraries (PS-SCLs) [13] and bead-bound libraries [14], could be employed. These synthetic combinatorial peptide libraries are valuable especially when the relevant target antigen is unknown. However, generally speaking, all the chemistry-based epitope mapping methods are expensive and nonrenewable, thus several expression cloning strategies have been developed to generate DNA libraries encoding either random peptides or antigenic proteins.

When the restriction MHC binding motif is known, a random peptide library with fixed MHC anchor and variable TCR-contacting residues can be constructed using degenerate primers. As with the combinatorial peptide library-based methods described above, the library screening usually results in the identification of T cell mimotopes (peptides that structurally mimic the antigenic epitope) rather than the native epitopes due to the random nature of the peptide library. While the mimotopes could be used to elucidate the sequence of biologically relevant epitopes [15, 16], they could also have very little resemblance [17]. To directly identify T cell epitopes, several groups have developed methods by either expressing a cDNA library in engineered APCs [18] or expressing viral cDNA libraries in *E. coli* to identify the antigenic protein first [19], so that the epitopes could be identified by testing chemically synthesized overlapping peptides. These expression cloning methods usually require either extensive engineering of a cell line for efficient processing and presenting MHCII restricted antigens that are usually of exogenous origin, or the use of professional APCs

that are often difficult to isolate, expand and maintain. Therefore, the screening process is often time-consuming, laborious, and reagent-intensive.

To address some of the limitations discussed above, we have developed a facile and high throughput CD4+ T cell epitope mapping method by displaying pathogen-derived peptide libraries in complex with the restriction MHCII protein on yeast cell surface [20]. It presents the advantages of both combinatorial peptide libraries and expression cloning, and allows direct epitope identification from either known or unknown pathogens. Here, we use hemagglutinin and influenza virus X31/A/Aichi/68 as an example for each case. Both antigens contain $HA_{306\text{-}318}$ peptide (PKYVKQNTLKLAT) that has been well characterized in complex with the human MHCII allele, DR1. The design of the epitope mapping method is shown in Fig. 1. DNA encoding the pathogenic protein(s) is obtained by PCR or RT-PCR using sequence-specific or random hexamer primers (Fig. 1a) for known and unknown antigens, respectively. For RT-PCR, RNase H and DNA polymerase I are used to generate the second strand cDNA. The resulting double stranded DNA is then randomly digested by DNase I into fragments with size ranging from 30 to 60 nucleotides that encode all the possible peptides with 10–20 amino acids from the antigens, which are then purified and blunt-end polished (T4 DNA polymerase and Klenow fragment). Meanwhile, the yeast display vector containing the gene encoding the restriction single-chain MHCII molecule, e.g., DR1 (Fig. 1b), is digested with NotI/SpeI (*see* **Note 1**), blunt-end polished, and then ligated to the DNA fragments (Fig. 1c). Following gene expression, yeast cells displaying the peptide library in complex of DR1 are analyzed by fluorescence-activated cell sorting (FACS) to enrich those containing peptides with high DR1-binding affinity based on their surface expression levels (*see* **Note 2**). This enrichment cycle is repeated for ~2–4 times to obtain a small subset of yeast cells displaying DR1 in complex with good binding peptides, which are then directly screened for their ability to induce antigen-specific T cell activation as indicated by IL-2 secretion using T hybridoma cells transfected with the TCR of interest (*see* **Note 3**). DNA sequence analysis of the selected positive clones leads to the identification of antigenic peptides. If necessary, a series of DNA fragments encoding overlapping peptide sequences can be used to refine those antigenic peptide sequences.

2 Materials

All buffers and media were prepared using ultrapure water (ddH_2O) with a resistivity of 18.2 mΩ cm at room temperature, and were sterilized by autoclaving at 121 °C and 15 psi for 15 min or

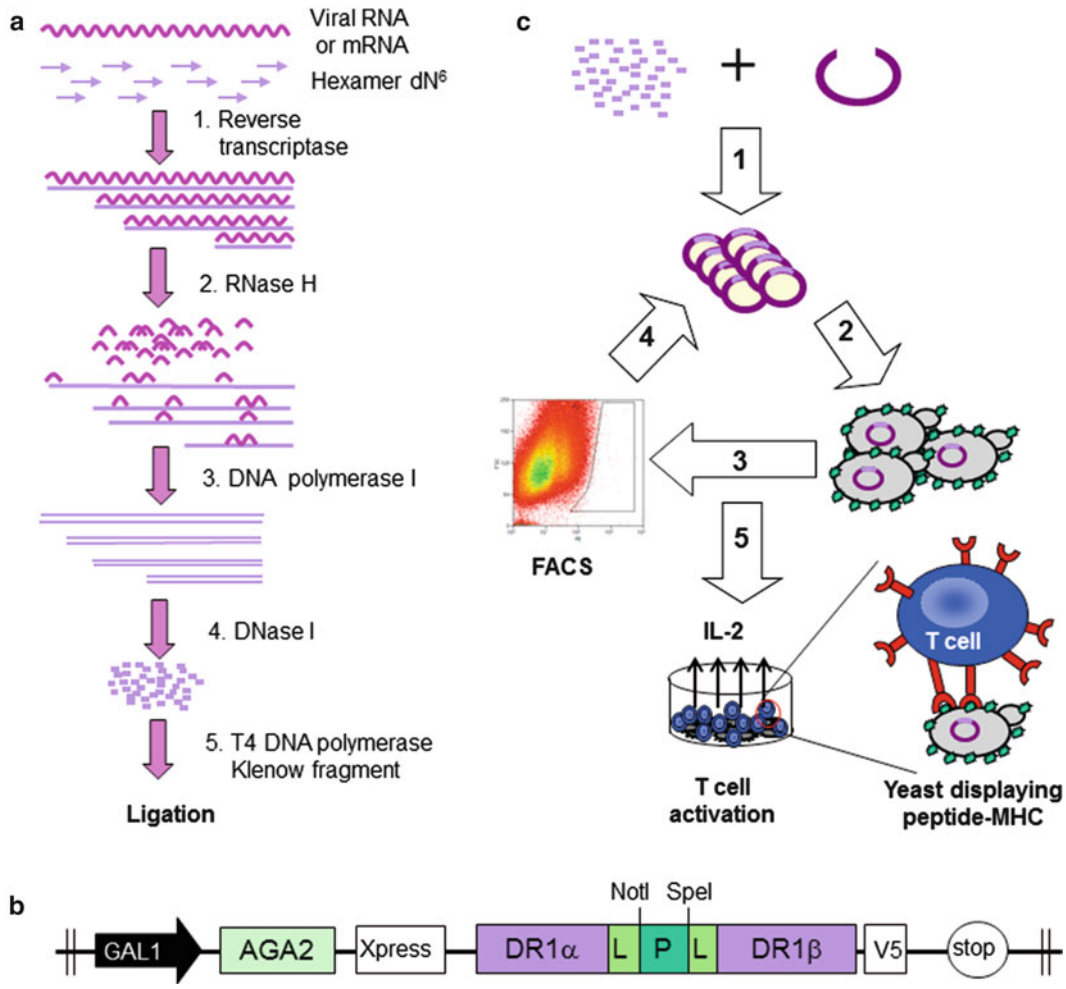


Fig. 1 Design of the CD4+ T cell epitope mapping method. (**a**) Generation of blunt-ended DNA fragments (~30–60 bp) encoding pathogen-derived peptides from genomic RNA or mRNA isolated from pathogens with unknown sequence. For antigens with known sequences, double stranded DNA could be readily amplified using specific primers by PCR and processed from *step 4* with DNase I treatment. (**b**) Schematic representation of the single-chain peptide-DR1 construct used for yeast display. GAL1 = yeast GAL1 promoter, AGA2 = an adhesion subunit of a-agglutinin of a-yeast-cells [25], Xpress = Xpress epitope, L = linker, P = peptide, V5 = V5 epitope. (**c**) Peptide library generation and screening. DNA fragments are ligated to the single-chain MHCII molecules in a yeast display vector (*step 1*). Following transformation and gene expression (*step 2*), yeast cells displaying the peptide library in complex of MHCII protein are analyzed by FACS (*step 3*) to identify a small subset of yeast cells containing peptides with high affinity toward the restriction MHCII. Plasmids are then recovered from these yeast cells (*step 4*) and analyzed for further enrichment. This enrichment cycle is usually repeated for ~2–4 rounds. Individual clones from the enriched library are then screened for their ability to induce antigen-specific T cell activation as indicated by IL-2 production using T cell hybridomas transfected with the TCR of interest (*step 5*) in a 96-well format

filtration through a 0.22 μm membrane. All restriction enzymes were obtained from New England Biolabs (Ipswich, MA) and stored at −20 °C and all antibodies were stored at 4 °C. All primers (listed in Table 1) were synthesized by Integrated DNA

Table 1
Primers used in cloning and peptide library creation

Name	Sequence
α-5BX	5′-GTACCAGGATCCAGTGTGGTGGAAAGGAAAGAAGAACATGTGATC-3′
β-3XH	5′-CCCTCTAGACTCGAGCTTGCTCTGTGCAGATTCAGAC-3′
pYD1For	5′-AGTAACGTTTGTCAGTAATTGC-3′
NotIRev	5′-TGCCAACTTCAGGGTGTTTTGCTTAACATACTTGGGGCGGCCGCCTCC TGAGCCTCCACC-3′
NotIFor	5′-GCAAAACACCCTGAAGTTGGCAACAGGTACCGGTGGCTCACTAG-3′
βRev73-67	5′-GGCCCGCCTCTGCTCCAGGA-3′
StfFor	5′-GGAGGCGGCCGCTTT TTG GATGGAGGAATTCATATG-3′
StfRev	5′-CTCACTAGTCGGGAAGACGTACGGGGTATACATGT-3′
AichiFor	5′-ATTCGCGGCCGCATGAAGACCATCATTGCTTTGAGCTACATTTTC-3′
AichiRev	5′-CTAATAACTAGTAATGCAAATGTTGCACCTAATGTTGCCTCTCTG-3′
2467For	5′-GGCCGCCCCAAGTATAGAAAGATGAACGCACGAAAGTTGGCAACAGGTA CCGGTGGCTCA-3′
2467Rev	5′-CTAGTGAGCCACCGGTACCTGTTGCCAACTTTCGTGCGTTCATCTT TCTATACTTGGGGC-3′
M4For	5′-GGCCGCGGAGGTTATAGACAGATGTCAGCACCAACTTTGGGAGGCGGTA CCGGTGGCTCA-3′
M4Rev	5′-CTAGTGAGCCACCGGTACCGCCTCCCAAAGTTGGTGCTGACATCTGTC TATAACCTCCGC-3′
YAKFor	5′-GGCCGCGCCGCATATGCCGCAGCGGCTGCCGCAAAGGCTGCCGCAGG TACCGGTGGCTCA-3′
YAKRev	5′-CTAGTGAGCCACCGGTACCTGCGGCAGCCTTTGCGGCAGCCGCTGCGG CATATGCGGCGC-3′
CIIFor	5′-GGCCGCGCTGGGTTTAAGGGGGAACAGGGACCTAAAGGAGAGCCTGG TACCGGTGGCTCA-3′
CIIRev	5′-CTAGTGAGCCACCGGTACCAGGCTCTCCTTTAGGTCCCTGTTCCCCCTT AAACCCAGCGC-3′
PKAFor	5′-GGCCGCCCCAAGGCTGTTAAGCAAAACACCCTGAAGTTGGCAACAGGT ACCGGTGGCTCA-3′
PKARev	5′-CTAGTGAGCCACCGGTACCTGTTGCCAACTTCAGGGTGTTTTGCTTAA CAGCCTTGGGGC-3′
LFor	5′-GGCCGCGGAGGTGGAGGCTCCGGAGGTGGAGGCTCAGGAGGTGGAGG TACCGGTGGCTCA-3′
LRev	5′-CTAGTGAGCCACCGGTACCTCCACCTCCTGAGCCTCCACCTCCGGA GCCTCCACCTCCGC-3′
MBPNotI	5′-GGAGGCGGCCGCGAAAACCCGGTTGTTCACTTCTTCAAAAACATCGTT ACCCCGCGTGGTACCGGTGGCTCACTAGTGA-3′

Technologies (Coralville, IA) and stored at −20 °C at a concentration of 100 μM. All other reagents were stored at room temperature unless indicated otherwise.

2.1 Vector Construction and DNA Fragments Generation

1. pJ3/238: A gift from M. Mage (NIH, Bethesda, MD) was used as the PCR template for amplifying the DNA encoding a single-chain polypeptide consisting of the α-chain of DR1, the peptide $HA_{306\text{-}318}$, and the β-chain of DR1 (*see* Fig. 1b).
2. 5× Phusion HF Reaction Buffer and Phusion DNA polymerase (New England Biolabs, Beverly, MA, USA).
3. 40× dNTPs premix: 10 mM each nucleotide (Promega, Madison, WI, USA).
4. 0.5 M ethylenediamine tetraacetic acid (EDTA) solution (pH 8.0): For a 500 mL of stock solution of 0.5 M EDTA, weigh out 93.05 g of EDTA disodium salt (MW = 372.2) and dissolve it in 400 mL of deionized water. Adjust to pH 8.0 with NaOH and correct the final volume to 500 mL. EDTA will not be dissolved completely in water unless the pH is adjusted to about 8.0.
5. 50× Tris-acetate-EDTA (TAE) stock solution: Dissolve 242 g of Tris base (MW = 121.14) in approximately 750 mL of deionized water. Carefully add 57.1 mL of acetic acid and 100 mL of 0.5 M EDTA, and add deionized water to make a final volume of 1 L. The pH of this buffer is not adjusted and should be about 8.5.
6. Working solution of TAE buffer (1×): Dilute the stock solution by 50-fold with deionized water. Final solute concentrations are 40 mM Tris acetate and 1 mM EDTA.
7. 0.7 % (1 %, 2 %) Agarose gel in 1× TAE buffer: Add 0.7 g (1 g, 2 g) of agarose into 100 mL of 1× TAE buffer and microwave until agarose is completely melted. Cool the solution to approximately 70–80 °C. Add 5 μL of ethidium bromide into the solution and mix well. Pour 25–30 mL of solution onto an agarose gel rack with appropriate 2-well (for gel purification) or 8-well (for checking PCR products) combs.
8. QIAquick Gel Extraction Kit (QIAGEN, Valencia, CA, USA).
9. NanoDrop2000c: Used to measure the concentration of DNA (Thermo Scientific, Wilmington, DE, USA).
10. Precision Molecular Mass Standard (Bio-Rad, Hercules, CA).
11. pYD1: Obtained from Invitrogen (Carlsbad, CA) and encodes AGA2, an adhesion subunit of a-agglutinin of a-yeast-cells, which enables yeast surface display of a target protein.
12. QIAquick PCR Purification Kit (QIAGEN, Valencia, CA, USA).

13. T4 DNA ligase and 10× T4 DNA ligase buffer (New England Biolabs, Ipswich, MA).
14. DH5α chemical competent cells: Obtained from Media Preparation Facility (University of Illinois, Urbana, IL) for cloning.
15. 100 mg/mL ampicillin stock solution: Dissolve 1 g of ampicillin powder in 10 mL of ddH_2O and filter-sterilize.
16. LB broth: Add 10 g of bacto-tryptone, 5 g of yeast extract, 10 g of NaCl into 1 L of ddH_2O and autoclave.
17. LB-Amp agar plates: Autoclave LB-agar and when the solution cools down to 70–80 °C, add 1 mL of 100 mg/mL ampicillin to 1 L of LB-agar. Pour 20–25 mL into each Petri dish.
18. QIAprep Miniprep Kit (QIAGEN, Valencia, CA, USA).
19. *S. cerevisiae* EBY100 (MATa ura3-52 trp1 leu2Δ 1 his3Δ 200 pep4::HIS3 prb1Δ1.6R can1 GAL (pIU211:URA3)): Obtained from Invitrogen for yeast surface display of protein of interest and plasmid construction through homologous recombination.
20. Zymoprep II yeast plasmid miniprep (Zymo Research, Orange, CA, USA).
21. 50 mg/mL kanamycin stock solution: Dissolve 1 g of kanamycin powder in 20 mL of ddH_2O and filter-sterilize.
22. SD-CAA medium: Dissolve 20 g dextrose, 6.7 g yeast nitrogen base, 10 g casamino acids in 1 L of ddH_2O and sterilize by autoclaving.
23. SD-CAA agar plates: autoclave SD-CAA-agar and let cool down to 70–80 °C, add 50 μg/mL kanamycin. Pour 20–25 mL into each Petri dish.
24. Optikinase and 10× reaction buffer: Obtained from USB (Cleveland, OH) and was used to phosphorylate DNA inserts for ligation.
25. Influenza A/Aichi/2/68 viral RNA: A gift from R. Donis (CDC, Atlanta, GA) was used as template to amplify the hemagglutinin gene.
26. SuperScriptIII One-Step RT-PCR System with Platinum Taq High Fidelity (Invitrogen, Carlsbad, CA).
27. Transcriptor First Strand cDNA Synthesis Kit (Roche, Indianapolis, IN).
28. DNase I, DNA polymerase I, RNase H, T4 DNA polymerase, Klenow fragment, and NEBuffer 2 were obtained from New England Biolabs (Ipswich, MA) and stored at –20 °C.
29. QIAEX II Gel Extraction Kit (QIAGEN, Valencia, CA, USA).
30. 10 mg/mL bovine serum albumin (BSA) solution (New England Biolabs, Ipswich, MA).

31. Shrimp alkaline phosphatase (SAP) and 10× SAP buffer (Fermentas, Glen Burnie, Maryland).
32. ElectroMax DH5α competent cells: Obtained from Invitrogen (Carlsbad, CA) for peptide library construction by electroporation.
33. Influenza virus X31/A/Aichi/68 genomic RNA: Obtained from Charles River Laboratories (Wilmington, MA) and used as the template for cDNA synthesis.

2.2 Cell Growth and Transformation

1. YPAD medium: Dissolve 10 g of yeast extract, 20 g of peptone, 20 g of dextrose, and 100 mg of adenine hemisulfate in 1 L of ddH_2O and sterilize by autoclaving.
2. LiAc solutions: To prepare 50 mL of 1 M solution, dissolve 3.3 g lithium acetate (MW = 65.99 g/mol) in ddH_2O and sterilize by filtration. Prepare 0.1 M solution by mixing 5 mL of the 1.0 M solution with 45 mL sterile ddH_2O.
3. ssDNA stock solution: Dissolve 200 mg of deoxyribonucleic acid Sodium Salt Type III from Salmon Testes (Sigma, St. Louis, MO) in 100 mL of Tris–EDTA (TE) buffer (10 mM Tris–HCl, 1 mM EDTA, pH 8.0) by vigorous mixing on a magnetic stirrer. Aliquot the ssDNA in 1 mL and store at −20 °C (*see* **Note 4**).
4. 50 % (w/v) polyethylene glycol (PEG): Add 50 g PEG (MW = 3,350 g/mol) into 30 mL ddH_2O, heat the solution to ~60 °C with constant stirring to ensure quick dissolution of PEG and once completely dissolved, adjust volume with ddH_2O to 100 mL and sterilize by filtration.
5. Gene Pulser II and Pulse controller plus: Obtained from Bio-Rad (Hercules, CA) and used to transform plasmids or ligation mixture into *E. coli* through electroporation.
6. 1 M glucose solution: Dissolve 90 g of d-glucose in 400 mL of ddH_2O, adjust to a final volume of 500 mL, and filter-sterilize it.
7. SOC medium: Add 20 g of Bacto-tryptone, 5 g of yeast extract, 0.5 g of NaCl, 186.4 mg of KCl into 980 mL of ddH_2O. Adjust the pH to 7.0 with NaOH. Autoclave at 121 °C for 15 min. After the solution cools down to 70–80 °C, add 20 mL of sterile 1 M glucose.

2.3 Expression and Function Analysis of pMHC Complexes on Yeast Cell Surface

1. YPG medium: Dissolve 10 g yeast extract, 20 g peptone and 20 g galactose in 1 L of ddH_2O and sterilize by autoclaving.
2. Phosphate-buffered saline (PBS) (VWR, Radnor, PA).
3. Bovine serum albumin (BSA): Obtained from Sigma (St. Louis, MO) and used at a final concentration of 0.5 % in PBS for all antibody staining and washing steps.

4. Antibodies: anti-V5 (Invitrogen, Carlsbad, CA), LB3.1 (American Tissue Culture Collection (ATCC), Manassas, VA), and goat-anti-mouse (GAM) IgG (Rockland, Gilbertsville, PA).
5. Streptavidin–phycoerythrin (SA-PE) conjugate (eBiosciences, San Diego, CA).
6. HA1.7 hybridoma T cells: A gift from J. Bill [17] was used as an indicator cell line for productive DR1-peptide-TCR interactions that produce IL-2 upon receptor engagement.
7. Complete Iscove's Modified Dulbecco's Media (IMDM) for HA1.7 culture: IMDM (Invitrogen, Carlsbad, CA) supplemented with 10 % fetal bovine serum (Biomeda, Forster City, CA) and Penicillin (10 U/mL), streptomycin (10 μg/mL).
8. Murine IL-2 enzyme-linked immunosorbent assay (ELISA) kit (eBiosciences, San Diego, CA).

3 Methods

3.1 Vector Construction for Yeast Surface Display of Single-Chain pMHC Complexes

1. Amplify DNA encoding single-chain DR1α-linker-$HA_{306-318}$-linker-DR1b using α-5BX and β-3XH as primers (*see* Table 1) and plasmid pJ3/238 as template. Set up the reaction mixture as following: 10 ng of DNA template, 1× HF Phusion buffer, 50 pmol of forward primer, 50 pmol of reverse primer, 0.25 mM (each) dNTPs, 2 U of Phusion DNA polymerase, and add ddH_2O to make up a final volume of 100 μL. Unless otherwise specified, this recipe was used for all PCR reactions.
2. PCR thermocycler program: an initial denaturation of 2 min at 98 °C, followed by 25–30 cycles of denaturation at 98 °C for 30 s, annealing at 55 °C for 30 s, and extension at 72 °C for 60 s, and a final 10-min elongation at 72 °C. Unless otherwise specified, this program was used for all PCR reactions with the addition of extension time of 30 s per additional 1 kb long gene.
3. Load the PCR reaction mixture onto 0.7 % agarose gels and perform electrophoresis in 1× TAE buffer at 120 V for 20 min.
4. Gel-purify the DNA band with correct size using the QIAquick Gel Extraction Kit and check the concentration using NanoDrop, as per the manufacturer's instructions (*see* **Note 5**).
5. Digest 1 μg of the purified PCR product (insert) and 1 μg of plasmid pYD1 (vector) with BstXI and XhoI and purify using the QIAquick PCR Purification Kit.
6. Ligate the digested insert and vector to generate plasmid pYD1scαHAβ at 16 °C overnight. Set up the ligation mixture as following: 50 ng of vector, 100 ng of insert, 1× T4 ligase buffer, 1 μL of T4 ligase, and add ddH_2O to make up a final volume of 20 μL.

7. Transform 5 μL of the ligation mixture into chemical competent DH5α cells (*see* **Note 6**) and select transformants on a LB-Amp+ agar plate in a 37 °C incubator.
8. Pick a single colony into 3 mL LB medium with 100 μg/mL ampicillin, grow at 37 °C with 250 rpm agitation overnight, isolate the plasmid using the QIAprep Miniprep Kit, as per the manufacturer's instructions, and confirm the sequence of pYD1scαHAβ by DNA sequencing.
9. Introduce a NotI restriction site directly upstream of $HA_{306\text{-}318}$ (*see* Fig. 1b) to facilitate swapping of different peptides in the single-chain pMHC construct as follows: Perform two PCR reactions using pYD1scαHAβ as template and pYD1For/NotIRev and NotIFor/βRev73-67 as primers (*see* Table 1), respectively. Purify the PCR products using QIAquick PCR Purification Kit and splice them as described elsewhere [21]. Co-transform the spliced DNA and pYD1scαHAβ digested with BstXI and SpeI into *S. cerevisiae* EBY100 to generate pYD1αHAβ. Refer to Subheading 3.2 for detailed yeast transformation protocol. Isolate pYD1αHAβ from a 3 mL yeast culture in SD-CAA using the Zymoprep II yeast plasmid kit.
10. Transform 5 μL of the yeast plasmid into chemical competent DH5α cells (*see* **Note 7**), and isolate the plasmid for DNA sequencing as described in **steps 7** and **8** to confirm the sequence of pYD1αHAβ.
11. Construct plasmid pYD1αSTFβ that contains a segment of stuffer DNA in place of the $HA_{306\text{-}318}$ peptide in pYD1αHAβ (*see* **Note 1**): The stuffer DNA (STF) is an unrelated gene (~1 kb) amplified from the phosphite dehydrogenase gene [22] using primers StfFor and StfRev (*see* Table 1). The PCR product was cloned into pYD1αHAβ via NotI and SpeI to generate plasmid pYD1αSTFβ.
12. To establish the correlation between the yeast surface expression level of the peptide-DR1 complexes and the peptide-binding affinity (*see* **Note 2**), plasmids pYD1α2467β, pYD1αM4β, pYD1αYAKβ, pYD1αCIIβ, pYD1αPKAβ, pYD1αLβ and pYD1αMBPβ were constructed in two steps. First, phosphorylate the oligonucleotides by Optikinase (USB, Cleveland, OH). Set up the phosphorylation reaction as follows: 2 μL of oligonucleotides (2.5 μM), 1× Optikinase buffer, 2.5 μL of dATP (10 mM), 1 μL of Optikinase and add ddH_2O to make a final volume of 25 μL. Incubate the mixture at 37 °C for 2.5–3.5 h and 65 °C for 15 min. Second, generate the double stranded DNA insert encoding the peptide by hybridization of the forward and reverse oligonucleotides (*see* Table 1). For pYD1αMBPβ, generate the insert using a self-annealing oligonucleotide MBPNotI. Ligate the resulting insert into pYD1αSTFβ digested by NotI and SpeI.

3.2 Yeast Transformation

1. Inoculate a single colony of *S. cerevisiae* EBY100 into 3 mL of YPAD medium and grow overnight in a shaker at 30 °C and 250 rpm (*see* **Note 8**).
2. Measure the OD_{600} of the overnight culture and inoculate the appropriate amount to 50 mL (this is enough for ten transformations, scale up or down proportionally for more or less transformations) of fresh YPAD medium to obtain an OD_{600} of 0.2.
3. Continue to grow the 50 mL of culture for approximately 4–5 h to obtain an OD_{600} of 0.8 (*see* **Note 9**).
4. Wash the cells once with 25 mL of sterile ddH_2O (3200 rcf, 5 min), once with 1 mL of 0.1 M LiAc (3350 rcf, 30 s), add 400 μL of 0.1 M LiAc and resuspend (it should result in a total of 500 μL of resuspended cells), and aliquot 50 μL into sterile Eppendorf tubes. Each tube is used for one transformation.
5. In the meantime, boil 1 mL of ssDNA stock solution in a water bath for 5 min and chill immediately on ice.
6. Spin down the 50 μL cells (3350 rcf, 15 s) and discard the supernatant.
7. Prepare the transformation mixture immediately as the following: Add 240 μL of 50 % PEG, 36 μL of 1 M LiAc, 50 μL of boiled ssDNA stock solution, 0.1–10 μg of plasmid DNA and add sterile ddH_2O to make a final volume of 360 μL (*see* **Note 10**). For homologous recombination, 5–20 ng of the digested vector and a 10–20-fold molar excess of the insert are used.
8. Mix the transformation mixture extensively by vigorous vortexing, incubate at 42 °C for 40 min, spin down the cells (3350 rcf, 30 s), and remove the transformation mixture carefully.
9. Resuspend the cell pellet in 1 mL ddH_2O gently with pipetting.
10. Plate 50–200 μL of the resuspension on a SD-CAA agar plate and incubate at 30 °C for 2–3 days until colonies appear.

3.3 Expression Analysis of pMHC Complexes on Yeast Surface

1. Grow *S. cerevisiae* EBY100 clones transformed with different plasmid constructs in 3 mL of SD-CAA at 30 °C until OD600 reaches ~5.
2. Wash the yeast cells twice (3350 rcf, 1 min) with 1 mL YPG medium.
3. To induce AGA2 fusion protein expression, resuspend the cells to an OD_{600} ~1.0 in 3 mL of YPG medium supplemented with 50 μg/mL kanamycin and incubate in a refrigerated rotary shaker at 20 °C and 250 rpm for ~48 h. The induced yeast cells could be stored at 4 °C up to 4 months for repeated analysis (*see* **Note 11**).

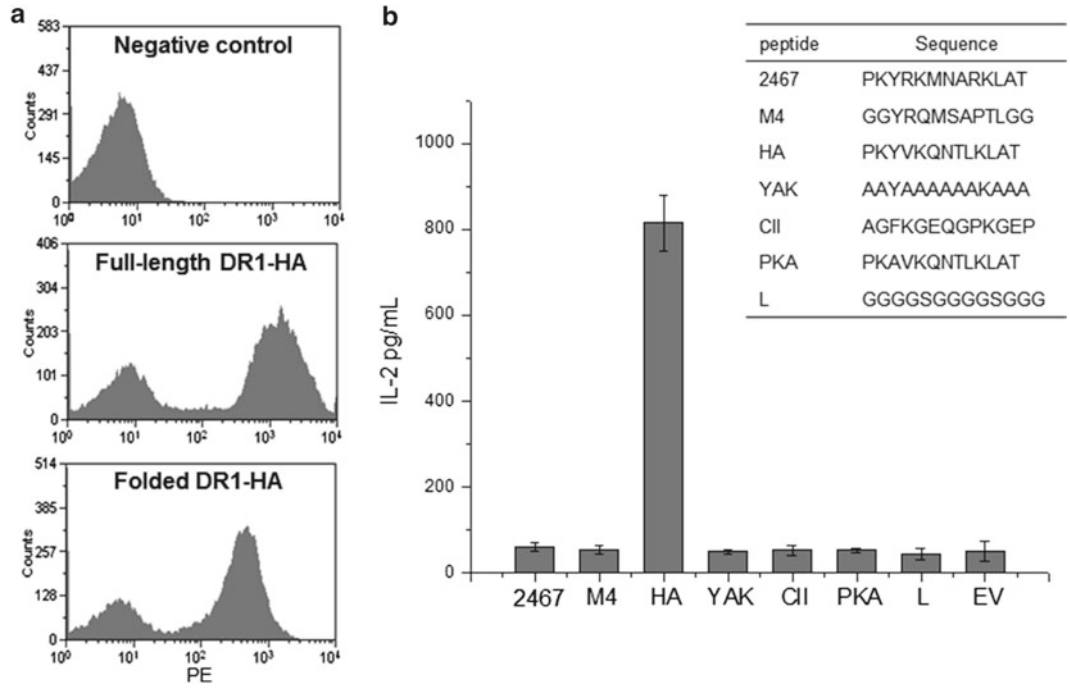


Fig. 2 Expression and function analysis of peptide-DR1 complexes displayed on yeast cell surface. (**a**) Yeast cells transformed with pYD1-scDR1αHAβ were stained with anti-V5 antibody to detect full-length protein expression, or with LB3.1 antibody to detect correctly folded complexes. Yeast cells stained only with a secondary antibody were used as a negative control. (**b**) Yeast cells displaying the peptide-DR1 complexes activated HA1.7 hybridoma in an antigen-specific manner. Yeast cells containing the empty plasmid pYD1 (EV) were cultured, induced, and analyzed in the same way as yeast displaying other constructs. Sequences of peptides used in the scDR1αpepβ constructs are shown in the *inset*

4. For flow cytometric analysis, wash ~2.5×10^6 cells in a 96-well V-bottom plate with 180 μL of PBS containing 0.5 % BSA and incubate with the primary antibody (anti-V5 for full-length protein detection and LB3.1 for correctly folded DR1 detection, *see* Fig. 2a) at 4 °C for 1 h at a 100-fold dilution rate.
5. Wash cells once with 180 μL of PBS + 0.5 % BSA and incubate with biotinylated GAM IgG (1:100 dilution) at 4 °C for 1 h.
6. Wash cells once with 180 μL of PBS + 0.5 % BSA and incubate with SA-PE (1:100 dilution) at 4 °C for 30 min.
7. Wash cells three times with PBS + 0.5 % BSA to remove unbound SA-PE.
8. Analyze fluorescently labeled yeast cells on a Coulter Epics XL flow cytometer at the Biotechnology Center of University of Illinois at Urbana-Champaign as per facility protocols (Urbana, IL) (*see* **Note 12**).

3.4 Stimulation of T Cell Hybridoma and IL-2 Detection

1. Aliquot ~10^6 yeast cells induced with YPG into 300 μL of PBS in a 96-well R-bottom tissue-culture plate. Prepare a triplicate for each pMHC construct.
2. Wash once with PBS and resuspend in 300 μL of PBS.
3. Incubate the plate at 4 °C overnight to allow the yeast cells to attach to the surface.
4. Wash away unbound yeast cells three times with 350 μL of PBS.
5. Wash HA1.7 hybridoma cells once with 10 mL of warm complete IMDM medium and resuspend to a density of 10^5 per 300 μL complete IMDM.
6. Add 300 μL of the HA1.7 hybridoma cell resuspension to each well and incubate for ~24 h at 37 °C with 5 % CO_2.
7. Spin down the cells and test the supernatant for IL-2 production using the murine IL-2 ELISA kit. IL-2 should only be detected in the supernatant from the wells where the yeast cells displaying functional peptide-MHCs specific for HA1.7 TCR, i.e., DR1-$HA_{306\text{-}318}$ (*see* Fig. 2b).

3.5 Peptide Library Construction from Hemagglutinin

1. Amplify DNA encoding hemagglutinin by one-step RT-PCR using influenza A/Aichi/2/68 viral RNA as the template and AichiFor and AichiRev as primers (Table 1). Set up the reaction mixture as follows: 2 μL of RNA, 1× reaction buffer, 0.2 μM of AichiFor, 0.2 μM of AichiRev, 1 μL of SuperscriptIII, and add ddH_2O to make up a final volume of 50 μL.
2. The thermocycler program: 55 °C for 30 min, an initial denaturation of 2 min at 94 °C, followed by 40 cycles of denaturation at 94 °C for 30 s, annealing at 55 °C for 30 s, and extension at 68 °C for 2 min, and a final 10-min elongation at 68 °C.
3. Gel-purify the DNA band with correct size using the QIAquick Gel Extraction Kit and quantify the concentration using NanoDrop, as per the manufacturer's instructions.
4. Fragmentize the purified DNA using DNase I as the following: mix 1 μg of DNA, 5 μL of 500 mM Tris–HCl, pH 7.4, 2.5 μL of 200 mM MnCl2, and ddH_2O to a total volume of 42.5 μL. Incubate the mixture at 15 °C for 15 min. Freshly dilute DNase I with ddH_2O to 0.01 U/μL and add 7.5 μL to the DNA mixture and incubate at 15 °C for 1 min. Heat inactivate at 90 °C for 10 min (*see* **Note 13**).
5. Run the DNase I digestion mixture immediately on a 2 % agarose gel and purify the DNA fragments with size between 30 and 60 bp using the QIAEX II Gel Extraction Kit, as per manufacturer's instructions.

6. Blunt-end polish the purified DNA fragments as follows: Mix 0.5 μL of 10 mg/mL BSA, 1× NEBuffer 2, 5 μL of 1 mM dNTP, 100 ng of DNA fragments, 0.1 U of T4 DNA polymerase and add ddH_2O to a final volume of 50 μL. Incubate at 25 °C for 10 min and add 1.25 μL of Klenow fragment (5 U/μL). Incubate at 25 °C for 10 min and 16 °C for 105 min.
7. Gel-purify the blunt-end polished DNA fragments using the QIAEX II Gel Extraction Kit.
8. In the meantime, blunt-end polish pYD1αSTFβ digested by SpeI and NotI in the same way as described in **step 5**, gel-purify using QIAquick Gel Extraction Kit, and dephosphorylate using 1 μL of SAP per 1 μg of DNA at 37 °C for 90 min. Heat inactivate at 70 °C for 10 min and clean up the mixture using the QIAquick PCR Purification Kit.
9. Ligate the DNA fragments from **step 6** into the vector from **step 7** as follows (*see* **Note 14**): Mix 1 μL of T4 DNA ligase buffer, 50 ng of vector, twofold molar excess of DNA fragments, 1.5 μL of PEG 8000, 1 μL of T4 DNA ligase, and add ddH_2O to a final volume of 10 μL. Incubate the ligation mixture at 16 °C for 16–20 h.
10. Clean up the ligation mixture as follows: add *n*-butanol to make a final volume of 500 μL, mix vigorously by vortexing for 30 s, centrifuge at the maximum speed in a benchtop centrifuge for 10 min, immediately remove the supernatant as much as possible, place the tube in a chemical hood for~1 h to dry, and resuspend the ligation mixture in 2 μL of ddH_2O (this is enough for one electroporation).
11. Transform the ligation mixture into ElectroMax DH5α competent cells by electroporation using Gene Pulser II.
12. After electroporation, immediately recover cells in 1 mL of pre-warmed SOC medium at 37 °C with 250 rpm agitation for 1 h.
13. For library creation, perform multiple electroporations, pool the cells into 400 mL LB medium supplemented with 100 μg/mL ampicillin, and grow overnight at 37 °C with 250 rpm agitation. For example, a library of 4.6×10^5 clones was obtained with ten electroporations in our study (*see* **Note 15**).
14. Isolate the plasmids using the QIAprep Miniprep Kit, as per manufacturer's instructions.
15. Perform plasmid transformations into the EBY100 competent yeast cells as described in Subheading 3.2 and pool the cells into 400 mL of SD-CAA medium. With 40 transformations, we obtained a library of 2.1×10^6.
16. After two passages in 400 mL of SD-CAA medium, induce protein expression using YPG as described in Subheading 3.3, and continue with FACS screening.

3.6 Peptide Library Construction from Influenza A Virus

1. Synthesize first strand cDNA using Transcriptor First Strand cDNA Synthesis Kit with random priming. Set up the reaction as follows: 2 μL of X31/A/Aichi/68 genomic RNA, 2 μL of hexamer, and 9 μL of PCR grade H_2O. Incubate the mixture at 65 °C for 10 min and put on ice immediately. Add 0.5 μL of PCR grade H_2O, 4 μL reaction buffer, 2 μL of dNTP, and 0.5 μL of Transcriptor RTase. Incubate the mixture at 25 °C for 10 min, 55 °C for 30 min, 85 °C for 5 min, and then cool down to 4 °C.
2. Synthesize the second strand DNA by incubating the cDNA with 10 units of DNA polymerase I, 0.32 units of RNase H, 0.25 mM dNTP, and 1× NEBuffer 2 in a final volume of 40 μL at 15 °C for 90 min.
3. Run the double stranded DNA on a 1 % agarose gel and purify the bands with different size ranges separately: <500 bp, 500 bp to 1 kb, and >1 kb.
4. Quantify the purified DNA using the Precision Molecular Mass Standard as per the manufacturer's instructions.
5. Fragmentize the DNA with DNase I using the same method as described in **step 4** of Subheading 3.6 except for approximately 6, 0.75, 1.7 units per 1 μg of DNA with size <500 bp, 500 bp to 1 kb, and >1 kb, respectively (*see* **Note 13**).
6. Generate the yeast display library for FACS screening using the same method as described in **steps 5–16** of Subheading 3.6.

3.7 Yeast Display Library Screening

1. Take an aliquot of the of induced peptide library culture (e.g., 7.4×10^6 yeast cells for the hemagglutinin library and 10^8 for the influenza A virus library) and stain with LB3.1 antibody followed by biotinylated GAM IgG and SA-PE as described in Subheading 3.3.
2. Sort the cells on a Coulter 753 bench FACS sorter (Flow Cytometry Facility, University of Illinois at Urbana-Champaign) and collect ~1.5 % of the population with the highest fluorescence (*see* Figs. 3a and 4a) in SD-CAA medium.
3. After protein expression induction in YPG, perform another two rounds of cell sorting (*see* **Note 16**) in the same way as described in **steps 1–2** except that the top 0.5 % of the population is collected into SD-CAA medium.
4. Induce protein expression in YPG and sort the top 0.5 % of the population into SD-CAA medium in a 96-well plate with no more than one cell in a well.
5. For hemagglutinin-derived peptide library, randomly pick ten of these cells and examine surface protein display level and their ability to stimulate T cell hybridoma HA1.7 as described in Subheadings 3.3 and 3.4 (*see* Fig. 3b, c). To determine the sequence of the peptide insert, perform DNA sequencing analysis.

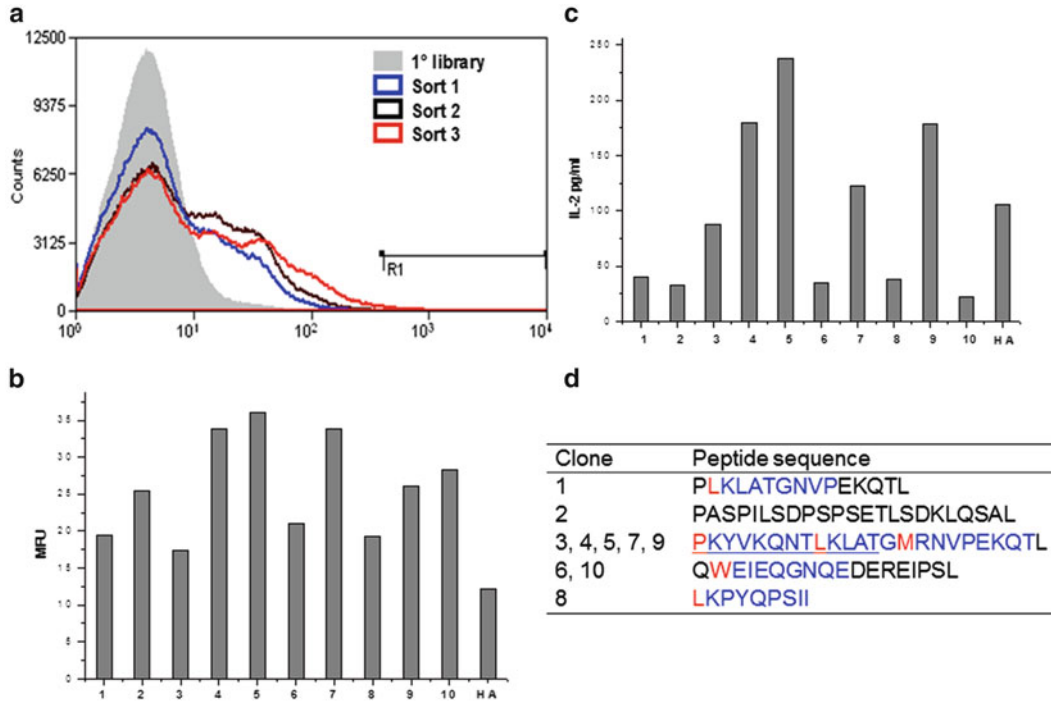


Fig. 3 Direct epitope identification from a single antigenic protein with known sequence—hemagglutinin. (**a**) FACS enrichment of potential good binders from the hemagglutinin-derived peptide library. LB3.1 antibody was used to stain cells as a measurement of surface expression level of correctly folded peptide-DR1 complexes. Surface expression (**b**), T cell hybridoma activation (**c**), and DNA sequence (**d**) analysis of ten clones randomly picked from the library after three rounds of cell sorting. The predicted binding 9-residue peptides are shown in blue with red letters corresponding to the amino acid residue at position P1

6. For influenza A virus-derived peptide library, sort four 96-well plates of cells from the top 0.5 % of the population and analyze them in the same way as described in **step 5** (*see* Fig. 4b, c).
7. Align peptide sequences from the active clones to identify epitope sequences using ClustalW (http://embnet.vital-it.ch/software/ClustalW.html) (*see* Figs. 3d and 4d).

4 Notes

1. To eliminate the possibility that the epitope $HA_{306\text{-}318}$ identified from the peptide library is derived from the undigested pYD1αHAβ, it is important to use plasmid pYD1αSTFβ that contains a segment of unrelated stuffer DNA in place of the $HA_{306\text{-}318}$ peptide as the backbone for blunt end ligation. In addition, the 1 kb stuffer DNA enables clear separation of digested plasmid by DNA electrophoresis.

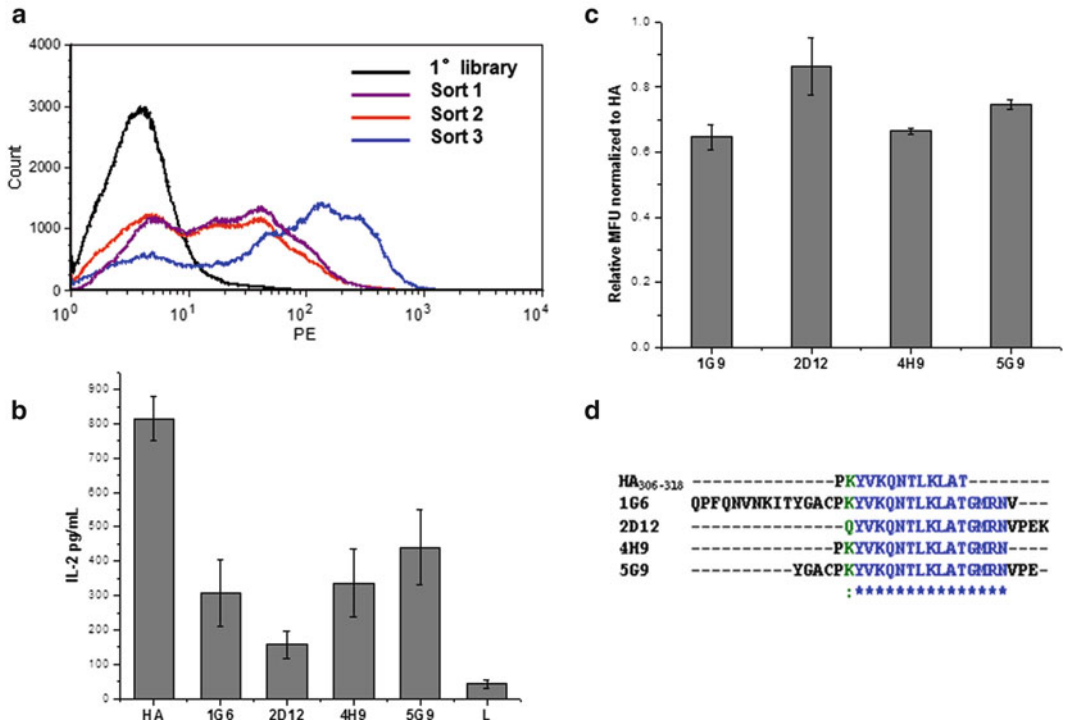


Fig. 4 Direct epitope identification from a pathogen with unknown sequence—influenza A virus. (**a**) FACS enrichment of potential DR1-binding peptides using LB3.1. (**b**) T-cell activation analysis of the four clones identified from the enriched peptide library. (**c**) The four active clones showed comparable surface expression levels as yeast displaying scDR1αHAβ. The relative mean fluorescence unit (MFU) was normalized to HA for direct comparison. (**d**) The peptide sequences from the four clones were aligned with the HA306-318 epitope sequence. *Asterisk* = fully conserved residues, *colon* = conservation of strong groups

2. A positive correlation between the yeast surface expression level of the peptide-DR1 complexes and the peptide-binding affinity was established [20]. Therefore, the yeast surface expression level of the single chain DR1-peptide protein, as measured by DR-specific antibody (LB3.1) staining, can be used as a proxy screening variable for DR1-binding peptides, significantly reducing the number of clones required in the function screening assay. This stabilizing effect of a binding peptide was also observed for DR4 [20] and DR2 [23].
3. The ability of yeast cells displaying DR1-$HA_{306\text{-}318}$ to activate T hybridoma cell HA1.7 greatly simplified the functional screening for epitope identification. This could be generally applicable to MHCII alleles, such as DR4 [20]. However, due to the polygenic and polymorphic nature of MHC proteins and the diversity of the binding peptides, it is difficult to establish that yeast cells could display all or most of pMHC complexes in a functional form to engage specific TCRs, such as low affinity self-reactive ones (e.g., DR2-$MBP_{85\text{-}99}$-Ob.1A12 [23]).

4. It is not desirable to freeze–thaw the carrier ssDNA frequently and usually it is discarded after 3–4 times of thawing.
5. When the DNA concentration is low (below 20 ng/μL), run 3–5 μL on a 0.7 % agarose gel and use the Precision Molecular Mass Standard for a more accurate estimation according to the manufacturer's instructions.
6. DH5α was used for DNA cloning in our experiments. However, any *E. coli* strain with recA and endA, such as Top10 and JM109, can also be used.
7. The number of DH5α transformants could vary from a few to several thousands. This is mainly due to the low quality of the isolated yeast plasmids. Sometimes, no colony was even obtained. In such cases, repeat the DH5α transformation.
8. Yeast competent cells need to be freshly prepared each time.
9. Normally, the doubling time for a *S. cerevisiae* laboratory strain is approximately 2 h.
10. It is important to follow the order by which the transformation mixture components are added, especially for 50 % PEG, which shield the competent cells from the toxic 1 M LiAc.
11. The storage time might vary depending on the stability of the target protein displayed on yeast cell surface. For the pMHC complexes in our study, no significant degradation was observed within 4 months.
12. To compare surface expression levels of different pMHC complexes, it is important to induce the protein expression, and perform the antibody staining in a single experiment run to minimize variation.
13. Since DNase I has very high activity, it is necessary to test and optimize the digestion conditions for different target DNA preparations and for different lots of DNase I.
14. The difficulties in cloning small pieces of DNA include: (a) optimization of the DNase I digestion step to have majority of the digested DNA fragments with a desired size range; (b) intramolecular circularization, which results in a high ligation background; (c) occurrence of multiple inserts, which results in splicing peptides that are not present in the natural peptide repertoire. Accordingly, when constructing the peptide library, it is important to (a) optimize the DNase I concentration and the digestion time; (b) dephosphorylate the vector and include 15 % PEG in the ligation mixture to promote intermolecular ligation over intramolecular circularization [24]; and (c) optimize the insert to vector molar ratio.
15. It is important to sample the diversity of the primary peptide library created in *E. coli*. To do so, an aliquot of the library is selected on a LB-Amp agar plate to obtain individual colonies

(this also allows an estimate of the library size) and 20 clones are randomly picked, sequenced, and mapped to the viral genome. There should be at least one peptide derived from each of the eight pieces of influenza viral genomic RNA displayed in either sense or antisense orientations, indicating a good representation of the whole viral genome.

16. The number of sorting cycles varies depending on how much improvement of the expression the enriched library has over the previous round. For example, there was a steady increase in the fluorescence intensity over the sorting rounds in Fig. 3a, so a total of four rounds of enrichment (the fourth round is indicated by region R1) were performed.

References

1. Germain RN (1994) MHC-dependent antigen processing and peptide presentation: providing ligands for T lymphocyte activation. Cell 76:287–299
2. Altman JD, Moss PA, Goulder PJ, Barouch DH, McHeyzer-Williams MG, Bell JI, McMichael AJ, Davis MM (1996) Phenotypic analysis of antigen-specific T lymphocytes. Science 274:94–96
3. Park H, Li Z, Yang XO, Chang SH, Nurieva R, Wang YH, Wang Y, Hood L, Zhu Z, Tian Q, Dong C (2005) A distinct lineage of CD4 T cells regulates tissue inflammation by producing interleukin 17. Nat Immunol 6: 1133–1141
4. Klein L, Kyewski B (2000) "Promiscuous" expression of tissue antigens in the thymus: a key to T-cell tolerance and autoimmunity? J Mol Med 78:483–494
5. Coulie PG, Karanikas V, Lurquin C, Colau D, Connerotte T, Hanagiri T, Van Pel A, Lucas S, Godelaine D, Lonchay C, Marchand M, Van Baren N, Boon T (2002) Cytolytic T-cell responses of cancer patients vaccinated with a MAGE antigen. Immunol Rev 188: 33–42
6. Huang CH, Peng S, He L, Tsai YC, Boyd DA, Hansen TH, Wu TC, Hung CF (2005) Cancer immunotherapy using a DNA vaccine encoding a single-chain trimer of MHC class I linked to an HPV-16 E6 immunodominant CTL epitope. Gene Ther 12:1180–1186
7. Jenkins MK, Khoruts A, Ingulli E, Mueller DL, McSorley SJ, Reinhardt RL, Itano A, Pape KA (2001) In vivo activation of antigen-specific CD4 T cells. Annu Rev Immunol 19:23–45
8. Henderson RA, Cox AL, Sakaguchi K, Appella E, Shabanowitz J, Hunt DF, Engelhard VH (1993) Direct identification of an endogenous peptide recognized by multiple HLA-A2.1-specific cytotoxic T cells. Proc Natl Acad Sci USA 90:10275–10279
9. Novak EJ, Liu AW, Gebe JA, Falk BA, Nepom GT, Koelle DM, Kwok WW (2001) Tetramer-guided epitope mapping: rapid identification and characterization of immunodominant CD4+ T cell epitopes from complex antigens. J Immunol 166:6665–6670
10. Babon JA, Cruz J, Orphin L, Pazoles P, Co MD, Ennis FA, Terajima M (2009) Genome-wide screening of human T-cell epitopes in influenza A virus reveals a broad spectrum of CD4(+) T-cell responses to internal proteins, hemagglutinins, and neuraminidases. Hum Immunol 70:711–721
11. Richards KA, Chaves FA, Sant AJ (2009) Infection of HLA-DR1 transgenic mice with a human isolate of influenza a virus (H1N1) primes a diverse CD4 T-cell repertoire that includes CD4 T cells with heterosubtypic cross-reactivity to avian (H5N1) influenza virus. J Virol 83:6566–6577
12. Bian H, Reidhaar-Olson JF, Hammer J (2003) The use of bioinformatics for identifying class II-restricted T-cell epitopes. Methods 29: 299–309
13. Sospedra M, Pinilla C, Martin R (2003) Use of combinatorial peptide libraries for T-cell epitope mapping. Methods 29:236–247
14. Hiemstra HS, Duinkerken G, Benckhuijsen WE, Amons R, de Vries RR, Roep BO, Drijfhout JW (1997) The identification of CD4+ T cell epitopes with dedicated synthetic peptide libraries. Proc Natl Acad Sci USA 94:10313–10318
15. Nino-Vasquez JJ, Allicotti G, Borras E, Wilson DB, Valmori D, Simon R, Martin R, Pinilla C

(2004) A powerful combination: the use of positional scanning libraries and biometrical analysis to identify cross-reactive T cell epitopes. Mol Immunol 40:1063–1074
16. Wang Y, Rubtsov A, Heiser R, White J, Crawford F, Marrack P, Kappler JW (2005) Using a baculovirus display library to identify MHC class I mimotopes. Proc Natl Acad Sci USA 102:2476–2481
17. Boen E, Crownover AR, McIlhaney M, Korman AJ, Bill J (2000) Identification of T cell ligands in a library of peptides covalently attached to HLA-DR4. J Immunol 165: 2040–2047
18. Wang RF, Wang X, Atwood AC, Topalian SL, Rosenberg SA (1999) Cloning genes encoding MHC class II-restricted antigens: mutated CDC27 as a tumor antigen. Science 284: 1351–1354
19. Koelle DM (2003) Expression cloning for the discovery of viral antigens and epitopes recognized by T cells. Methods 29:213–226
20. Wen F, Esteban O, Zhao H (2008) Rapid identification of CD4+ T-cell epitopes using yeast displaying pathogen-derived peptide library. J Immunol Methods 336:37–44
21. Horton RM, Cai ZL, Ho SN, Pease LR (1990) Gene splicing by overlap extension: tailor-made genes using the polymerase chain reaction. Biotechniques 8:528–535
22. Woodyer R, van der Donk WA, Zhao H (2003) Relaxing the nicotinamide cofactor specificity of phosphite dehydrogenase by rational design. Biochemistry 42:11604–11614
23. Wen F, Sethi DK, Wucherpfennig KW, Zhao H (2011) Cell surface display of functional human MHC class II proteins: yeast display versus insect cell display. Protein Eng Des Sel 24:701–709
24. Upcroft P, Healey A (1987) Rapid and efficient method for cloning of blunt-ended DNA fragments. Gene 51:69–75
25. Boder ET, Wittrup KD (1997) Yeast surface display for screening combinatorial polypeptide libraries. Nat Biotechnol 15:553–557

Chapter 16

Profiling of Cytokine and Chemokine Responses Using Multiplex Bead Array Technology

Greg Harris and Wangxue Chen

Abstract

Multiplex bead array technology expands upon the principles of the enzyme-linked immunosorbent assay by allowing the simultaneous quantification of a large number of cytokines and chemokines within a single sample. This allows for the researcher more freedom to investigate complex immune responses both *in vivo* and *in vitro*. Here we describe the detailed assay protocol and technical tips for simultaneous quantification of multiple cytokines and chemokines in mouse biological fluids such as sera, bronchoalveolar lavage fluid, tissue homogenate supernatant, and tissue culture supernatant, using a multiplex bead array assay.

Key words Multiplex assay, Cytokine, Chemokine, Bead array, Immunoassay, Luminex

1 Introduction

Multiplex bead array assays (MBAA) are a highly efficient alternative to traditional enzyme-linked immunosorbent assays (ELISA) for quantification of cytokines and chemokines in various types of biological samples [1, 2]. In a traditional ELISA, usually one cytokine is studied at a time and the number of possible analytes is limited by sample volume. Typically, 50–100 µL of sample at a working dilution is needed for a single cytokine ELISA; therefore, only a few ELISA assays can be carried out when the sample has an inherently small volume (such as mouse sera and human cerebrospinal fluid). MBAA allow for the simultaneous quantification of up to 100 analytes in the same sample and, generally, no more than 50 µL of sample is required for the entire panel.

Commercial MBAA reagents became first available in the late 1990s, developed by Luminex® (Austin, TX) under the name FlowMetrix™ [3–6], now known as xMAP® technology (Multi-Analyte Profiling) (*see* **Note 1**). Initially, the assays were run and analyzed on a standard flow cytometer equipped with specialized software. Luminex® later developed the Luminex® 100, a flow cytometer specifically equipped to acquire and analyze microspheres

Kelly M. Fulton and Susan M. Twine (eds.), *Immunoproteomics: Methods and Protocols*, Methods in Molecular Biology, vol. 1061,
DOI 10.1007/978-1-62703-589-7_16,

in a MBAA format. These units went into common use in the early 2000s [7, 8]. Since its inception, MBAA technology has been evaluated extensively [9–13] and has been cited in thousands of scientific publications [14]. MBAA has broad applications within the field of life science research, including assays for cell signalling molecules, cancer markers, metabolic markers, and cytokine and chemokine responses, as well as pathogen detection [2, 15–19]. Users have the option of customizing their own assay panels by purchasing beads and adding relevant capture and detection antibodies, or purchasing premade kits from commercial vendors who have partnered with Luminex® (Table 1).

Table 1
Common vendors of Luminex-compatible MBAA kits

Product line/company	Assay types	Species range	Web site
Bio-Plex/Bio-Rad (Hercules, CA)	Acute phase Angiogenesis Cancer Cell signalling Cytokine/chemokine Diabetes Isotyping	Human, mouse, and rat	www.bio-rad.com
Fluorokine/R&D Systems (Minneapolis, MN)	Angiogenesis Cardiovascular Cytokine/chemokine Kidney toxicity MMP Obesity	Human, mouse, and rat	www.rndsystems.com
Innogenetics (Belgium)	Alzheimer's	Human	www.innogenetics.com
Milliplex/Millipore (Billerica, MA)	Adipokine Bone metabolism Cancer Cardiovascular Cell signalling Cellular metabolism Cytokine/chemokine Endocrine Isotyping Skin Toxicity	Human, mouse, rat, nonhuman primate, dog, guinea pig, hamster, pig, and rabbit	www.millipore.com
Myriad RBM (Austin, TX)	Autoimmune and arthritis Bone disease Cancer Cardiovascular Cytokine/chemokine Diabetes and metabolic markers Endocrine Gastrointestinal	Human	www.myriadrbm.com

(continued)

Table 1
(continued)

Product line/company	Assay types	Species range	Web site
Novex/Life Technologies (Carlsbad, CA)	Cell signalling Cytokine/chemokine Transcription factors/nuclear receptors	Human, mouse, and rat	www.lifetechnologies.com
Origene (Rockville, MD)	Genotyping Transcription factors/nuclear receptors	Human	www.origene.com
Procarta/Affymetrix (Santa Clara, CA)	Cell signalling Cytokine/chemokine Gene expression profiling Transcription factors/nuclear receptors	Human, mouse, rat, nonhuman primate, dog, and pig	www.affymetrix.com
Tellgen (Shanghai)	Cancer	Human	www.tellgen.com
Widescreen/Merck Millipore (Billerica, MA)	Cancer Cardiovascular Cell signalling Endocrine Kidney toxicity MMP	Human, mouse, and rat	www.merckmillipore.com

The principle of the MBAA is similar to that of an ELISA [20], except that capture antibodies are covalently coupled with the surface of beads instead of the bottom of a 96-well plate (Fig. 1). Each well contains a mixture of up to 100 bead types, each type coupled with a different capture antibody. The bead conjugates are distinguished from one another by the fluorescence intensity ratio of two different fluorescent dyes embedded within the bead. Sample incubation is similar to ELISA, as well as the detection steps, which use biotinylated detection antibodies and a streptavidin–phycoerythrin (PE) conjugate reporter. Sample acquisition is performed using a specialized flow cytometer, which performs a series of analyses similar to flow cytometric analysis of cell targets. Individual beads are first gated for the proper size, then the bead type is determined based on the ratio of the two internal fluorescent dyes, and finally the fluorescence intensity of the reporter is determined for each individual bead. This last number correlates with the concentration of a given cytokine in solution, similar to the optical density readings in an ELISA assay. A standard curve is generated to calculate the final cytokine/chemokine concentration.

There are two different types of beads available for a MBAA: polystyrene and magnetic. The nonmagnetic, polystyrene beads are currently being phased out by most manufacturers. Magnetic

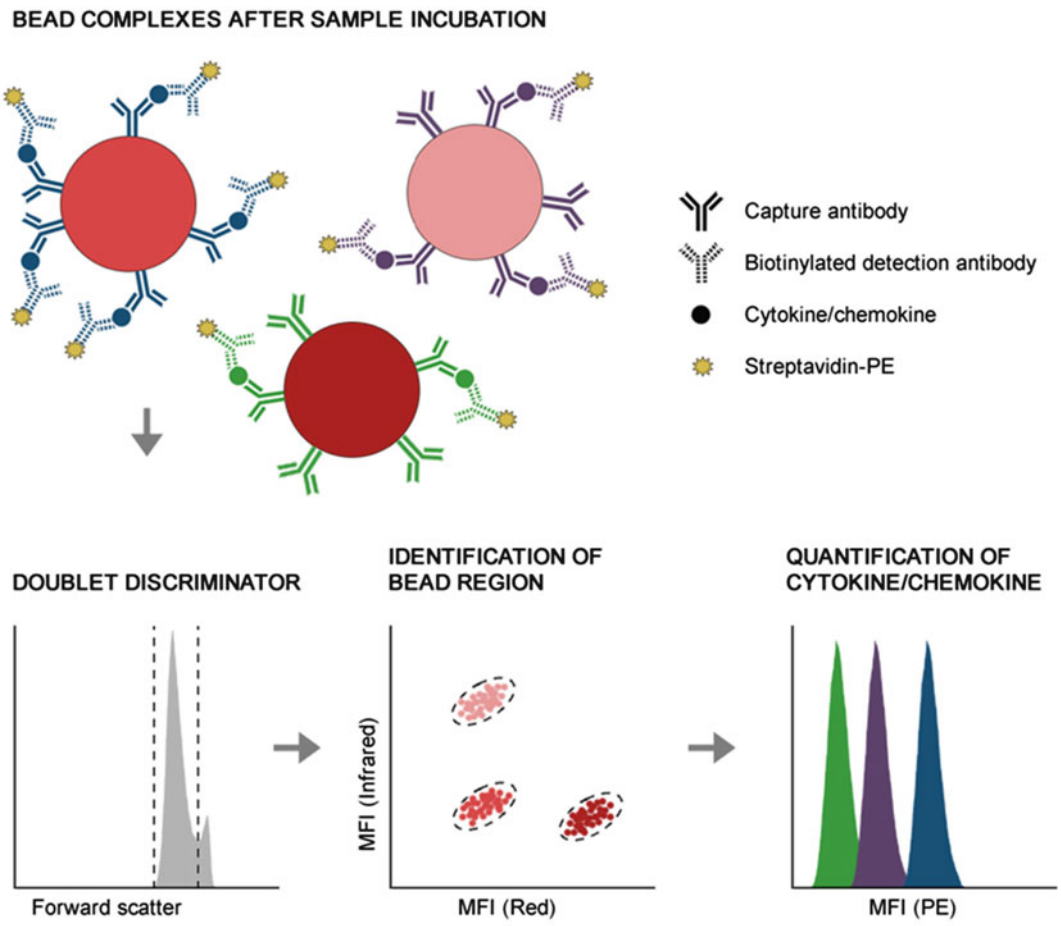


Fig. 1 The principle of multiplex bead array assay (MBAA). Monoclonal antibodies are covalently coupled to the surface of beads, so that each bead type has a unique type of monoclonal antibody attached to its surface. Beads are eventually distinguished from one another based on the relative amounts of red and infrared dyes embedded within. Beads are mixed together and the cocktail is added to each well of the assay, followed by incubation with samples, then with a biotinylated detection antibody and finally with streptavidin–PE as the reporter. Bead complexes are passed through the fluidics system of a specialized flow cytometer. Beads are gated on their size and doublets are screened out. Beads are then selected based on predetermined gates defined by the ratio of red and infrared dyes. Finally, a reporter laser and photomultiplier determine the fluorescence intensity of the reporter color (PE) for each bead. When extrapolated on a standard curve, the fluorescence intensity correlates directly to the concentration of the specified analyte (i.e., cytokine or chemokine) in the sample

beads are a newer technology, and are becoming the industry standard, as they employ a washing technique that is easier, less error-prone, and does not require costly filter plates like the traditional nonmagnetic bead-based platform.

Magnetic beads are also compatible with a newer type of analyzer, called MAGPIX®, which uses a magnetic plate to capture all beads simultaneously in a grid, LED lights to illuminate the

chamber and a CCD camera system to capture a digital image and determine the fluorescence intensity of each bead. MAGPIX is more compact and affordable than the Luminex100/200 system, but only allows for analysis of 50 analytes at the same time. Recently, Luminex has also developed a technology called FLEXMAP 3D®, which escalates the capabilities of Luminex 100/200 by combining three dyes within the microspheres, allowing for a multiplex number as high as 500.

MBAA reagents for common cytokines, chemokines and immunological signalling pathways are available in preassembled commercial kits, which are usually highly customizable for a wide range of analytes. In addition, there is a vast array of non-immunological applications for MBAA analysis of various biomarkers implicated in human health (Table 1). Availability of analytes varies from one vendor to the next, but the majority of commercially available kits are intended for use on Luminex 100/200® analyzer or its equivalent so there is a cross-vendor conformity in the general principles of the assay, particularly in the beads used.

As with ELISA, the individual steps in the assay procedure are optimized for each commercial kit, based on the manufacturer's protocols. Furthermore, there are various options for the type of system used to acquire data, and for the type of software used to analyze data. Here we outline the steps of a MBAA using a Milliplex®-brand Mouse Cytokine/Chemokine kit (Millipore, Billerica MA), which is compatible with a Luminex 100/200 analyzer for data acquisition.

2 Materials

1. Laboratory rocker or shaker (e.g., Lab-Line Instruments model #4625).
2. Aluminum foil.
3. Vortex mixer.
4. Sonicator (e.g., Branson Ultrasonic Cleaner model #B200).
5. (a) For nonmagnetic beads: vacuum filtration unit (e.g., Millipore Vacuum Manifold cat # MSVMHTS00). (b) For magnetic beads: handheld magnetic separation block (e.g., Millipore cat # 40-285), or automatic plate washer for magnetic beads (e.g., Bio-Tek ELx405, Millipore cat #40-015).
6. Luminex 100/200 analyzer (*see* **Note 2**).
7. Analysis software (e.g., Luminex xPONENT® software v4.2 or Millipore Analyst™ software v3.5.5.0).
8. Milliplex Mouse Cytokine/Chemokine custom kit (Millipore).
9. Calibration/verification beads (Luminex).

10. Phosphate-buffered saline (PBS).
11. 70 % ethanol.
12. Deionized water.
13. Sheath fluid (Luminex).

3 Methods

All steps are to be carried out at room temperature (20–25 °C) unless otherwise stated. Ensure that all procedures after **step 4** are carried out under low light conditions to preserve fluorescence intensity of beads and streptavidin–PE.

3.1 Set-Up and Incubation

1. Acclimate kit components to room temperature before use (*see* **Note 3**).
2. Pre soak plate with 200 µL wash buffer, and place on a plate shaker or rotator at a medium-high speed (*see* **Note 4**) for 10 min.
3. Dilute lyophilized standards and quality control samples with 250 µL deionized H_2O, mix gently, leave at room temperature for at least 5 min, then transfer to polypropylene microcentrifuge tubes.
4. Prepare standard curve dilutions in polypropylene microcentrifuge tubes, as outlined in the kit instructions.
5. Mix beads if necessary (*see* **Note 5**) by adding 60 µL of each bead type to the mixing bottle provided with the kit, then adding the appropriate volume of assay buffer for a total volume of 3 mL. Sonicate cocktail for 30 s in an ultrasonic cleaner and vortex at high speed for 60 s (*see* **Note 6**). Unused beads can be stored at 4 °C for up to 1 month (*see* **Note 7**).
6. Remove wash buffer from plate by either vacuum aspiration or by flicking off, depending on the type of kit being used (magnetic or nonmagnetic, *see* **Note 8**), and add 25 µL of sample diluent or standard diluent to appropriate wells (*see* **Note 9**).
7. Add 25 µL of standards, controls and samples (*see* **Note 10**) to appropriate wells, in duplicate.
8. Add 25 µL of bead cocktail to all wells.
9. Seal plate with a plastic plate sealer (*see* **Note 11**), protect plate from light by wrapping or covering with aluminum foil, and secure on plate shaker or rocker. Shake at a medium-high speed for 2 h at room temperature, or overnight at 4 °C (*see* **Note 12**).
10. Remove well contents by vacuum aspiration or by flicking off and wash twice with 200 µL of wash buffer (*see* **Note 13**).

11. Add 25 µL detection antibody cocktail (provided premixed in the kit). Reseal plate, protect from light and secure on plate shaker or rocker. Shake at a medium-high speed for 1 h at room temperature (*see* **Note 14**).
12. Add 25 µL streptavidin–PE conjugate to all wells. Do NOT perform a wash step at this point. Reseal plate, protect from light and secure on plate shaker or rocker. Shake at a medium-high speed for 30 min at room temperature.
13. Wash the wells 2× as in **step 10**.
14. Resuspend beads in 150 µL of PBS, reseal plate, protect from light and secure on plate shaker or rocker. Shake at a medium-high speed for 5 min at room temperature, then acquire (*see* **Note 15**).

3.2 Acquisition

Setup and operation of the Luminex 100/200 analyzer is similar to that of a flow cytometer. Please refer the user manual provided with the analyzer for step-by-step detailed instructions. The following is a basic overview of the steps involved and the estimated time required.

1. System setup (30 min): Warming up lasers, priming fluidics with sheath fluid, cleaning with 70 % ethanol and washing with PBS (cleaning and priming can be performed while lasers are warming up).
2. Assay settings and plate layout (10–20 min): Each type of commercial kit has its own optimal settings (e.g., number of beads to be acquired, bead size gates, sample volume collected) which need to be inputted by the user. Also, bead ID data (i.e., which cytokine/chemokine is associated with which bead type) must entered by the user. Finally, plate information is entered, including sample IDs, and information about standard curve dilutions and controls.
3. Calibration/verification (5 min): The analyzer requires calibration and verification every time is it turned on, as well as once a month if left on during that time. This involves running two sets of calibration beads, one set to calibrate the doublet discriminator and classification channels (which determine bead size and region, respectively) and the other set to calibrate the reporter channels (which determine fluorescence intensity of PE). Control beads are then run in order to verify calibration.
4. Data acquisition (45–90 min): The plate is loaded into the analyzer and once the data acquisition starts, the entire process is automated (*see* **Note 16**). Total acquisition time can vary from one assay to the next, but generally takes no more than 1 min per well (*see* **Note 17**).

5. System shut down (15 min): The fluidics systems are flushed with decontaminating solution (70 % ethanol), and flushed thoroughly with deionized water to remove any remaining salts. The system is then shut down, or left on if it will be used in the near future (the lasers power down automatically if the system is left idle for a few hours).

3.3 Analysis

If using Luminex xPONENT software for analysis, experiment-specific settings (sample IDs, standard curve concentrations, controls) are entered into the system before sample acquisition begins. Standard curves and sample concentrations are calculated at the end of the acquisition step and are exported as a portable document file (.pdf).

If using a second party software, such as Milliplex Analyst, raw data pertaining to bead counts and fluorescence intensity is imported as a comma-separated values (.csv) file, and experiment-specific settings (samples IDs, standard curve concentrations, controls) are applied to the imported values. Standards and samples are then analyzed and results displayed in a table format, which can be exported to standard spreadsheet software (*see* **Note 18**). *See* Fig. 2 for examples of standard curves produced by the analysis software, and Fig. 3 for a set of sample data for quantification of six cytokines/chemokines in mouse sera.

4 Notes

1. The BD Cytometric Bead Array (CBA) is a similar technology available through BD Biosciences and utilizes the BD flow cytometry instruments. However, Luminex reagents are more widely used than BD CBA, and the associated product line covers a broader spectrum of both life sciences research areas and species range. This chapter will focus solely on Luminex technology.
2. The Luminex 100 is an older model than the Luminex 200 and is not equipped to run magnetic beads. Because magnetic beads are being phased in as the new industry standard, the Luminex 200 is recommended. A more cost-effective alternative would be the MagPix which can assay up to 50 analytes simultaneously.
3. To speed up this step, float reagents in a room temperature water bath for 10 min.
4. The ideal speed is one where maximal shaking occurs without splashing of well contents. It is recommended that the user determines the optimal speed for their shaker or rotator on a blank plate with 100 µL/well of water before beginning the assay. While a rotator is generally preferred, we use a plate

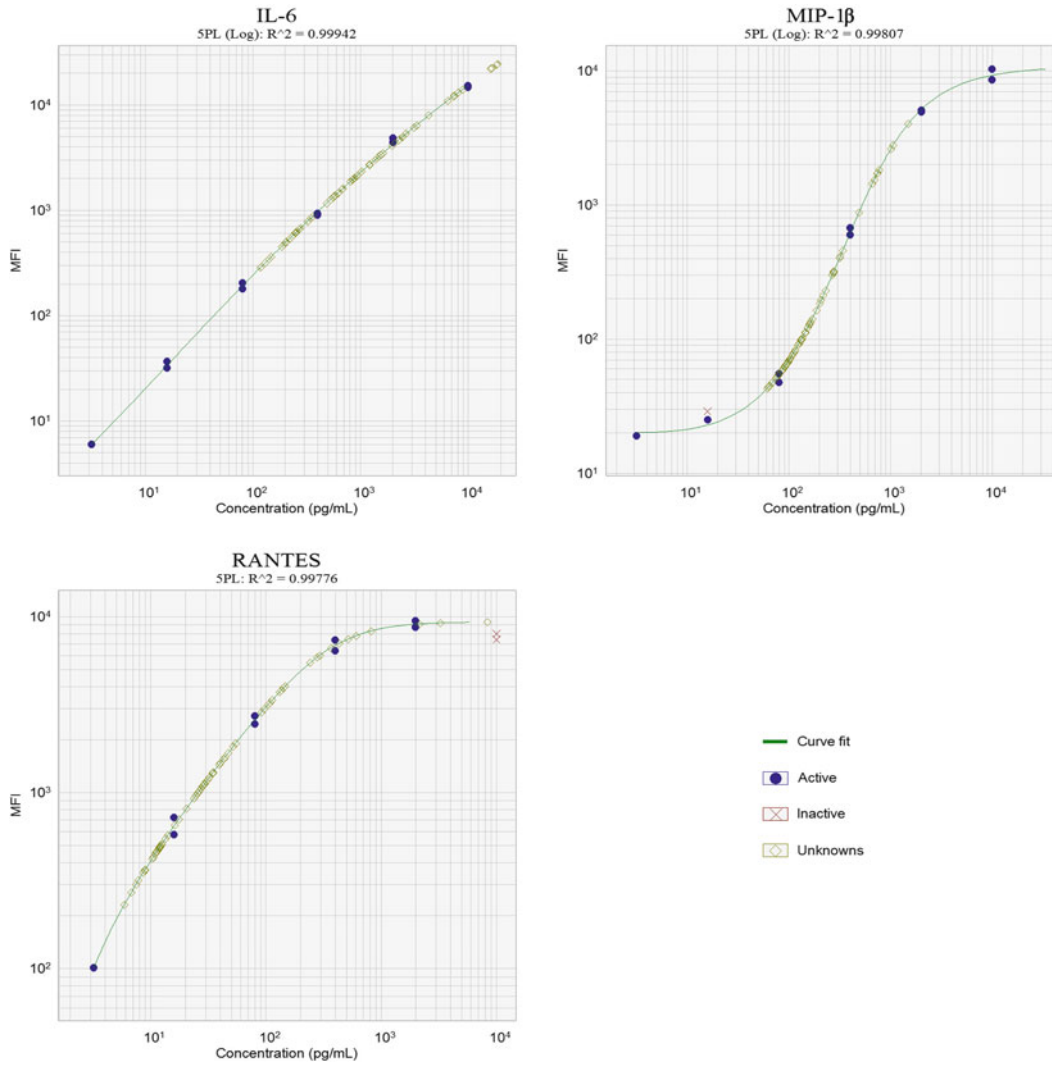


Fig. 2 Sample standard curves from a MBAA. Standards curves were generated using a Milliplex custom mouse cytokine/chemokine panel, and mouse serum samples. Grouped *blue dots* indicate duplicate standard dilutions, open diamonds indicate samples as extrapolated onto the standard curve, and *red x*'s indicate standard curve dilutions which were omitted from curve calculation, due to low fluorescence intensity or variation between duplicates. Curves were generated using Beadview® software (v.1.03, Upstate)

rocker at 120 rpm without complications. Be sure to secure the plate with tape or elastic bands.

5. Omit this step if a premixed bead cocktail is supplied with the kit although custom-designed kits generally require the preparation of bead cocktail by the end-user.
6. Polystyrene beads can stick together after long-term storage. While the analyzer is programmed to gate out any doublets (i.e., made up of two different bead types, which would give a

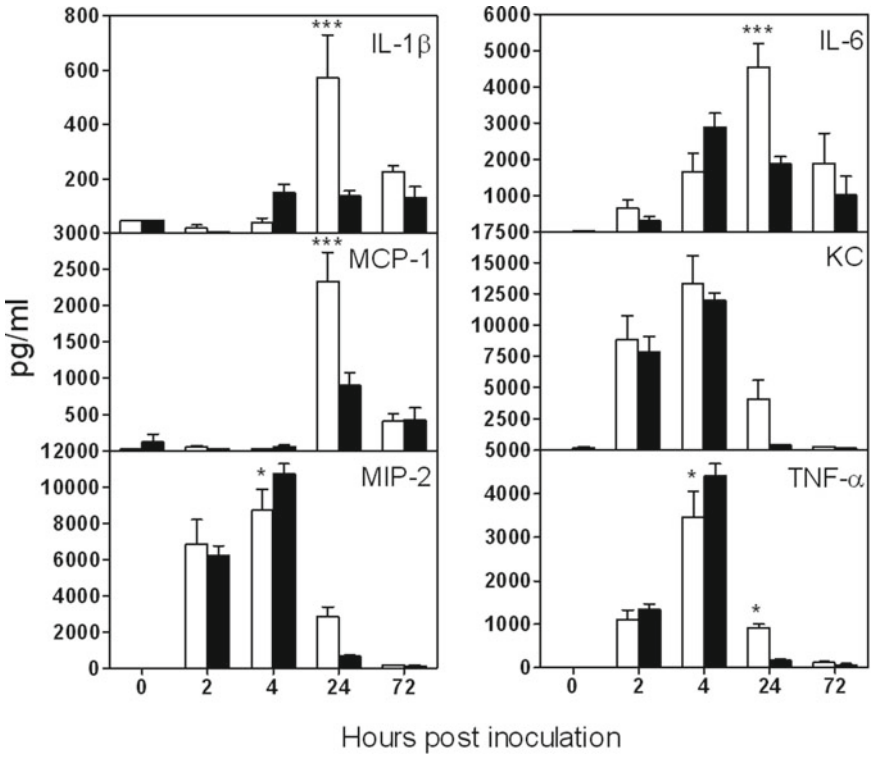


Fig. 3 Cytokine and chemokine levels in bronchoalveolar lavage fluid of A/J and C57BL/6 mice following i.n. inoculation with *A. baumannii.* Groups of A/J or C57BL/6 mice were i.n. inoculated with 2.5×10^7 CFU of *A. baumannii*. BAL samples were collected at 0, 2, 4, 24 and 72 h, and cytokine and chemokine levels were determined using the mouse panel of Fluorokine MAP Multiplex Kits (R&D Systems, Inc. Minneapolis, MN) on a Luminex 100 IS instrument. Data are expressed as mean ± SEM of five mice at each time point. The detection limits of the assays were 2.5–15 pg/mL. *$P < 0.05$ and ***$P < 0.005$ vs. C57BL/6 mice. From ref. 15

false reading), sonication helps to minimize the total amount of doublets, thereby increasing the total bead count and the assay accuracy.

7. Alternatively, the user can calculate the exact number of beads needed for the assay and only mix those into the cocktail. This is recommended when only a small number of samples are being assayed.

8. For nonmagnetic kits: use a vacuum manifold designed for this purpose. *See* **Note 13** for liquid removal instructions. After gentle removal of wash buffer, blot bottom of plate on paper towel to remove any droplets left behind (otherwise, these droplets can act as a wick and draw out further liquid once the wells are filled, leading to a reduced total volume). Never leave filter plates resting on paper towel or other absorbent material, as this can exacerbate the wicking effect.

   For magnetic kits, wash buffer can be removed by flicking plate contents into a sink (this step only—for future washes, a magnetic washer must be used, *see* **Note 13**).

9. Samples are diluted 1:2 in assay buffer, whereas standards and quality controls are diluted 1:2 in the buffer that was used to prepare the samples (i.e., serum matrix if serum samples are being assayed; bronchoalveolar lavage (BAL) fluid if BAL samples are being assayed; PBS or saline if lung homogenate supernatants are being assayed; complete tissue culture medium if tissue culture supernatants are being assayed).
10. We have assayed mouse cytokine and chemokine levels in the serum, BAL fluids, the supernatants of tissue homogenates (lung, spleen, liver and skin) and tissue culture supernatants using the MBAA system. It is generally recommended to centrifuge the samples after thawing to remove precipitating materials before assay even if no precipitate is visible. The amount of precipitate varies from sample to sample, with larger amounts in lung homogenates, smaller amounts in spleen homogenates, and very small amounts in mouse serum, BAL fluid and tissue culture supernatants. In addition, the tissue homogenate pellets tend to be loose, especially with lung and liver samples. Therefore, remove supernatant gently without dislodging the pellet.
11. For nonmagnetic kits, the pressure introduced into the wells by sealing the plate will force a small amount of liquid out the bottom of the wells. This must be blotted off with a paper towel, or a wicking effect will further draw out liquid over time.
12. An overnight incubation is recommended by the manufacturer for a higher assay sensitivity.
13. For nonmagnetic kits, use a vacuum manifold designed for this purpose. Be careful not to apply too much vacuum pressure, as this can tear the membrane of the filter plate. Optimally, it should take 2–3 s to remove 200 µL of wash buffer. Place plate on vacuum manifold, press downward with one hand while slowly opening the vacuum valve with the other. If properly sealed, vacuum pressure should occur easily. If no pressure is observed, reposition plate and press down until suction is achieved. With certain sample types, especially sera, it is possible to get a few clogged wells on the plate. After each wash, ensure that all wells are emptied. If any wells are clogged, use the wide end of a 200 µL pipette tip to scrape against the valve on the underside of the well. This should dislodge any particulates that are clogging the well. Reapply vacuum pressure and ensure all wells have emptied. After washes are complete, blot plate bottom with paper towel. If only using a portion of the plate, seal unused wells and keep sealed throughout the assay. Exposed, dry wells will reduce the overall suction pressure, making washes more difficult. Alternatively, wet these wells and keep wet throughout the assay by washing during wash steps. However, once wells are wet, the empty plate portion cannot be reused.

For magnetic kits, use a handheld magnetic separation block. Affix the plate to the handheld separation block, and allow beads to settle at the bottom of wells for 60 s. Flick off liquid into a sink or similar receptacle. Typically this will involve three sharp downward motions to remove all the liquid and remaining droplets. Despite the fact that beads are magnetically bound to the plate bottom, do not tap the plate on a paper towel to remove excess liquid, as this can dislodge beads. Remove plate from separation block, add 200 μL of wash buffer, reaffix plate to separation block, and repeat wash steps for the appropriate number of times. Magnetic beads can also be washed with an automatic plate washer for magnetic beads, and are compatible with the vacuum manifold used for nonmagnetic beads (however, the filter plate required for the vacuum manifold does not come packed with the magnetic beads kit, and must be purchased separately).

14. This incubation step is a convenient time to set up the analyzer, as it requires a 30 min warm-up time for the laser, as well as approximately 10–20 min for plate template setup.
15. If immediate acquisition is not an option, plates can be stored at 4 °C and read at a later time, after a period of shaking to resuspend the beads. Milliplex manufacturers recommend acquisition no more than 24 h later; however, we have read plates 72 h after incubation, with no complications or loss of fluorescence intensity.
16. As with other flow cytometry systems, the acquisition software features real-time display of events as they occur, including progress bars for acquisition of each bead type, as well as the bead classification gates. During setup, the bead ID numbers are assigned to the software, and predetermined gates are used to gate out any beads that don't fall within any expected regions. It is recommended that the user watch the acquisition of the first set of samples, to ensure that all bead ID numbers are properly entered. If an error was made during bead ID number assignment, then the improper gate would be displayed, and the corresponding bead populations would fall outside of the gate (*see* Fig. 4). If this occurs, the acquisition must be shut down immediately, and the assay settings reentered. This may result in sample loss from the first well acquired, but the samples can still be rerun.
17. Typically, a minimum of 100 beads is acquired per bead type. However low bead counts are known to occur, and the system is set to "time out" if a certain period of time (typically 1 min) passes before the minimum number of beads is reached. In this case, the data is still useable despite the low bead count. We have had instances where only 20 beads were acquired, and the ensuing results were in line with the next well containing a duplicate sample where 100 beads were acquired.

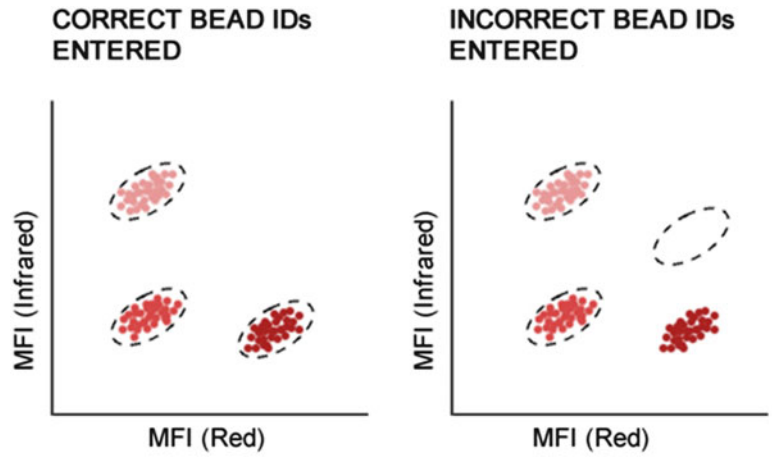


Fig. 4 Example of gate and bead locations from correctly and incorrectly entered bead ID numbers. Bead region gates (*dotted lines*) are preprogrammed into the software and are displayed based on the bead ID numbers entered by the user. *Colored spots* indicate all bead events, and are displayed in real time on the screen during sample acquisition. The empty gate in the *right hand panel* indicates that the incorrect bead region was entered by the user. Note that this is a simplified portrayal of the data-acquisition display screen

18. Before exporting data, the user should examine the standard curves generated by the software. The curves are typically sigmoidal, and each will have its own formula for a best fit, depending on binding behavior of each antibody type (best fit can be assigned by the software, or selected by the user). There are occasions where one or both of the standard curve duplicates falls outside of the calculated curve. *See* the RANTES standard curve in Fig. 2 as an example; in this instance, the top two points of the standard curve were removed, as their mean fluorescence intensity (MFI) registered lower than expected. The curve was recalculated using five points instead of six. If only one point needs to be removed (*see* MIP-1β standard curve, Fig. 2), the user should remove the point which results in a higher coefficient of determination (R^2) value when the curve is recalculated.

    Samples should also be checked for their location on the standard curve. Ideally, all samples would fall within the linear range of the curve but they can sometimes fall within the upper or lower plateau (*see* RANTES standard curve in Fig. 2). In the case of the upper plateau, a small difference in MFI between two duplicates can result in a very large difference in the calculated concentration. Therefore careful attention must be paid to the %CV between duplicates. Most analysis software will offer a display screen that shows the %CV of each set of replicates.

References

1. Khan SS, Smith MS, Reda D, Suffredini AF, McCoy JP Jr (2004) Multiplex bead array assays for detection of soluble cytokines: comparisons of sensitivity and quantitative values among kits from multiple manufacturers. Clin Cytom 61B:35–39
2. Elshal MF, McCoy JP Jr (2006) Multiplex bead array assays: performance evaluation and comparison of sensitivity to ELISA. Methods 38:317–323
3. Fulton RJ, McDade RL, Smith PL, Kienker LJ, Kettman JR Jr (1997) Advanced multiplexed analysis with the FlowMetrix™ system. Clin Chem 43:1749–1756
4. Keij JF, Steinkamp JA (1998) Flow cytometric characterization and classification of multiple dual-color fluorescent microspheres using fluorescence lifetime. Cytometry 33:318–323
5. Kettman J, Davies T, Chandler D, Oliver K, Fulton R (1998) Classification and properties of 64 multiplexed microsphere sets. Cytometry 33:234–243
6. Oliver KG, Kettman JR, Fulton RJ (1998) Multiplexed analysis of human cytokines by use of the FlowMetrix system. Clin Chem 44: 2057–2060
7. Bellisario R, Colinas RJ, Pass KA (2000) Simultaneous measurement of thyroxine and thryotropin from newborn dried blood spot specimens using a multiplexed fluorescent microsphere immunoassay. Clin Chem 46:1422–1424
8. Taylor JD, Briley D, Nguyen Q, Long K, Iannone MA, Li M-S et al (2001) Flow cytometric platform for high-throughput single nucleotide polymorphism analysis. Biotechniques 30:661–669
9. de Jager W, te Velthuis H, Prakken BJ, Kuis W, Rijkers GT (2003) Simultaneous detection of 15 human cytokines in a single sample of stimulated peripheral blood mononuclear cells. Clin Diagn Lab Immunol 10:133–139
10. de Jagger W, Prakken BJ, Bijlsma JW, Kuis W, Rijikers GT (2005) Improved multiplex immunoassay performance in human plasma and synovial fluid following removal of interfering heterophilic antibodies. J Immunol Methods 300:124–135
11. DuPont NC, Wang KH, Wadhwa PD, Culhane JF, Nelson EL (2005) Validation and comparison of Luminex multiplex cytokine analysis kits with ELISA: determinations of a panel of nine cytokines in clinical sample culture supernatants. J Reprod Immunol 66:175–191
12. Codorean E, Nichita C, Albulescu L, Raducan E, Popescu ID, Lonita AC et al (2010) Correlation of xMAP and ELISA cytokine profiles; development and validation for immunotoxicoligical studies in vitro. Roum Arch Microbiol Immunol 69:13–19
13. Richens JL, Urbanowiscz RA, Metcalf R, Corne J, O'Shea P, Fairclough L (2010) Quantitative validation and comparison of multiplex cytokine kits. J Biomol Screen 15:562–1572
14. Searchable bibliography at Luminex website http://www.luminexcorp.com/bibliography
15. Qiu H, KuoLee R, Harris G, Chen W (2009) High susceptibility to respiratory *Acinetobacter baumannii* infection in A/J mice is associated with a delay in early pulmonary recruitment of neutrophils. Microbes Infect 11:946–955
16. Liu M, Xydakis A, Hoogeveen R, Jones P, Smith E, Nelson K et al (2005) Multiplexed analysis of biomarkers related to obesity and the metabolic syndrome in human plasma, using the Luminex-100 system. Clin Chem 51:1102–1109
17. Landlinger C, Preuner S, Willinger B, Haberpursch B, Racil Z, Mayer J et al (2009) Species-specific identification of a wide range of clinically relevant fungal pathogens by use of Luminex xMAP technology. J Clin Microbiol 47:1063–1073
18. Joslin E, Opresko L, Wells A, Wiley H, Lauffenburger D (2007) EGF-receptor-mediated mammary epithelial cell migration is driven by sustained ERK signaling from autocrine stimulation. J Cell Sci 120:3688–3699
19. Scholler N, Crawford M, Sato A, Dreschner C, O'Briant K, Kiviat N et al (2006) Bead-based ELISA for validation of ovarian cancer early detection markers. Clin Cancer Res 12:2117–2124
20. Vignali DAA (2000) Multiplexed particle based flow cytometric assays. J Immunol Methods 243:243–255

Chapter 17

Preparation of the Low Molecular Weight Serum Proteome for Mass Spectrometry Analysis

Timothy J. Waybright, King C. Chan, Timothy D. Veenstra, and Zhen Xiao

Abstract

The discovery of viable biomarkers or indicators of disease states is complicated by the inherent complexity of the chosen biological specimen. Every sample, whether it is serum, plasma, urine, tissue, cells, or a host of others, contains thousands of large and small components, each interacting in multiple ways. The need to concentrate on a group of these components to narrow the focus on a potential biomarker candidate becomes, out of necessity, a priority, especially in the search for immune-related low molecular weight serum biomarkers. One such method in the field of proteomics is to divide the sample proteome into groups based on the size of the protein, analyze each group, and mine the data for statistically significant items. This chapter details a portion of this method, concentrating on a method for fractionating and analyzing the low molecular weight proteome of human serum.

Key words Low molecular weight proteome, Mass spectrometry, Strong cation exchange, Human serum, Liquid chromatography, Protein depletion

1 Introduction

The search for important insights into the mechanisms of diseases through the study of biological samples can be overwhelming due to the plethora of data generated and the ability to mine meaningful data from the results. Investigators must also contend with the problematic flux of normal and disease-related components within the test subjects. This challenge is exacerbated with very large cohorts, leading to phrases being used such as "looking for a needle in a haystack," or using a "shotgun approach." While these studies are definitely high risk, they are also potentially highly rewarding. Unfortunately, the recent era of biomarker discovery has been witness to too much failure and too little reward. The reasons for this imbalance are many and lead to the obvious question of how to design experiments to look at complex biological samples in a methodical, meaningful, and useful fashion that will produce validatible biomarkers to diagnose human diseases such as cancer.

Kelly M. Fulton and Susan M. Twine (eds.), *Immunoproteomics: Methods and Protocols*, Methods in Molecular Biology, vol. 1061, DOI 10.1007/978-1-62703-589-7_17,

One necessary step when analyzing blood samples for biomarkers is to reduce the overall complexity of the sample so that a specific group of molecules can be studied [1–4]. Serum (as well as plasma) is an excellent candidate for complexity reduction since it contains thousands of compounds of highly diversified chemotypes, chief among them being very high protein content [5, 6]. The most abundant proteins in serum comprise approximately 90–95 % of the total protein [7–9]. Low level proteins that may be potential candidates for disease biomarkers may not be detected because of this, necessitating a pre-analytical clean-up step to reduce the amount of the high-abundance proteins [10, 11].

In this study, the low molecular weight portion of human serum is analyzed to provide a characterization of that group of protein. While the term low molecular weight is a relative designation [5], an arbitrary cutoff was established at 30 kDa. The low molecular weight proteins in human serum have been linked with certain diseases involving functional components of the immune system such as cytokines, chemokines, and related signaling molecules [12], diabetes [13], cancer [14], and various cardiovascular ailments [15]. The procedure outlined in this chapter demonstrates that the removal of the high-abundant proteins in serum (such as albumin, haptoglobin, and transferrin) enhances the enrichment of proteins of low molecular weight and low abundance.

2 Materials

2.1 Serum Preparation and Filtration

1. Standard human serum (909b, human lyophilized serum, NIST, Gaithersburg, MD).
2. Centriplus centrifuge filters with a molecular weight cutoff (MWCO) of 30 kDa and collection tube.
3. Blue Max 15 mL polypropylene conical tube.
4. Centrifuge (Avanti J301, Beckman Coulter, Fullerton, CA) with a fixed-angle (JA30.50) rotor.
5. Speed vacuum centrifuge.

2.2 Protein Analysis Using Sodium Dodedcyl Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE) to Confirm Removal of Serum High Molecular Weight Proteins

1. Pre-cast 4–12 % Bis-Tris gel.
2. Pre-cast 10–20 % gradient tricine gel.
3. 20× 3-(*N*-morpholino)propanesulfonic acid-sodium dodecyl sulfate (MOPS SDS) running buffer.
4. Xcell SureLock electrophoresis cell.
5. PowerEase 500 high-voltage power supply.
6. Mark 12 unstained molecular weight (MW) standard.
7. 10× NuPAGE sample reducing agent.

8. 4× NuPAGE lithium dodecyl sulfate (LDS) sample buffer.
9. 2.0 mL Safe-lock polypropylene tubes.
10. 0.65 mL Safe-lock polypropylene tubes.
11. 25 mM ammonium bicarbonate (NH_4HCO_3) at pH 8.2
12. SimplyBlue safestain.

2.3 Enzymatic Digestion and Desalting

1. Porcine sequencing grade modified trypsin.
2. Incubator.
3. Speed vacuum centrifuge.
4. Bond-Elut C18 reversed-phase solid phase extraction (SPE) column.
5. Ultrapure water (double distilled, deionized >18 Ω, NANOPure Diamond water system).
6. 2.0 mL Safe-Lock Polypropylene tubes.
7. 1 M dithiothreitol (DTT).
8. 1 M iodoacetamine.
9. 50 mM ammonium bicarbonate (NH_4HCO_3).
10. 0.1 % (v/v) trifluoroacetic acid (TFA).
11. 60 % acetonitrile (CH_3CN) in water.
12. 60 % methanol (CH_3OH) in water.

2.4 Strong Cation Exchange (SCX) Fractionation

1. Polysulfoethyl A Ion Exchange column (150 mm × 1 mm, 300 Å, 5 µm particle, PolyLC, Inc., Columbia, MD).
2. 96-well polypropylene V-bottomed plates.
3. Speed vacuum centrifuge.
4. Model 1100 capillary liquid chromatography (LC) system (Agilent Technologies, Santa Clara, CA).
5. Fraction collector 96-well format (Model Foxy Jr., Teledyne ISCO Lincoln, NE).
6. Custom-built UV laser-induced fluorescence (LIF) detector: excitation/266 nm, emission filter/UG11. UV-LIF detector is also commercially available. A UV detector (280 nm) with a small flow cell may be used for the detection of high peptide content.
7. Solvent A: 25 % acetonitrile.
8. Solvent B: 25 % acetonitrile, 0.5 M ammonium formate, pH 3.0.
9. 0.1 % formic acid (FA), 25 % acetonitrile.
10. Ultrapure water (double distilled, deionized >18 Ω, NANOPure Diamond water system).

2.5 Microcapillary LC-MS/MS Analysis

1. Autosampler vials.
2. Branson Sonicator (Thomas Scientific, Swedesboro, NJ).
3. Ion Trap Mass Spectrometer (ThermoFinnigan, San Hose, CA).
4. Model 1100 capillary LC system (Agilent Technologies, Santa Clara, CA).
5. Fused-silica capillary tubing (Polymicro Technologies, Phoenix, AZ).
6. Jupiter C-18 column packing material (Phenomenex, Torrance, CA).

3 Methods

3.1 Serum Preparation and Filtration

1. Reconstitute the lyophilized sample (received from NIST in a sealed vial) to 10 mL with ultrapure double-distilled water (ddH_2O) as instructed in the kit (*see* **Note 1**).
2. Place a 30 kDa Centricon filter in the filtrate retention vial.
3. Add 2–3 mL of water to the filter.
4. Centrifuge the vial at 1,000 × *g* for 2 min (*see* **Note 2**). Repeat this step.
5. Discard the water and place the filter in a new filtrate retention vial.
6. Add 4 mL of 25 mM NH_4HCO_3/acetonitrile (4:1 v/v) to a 15 mL conical tube.
7. Transfer 1 mL of the reconstituted serum to the 15 mL conical tube above. Mix gently by inverting 5–10 times.
8. Transfer the diluted serum solution to the Centricon filter and centrifuge at 3,000 × *g* until approximately 90 % of the solution has passed through the filter membrane.
9. Lyophilize the filtrate to complete dryness under vacuum.

3.2 Protein Analysis Using SDS-PAGE

1. Resolubilize the lyophilized serum filtrate in 1 mL of 25 mM NH_4HCO_3.
2. Aliquot 20 µL of the solution to a 0.65 mL microcentrifuge tube.
3. Aliquot 7 µL of the NuPAGE LDS sample buffer into the tube.
4. Aliquot 3 µL of the NuPAGE sample reducing agent into the tube.
5. Gently mix the sample and place it into boiling water for 10 min. Remove the sample from the boiling water bath and allow it to cool to room temperature.

6. Remove the gel (either 4–12 % Bis-Tris or 10–20 % Gradient Tricine) gel from its package and rinse it with water. Remove the protective tape at the bottom of the gel and the lane divider/holder from the top of the gel (*see* **Note 3**).
7. Prepare a MOPS solution by adding 50 mL of the concentrated (20×) stock solution to 950 mL of water and mix gently.
8. Assemble the electrophoresis cell apparatus, as per the manufacturer's instructions, and add the diluted MOPS solution.
9. To one lane of the gel, add 10 µL of the Mark 12 Unstained MW Standard.
10. To another lane of the gel, add the serum mixture.
11. Cover the electrophoresis cell with the lid, attach the connectors to a power source, and run the gel at approximately 170 V for approximately 1 h (*see* **Note 4**).
12. When the dye front reaches the bottom of the gel, remove the gel, wash it in water, and stain with SimplyBlue Safestain, as per the manufacturer's instructions, to confirm the removal of high molecular weight proteins from the serum sample.

3.3 Enzymatic Digestion and Desalting

1. Boil the remaining resolubilized serum filtrate (i.e., approximately 1 mL) for 5 min and let it cool to room temperature.
2. Add the 1 M DTT (prepared fresh) to the 1 mL serum filtrate for a final DTT concentration of 10 mM (*see* **Note 5**).
3. Heat the samples for 1 h at 56 °C and let it cool to room temperature.
4. Add the 1 M iodoacetamide (prepared fresh) to the serum filtrate for a final iodoacetamide concentration of 10 mM (*see* **Note 6**).
5. Incubate the sample for 1 h at room temperature.
6. Prepare a fresh solution of sequencing grade trypsin by solubilizing it in 50 mM NH_4HCO_3, pH 8.4 (*see* **Note 7**).
7. Add the trypsin solution to the serum filtrate so that the enzyme:protein ratio is 1:50 (*see* **Note 8**).
8. Digest the serum filtrate for 16 h at 37 °C with gentle shaking.
9. Acidify the serum filtrate with TFA to a final concentration of 0.1 %.
10. Prepare an SPE cartridge by wetting the packing material with 2 mL of 60 % CH_3OH/H_2O and flushing it with 2 mL of 0.1 % (v/v) TFA.
11. Apply the digested serum sample to the cartridge and remove the salts by flushing with 2 mL of 0.1 % (v/v) TFA.

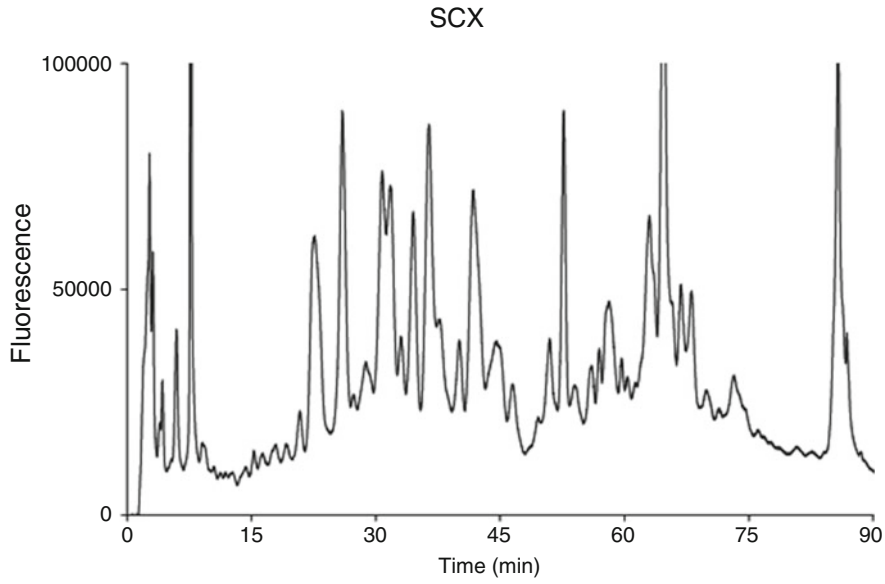


Fig. 1 Typical strong cation exchange chromatogram for the separation of low molecular weight serum tryptic peptides

12. Elute the digested peptides with 400 µL of 60 % acetonitrile/H_2O.
13. Lyophilize the samples at room temperature.

3.4 Strong Cation Exchange Fractionation

1. Prepare solvent A (25 % acetonitrile) and solvent B (25 % acetonitrile in 0.5 M ammonium formate, pH 3.0).
2. Attach the Polysulfoethyl A ion exchange column to the LC system and equilibrate the column with 3 % Solvent B at a flow rate of 50 µL/min.
3. Solubilize the lyophilized samples from Subheading 3.3, **step 13** in 300 µL 0.1 % FA/25 % acetonitrile.
4. Inject 40 µL of the resolubilized sample onto the column and elute the peptides into a 96-well V-bottomed plate. Collect a total of 96 fractions at 1-min interval (*see* Fig. 1), using the following step gradient:
   (a) 0–3 min—3 % solvent B
   (b) 3–46 min—3–10 % solvent B.
   (c) 46–86 min—10–45 % solvent B.
   (d) 86–87 min—45–100 % solvent B.
   (e) 87–96 min—100 % solvent B.

   Adjust the above gradient if necessary to reach optimal separation.

5. Pool every five fractions into a 0.65 mL microcentrifuge tube, lyophilize, and store at −80 °C until ready for mass spectrometry analysis.

3.5 Final Sample Preparation for Proteomics Analysis

1. Add 100 µL of 0.1 % (v/v) formic acid to each pooled sample from the 96-well plate.
2. Sonicate each sample for 1 min, then gently vortex and centrifuge briefly.
3. Transfer the resolubilized samples to labeled autosampler vials and cap tightly.
4. Place on the LC autosampler for immediate analysis.
5. Peptide separation is performed using a fused-silica capillary column prepared in-house with Jupiter C-18 (5 µm particle size, 300 Å) packing material and with 75 µm OD tubing pulled to a tip with an opening of approximately 7–10 µm and cut to a length of 10 cm (*see* **Note 9**) coupled directly online with a mass spectrometer of the user's choice for protein identification (*see* **Note 10**).

4 Notes

1. The kit includes sealed vials containing 10 mL of deionized autoclaved water for the purpose of reconstituting the serum. To prepare the reconstituted serum, the content of one vial of water (10 mL) is transferred to a vial containing the lyophilized serum and the solution is gently agitated for 30 min at room temperature.
2. The centrifugation times may vary for the filtration cells based on filter pore size, temperature, and speed. Also, some filtration units may have sodium azide in the filter acting as a preservative. These filters should be flushed multiple times with water prior to use.
3. Either gel may be used for this step. Both gels gave near-identical results during initial tests.
4. Gel running times and voltages may be varied at this step based on the equipment used, manufactures recommendations, and gel types.
5. *In lieu* of using DTT, the protein disulfide bonds may be broken during the boiling step by using a commercially available solution such as Bond-Breaker tris(2-carboxyethyl)phosphine (TCEP) Solution (Thermo Scientific, West Palm Beach, FL). This is provided as a 0.5 M solution and provides a number of advantages over DTT or β-mercaptoethanol in that it can be

used in either acidic or basic conditions, has no pungent odor compared to DTT, and the breakage is almost irreversible, removing the need for alkylation.

6. Care must be taken when using iodoacetamide due to its high toxicity and sensitivity to light.
7. The trypsin solution was prepared by adding 1 mL of 50 mM NH_4HCO_3 (pH 8.4) to a vial containing 20 μg of trypsin. The vials are gently shaken and sonicated to facilitate solubilization. Excess trypsin solution prepared in 50 mM NH_4HCO_3 may be frozen and stored at −20 °C for use within 2–3 weeks.
8. This protein:enzyme ratio is not fixed. It may be varied based on the nature of individual samples and the amount of protein in the sample. Ratio is often adjusted and optimized to reach complete or nearly complete digestion of the proteins.
9. Based on the conditions and desired results, the investigator may wish to use their own preferred brand of column. The columns used for this experiment were prepared in-house with a length of 10 cm and used at a flow rate of 0.5 nL/min.
10. The data collected from the experiments was searched against the human proteome database using SEQUEST on a 10-node Beowulf cluster (ThermoFinnigan, San Jose, CA). Peptide identification was considered legitimate if the cross-correlation (Xcorr) and Delta correlation values (DeltaCN) matched or exceeded the set limits. A detailed explanation of the SEQUEST scoring criteria are listed on the website at http://omics.pnl.gov/software/SynopsisAndFirstHitsFiles.php. Table 1 shows a representative list of immune-related proteins identified using the described protocol here in this chapter.

Acknowledgment

This project has been funded in whole or in part with federal funds from the National Cancer Institute, National Institutes of Health, under contract number HHSN261200800001E. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the US Government.

Table 1
A representative list of immune-related proteins identified in the study

Protein	Gene ID
A kinase (PRKA) anchor protein 10	6005707
Absent in melanoma 1 protein	12643308
Adapter-related protein complex 2 alpha 2 subunit	12643300
Alpha-1D-adrenergic receptor	4501957
Androgen receptor A	86836
Androgen receptor; dihydrotestosterone receptor	4557331
Apoptotic protease activating factor isoform b	4502123
Brain-cadherin precursor	1705551
Bullous pemphigoid antigen 1	4502443
Chloride channel	4758002
Diacylglycerol kinase	4557519
EGF-containing fibulin-like extracellular matrix protein 1 precursor	9665262
Epidermal growth factor receptor substrate 15	1169540
Follicle Stimulating Hormone Receptor Precursor	544350
Fyn-related kinase	4503787
G protein-coupled receptor 50	4758468
G protein-coupled receptor GPR2	1082383
G protein-coupled receptor kinase	7661712
Gamma-aminobutyric acid receptor	14757381
Glucocorticoid receptor	4504133
Hepatocyte growth factor-regulated	4758528
HGF activator	4504383
Insulin receptor substrate 1	5031805
Insulin-like growth factor 1	11024682
Insulin-like growth factor 2 receptor	4504611
Insulin-like growth factor binding protein 5	10834982
Integrin alpha-3 precursor	11467963
Interleukin 15	10835153
Interleukin 4 receptor precursor	4557669
Kangai 1; CD82 antigen	4504813
Kinase-related transforming protein	88061
Leukemia inhibitory factor (cholinergic differentiation factor)	4504991

(continued)

Table 1
(continued)

Protein	Gene ID
Leukemia inhibitory factor receptor	4504993
Lymphoid blast crisis oncogene	5803058
MHC class II transactivator	4557749
Myristoylated alanine-rich protein kinase C substrate	11125772
Neural cell adhesion molecule 2	4758764
Olfactory receptor	7443964
Phosphoinositide-3-kinase, catalytic	5453892
Platelet-derived growth factor receptor alpha precursor	5453870
Potential phospholipid-transporting ATPase	8134335
Pre-B-cell leukemia transcription factor-2	1352729
Progesterone receptor	12644100
Progesterone receptor form B -human	625331
Protein tyrosine phosphatase, non-receptor type-3	4506293
Protein tyrosine phosphatase, receptor type	4506319
Protein tyrosine kinase, receptor type	87779
Protein tyrosine phosphatase	7439346
Protocadherin beta 7 precursor	11036656
Rho GTPase activating protein 1	4757766
Rho-associated, coiled-coil containing protein kinase 1	4885583
Secretory phospholipase A2 receptor precursor	1082778
Serine/threonine kinase 14 alpha	4506737
Signal recognition particle receptor	4507223
Signal transducer and activator of transcription 5A	14786068
Sodium channel, voltage-gated, type III	14732217
T-cell receptor V delta 5-C alpha region	477453
Thyroid Receptor Interacting Protein 4 (TRIP4)	2499057
Tumor necrosis factor-inducible protein TSG-6 precursor	1351315
Type II cAMP-dependent protein kinase RII anchoring protein	284481
Voltage-dependent anion channel 3	5032221
Voltage-dependent T-type calcium channel alpha-1G subunit	12644067

References

1. Bouamrani A, Hu Y, Tasciotti E, Li L, Chiappini C, Liu XW, Ferrari M (2010) Mesoporous silica chips for selective enrichment and stabilization of low molecular weight proteome. Proteomics 10:496–505
2. Chertov O, Simpson JT, Biragyn A, Conrads TP, Veenstra TD, Fisher RJ (2005) Enrichment of low-molecular-weight proteins from biofluids for biomarker discovery. Expert Rev Proteomics 2:139–145
3. Greening DW, Simpson RJ (2010) A centrifugal ultrafiltration strategy for isolating the low-molecular weight (≤25 K) component of human plasma proteome. J Proteomics 73: 637–648
4. Harper RG, Workman SR, Schuetzner S, Timperman AT, Sutton JN (2004) Low-molecular-weight human serum proteome using ultrafiltration, isoelectric focusing, and mass spectrometry. Electrophoresis 25:1299–1306
5. Drake RR, Cazares L, Semmes OJ (2007) Mining the low molecular weight proteome of blood. Proteomics Clin Appl 1:758–768
6. Lee JE, Kellie JF, Tran JC, Tipton JD, Catherman AD, Thomas HM, Ahlf DR, Durbin KR, Vellaichamy A, Ntai I, Marshall AG, Kelleher NL (2009) A robust two-dimensional separation for top-down tandem mass spectrometry of the low-mass proteome. J Am Soc Mass Spectrom 20:2183–2191
7. Adkins JN, Varnum SM, Auberry KJ, Moore RJ, Angell NH, Smith RD, Springer DL, Pounds JG (2002) Toward a human blood serum proteome – analysis by multidimensional separation coupled with mass spectrometry. Mol Cell Proteomics 1:947–955
8. Davis MT, Auger P, Spahr C, Patterson SD (2007) Cancer biomarker discovery via low molecular weight serum proteome profiling – where is the tumor? Proteomics Clin Appl 1:1545–1558
9. Stempfer R, Kubicek M, Lang IM, Christa N, Gerner C (2008) Quantitative assessment of human serum high-abundance protein depletion. Electrophoresis 29:4316–4323
10. Di Girolamo F, Alessandroni J, Somma P, Guadagni F (2010) Pre-analytical operating procedures for serum low molecular weight protein profiling. J Proteomics 73:667–677
11. Luque-Garcia JL, Neubert TA (2007) Sample preparation for serum/plasma profiling and biomarker identification by mass spectrometry. J Chromatogr A 1153:259–276
12. Tirumalai RS, Chan KC, Prieto DA et al (2003) Characterization of the low molecular weight human serum proteome. Mol Cell Proteomics 2:1096–1103
13. Basso D, Valerio A, Seraglia R et al (2002) Putative pancreatic cancer-associated diabetogenic factor: 2030 MW peptide. Pancreas 24:8–14
14. Petricoin EF 3rd, Ornstein DK, Paweletz CP et al (2002) Serum proteomic patterns for detection of prostate cancer. J Natl Cancer Inst 94:1576–1578
15. Rubin DB, Reznik G, Weiss EA, Young PR (2000) Non-protein thiols flux to S-nitrosothiols in endothelial cells: an LPS redox signal. Shock 14:200–207

Chapter 18

Identification of Protein Biomarkers in Human Serum Using iTRAQ and Shotgun Mass Spectrometry

Theodoros A. Koutroukides, Julian A.J. Jaros, Bob Amess, Daniel Martins-de-Souza, Paul C. Guest, Hassan Rahmoune, Yishai Levin, Mike Deery, Philip D. Charles, Svenja Hester, Arnoud Groen, Andy Christoforou, Julie Howard, Nick Bond, Sabine Bahn, and Kathryn S. Lilley

Abstract

Blood serum is one of the easiest accessible sources of biomarkers and its proteome presents a significant parcel of immune system proteins. These proteins can provide not only biological explanation but also diagnostic and drug response answers independently of the type of disease or condition in question. Shotgun mass spectrometry has profoundly contributed to proteome analysis and is presently considered as an indispensible tool in the field of biomarker discovery. In addition, the multiplexing potential of isotopic labeling techniques such as iTRAQ can increase statistical relevance and accuracy of proteomic data through the simultaneous analysis of different biological samples. Here, we describe a complete protocol using iTRAQ in a shotgun proteomics workflow along with data analysis steps, customized for the challenges associated with the serum proteome.

Key words iTRAQ, Mass spectrometry, Serum, Depletion, Strong cation exchange

1 Introduction

An increasing interest has been shown in the human blood serum proteome for identification of potential protein markers and biochemical pathways associated with the disease state [1]. Blood contains the largest version of the human proteome, considering the presence of leakage markers from many cell types and the variety of numerous immunoglobulin sequences [2]. A relatively high concentration of proteins is present in the serum, typically in the range of 60–80 mg of protein per mL. Protein-based biomarker studies in different therapeutic areas can be carried out through minimally invasive means from the bloodstream. In a clinical setting, blood is easy to access and can be processed quickly to serum.

Kelly M. Fulton and Susan M. Twine (eds.), *Immunoproteomics: Methods and Protocols*, Methods in Molecular Biology, vol. 1061, DOI 10.1007/978-1-62703-589-7_18,

Furthermore, serum contains circulatory analytes and components of tissue proteomes, and studies have shown that potential protein-based biomarkers are secreted or leaked into the blood stream in an altered manner in response to disease, drug treatment, or other causes [3, 4].

A special concern in the case of serum proteome is that the presence of some protein biomarkers can be obscured, given the large dynamic range of individual protein concentrations in blood (more than 10^{10} orders of magnitude)[2, 5, 6]. There is also the confounding presence of some high-abundance proteins (HAP) such as serum albumin and the various forms of immunoglobulins that alone constitute up to 80 % of the total protein mass in serum. Consequently, identification of the low abundance proteins (LAP), which comprise thousands of individual molecules, is difficult without some form of separation or purification. Therefore, the first essential stage in the workflow is to deplete the HAP to unmask this potential biomarker source in the LAP fraction [7]. This procedure works well, although most often only the LAP fraction has been analyzed. We suggest that it is important to also investigate the HAP fraction. One reason for this is obvious as some potential biomarkers of hydrophobic nature could arise from the HAP fraction. Another reason has arisen from studies which have shown evidence of nontargeted co-depletion of LAP which can happen due to several reasons [8].

A significant fraction of the blood proteome plays roles in immune system function. Studies of these proteins are not only important for increasing our understanding of inflammatory functions and autoimmune diseases but also for unraveling their function in other diseases such as central nervous system disorders [9, 10], microbial infection [11, 12], and cancer [4, 13] which are known to be associated with an altered immune response. Consequently, searching for biomarkers in serum will lead to the identification of immune-related proteins even if the stressor agents or disease conditions are not directly related to the immune system. For example immune- or inflammation-related proteins have been shown to be altered in studies of schizophrenia [14, 15], Alzheimer's disease [16], and heart conditions [17]. However, the multiplex nature of proteomic mass spectrometry (MS) profiling [18] allows comprehensive coverage of the proteome and the detection of changes in whole protein patterns reflecting the biological signatures underpinning the disease etiology. This should enable the detection of protein patterns such as those of the immune system, which are unique for each specific disorder.

Although two-dimensional gel electrophoresis is still the most widely used proteomic methodology, there have been recently several versions of shotgun proteomics platforms which have come into more frequent use in biomarker discovery studies [19]. Non-labeling spectral counting approaches (also known as label-free

shotgun proteomics) [20] have been successfully used, but the need for more accurate relative and absolute quantification led to the development of a number of systematic approaches such as (a) reduction of sample complexity using affinity tag labeling methods [21–23], (b) modification of the fragmentation behavior of peptides using charge derivatization [24–26], or (c) employing protein/peptide isotopic labeling to increase quantitative accuracy [27–29].

iTRAQ (Isobaric Tags for Relative and Absolute Quantification) [28] and TMT (Tandem Mass Tags) [30] are two of the most prevalent chemical labeling methods. Compared to label-free approaches, the multiplexing potential of isotopic labeling techniques increases statistical relevance and accuracy through the analysis of different biological samples simultaneously and through normalization to an internal standard. The iTRAQ technology provides up to eight isobaric tags that lead to labeling of peptides identical in mass and, at the same time, each tag yielding a distinct, low-mass, diagnostic MS/MS signature ion. Labeling involves covalent bonding of the NHS-ester based tags to the N-terminus and lysine side chains of all peptides containing these in a digest mixture. After labeling peptides from each sample can be pooled, reducing technical variance for the downstream LC-MS/MS analysis. As all tags are initially nearly isobaric (i.e., of equal mass), the labeled peptides are at first indistinguishable. However, each tag consists of labile asymmetric reporter and balance regions of different masses. This enables release of reporter ions of distinct masses upon fragmentation of peptides in the MS collision chamber [28].

The relative quantification of the labeled peptides in an iTRAQ MS experiment is performed by calculating reporter ion peak intensity ratios in product ion scans, based on the fact that all differentially labeled forms of the peptide are indistinguishable in this mode. This can be feasible only using an MS/MS setup, which can isolate precursor ions and fragment these at a given collision energy. The MS/MS survey scan used in iTRAQ experiments is also referred to as Data-Dependent Analysis (DDA) [31, 32], which is in contrast to the label-free approach which uses an MS scan in expression mode (MS^E) [33]. In MS^E, the MS switches rapidly from low to high collision energy without isolation of the precursors but with scanning at a range of set masses [34]. On the other hand, DDA is an automated mode in which the instrument switches automatically to MS/MS according to user set parameters such as "product ion decay" and "neutral loss." In the following protocol, we have used a Quadrupole Time-of-Flight (Q-TOF) MS. The two main advantages of this configuration are the high specificity obtained by the quadrupole mass analyzer in selection and filtering of peptide ions according to mass/charge (m/z), and the capability of the TOF mass analyzer to accurately measure the m/z of the fragmented precursor ions [35].

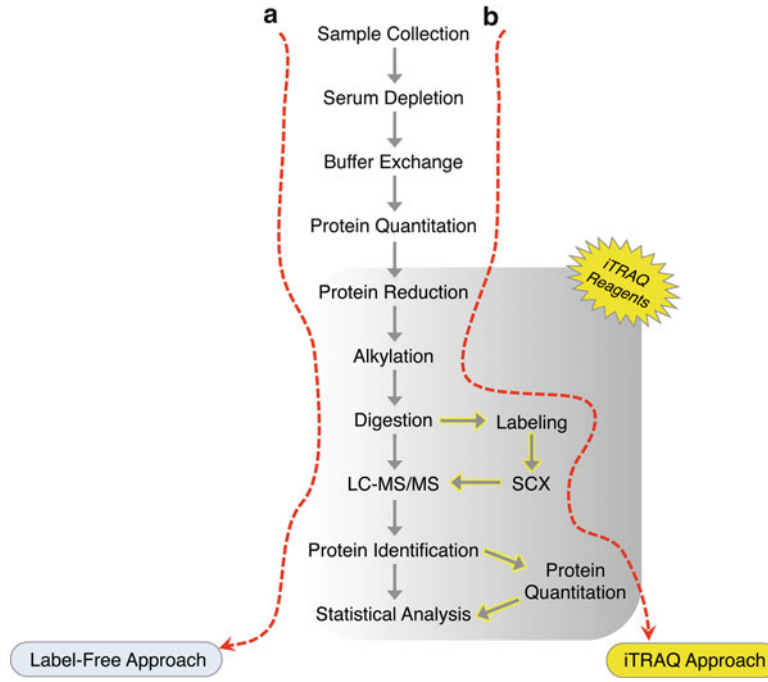


Fig. 1 Flowchart outlining methods from sample preparation to data analysis of human serum performed using: (*A*) a conventional label-free approach and (*B*) the iTRAQ reagent

We have developed a protocol for serum analyses using iTRAQ technology in combination with a liquid chromatography (LC) Q-TOF MS/MS instrument in a "bottom-up" shotgun approach, in which all proteins are enzymatically cleaved into smaller peptides prior to iTRAQ labeling (*see* Fig. 1). This approach takes advantage of the high separation efficiency of LC on the peptide level, which facilitates analysis of multiple peptides from the same protein for increased quantitative accuracy. As we observed experimentally [36, 37], this workflow is suitable for immunoproteomics, since a great parcel of the detected and quantified proteins are immune system related.

2 Materials

2.1 Serum Depletion

1. Human plasma or serum samples, 40 µL per injection on the column given below.
2. Agilent MARS Hu14 column (4.6 mm inner diameter, length 100 mm, 1.66 mL bed volume, capacity of 40 µL of plasma/serum).
3. Buffer solution A by Agilent (1 L bottle): neutral salt solution containing phosphate, pH 7.4 (*see* **Note 1**).

4. Buffer solution B by Agilent (1 L bottle): a concentrated urea solution, pH 2.2 (*see* **Note 1**).
5. Filter centrifuge tubes with 0.22 µm pore size (available from various manufacturers including Agilent) for removing particulates in the samples before the depletion.
6. HPLC system (ÄKTA™ purifier UPC 10, GE Healthcare, Unicorn control software (v5.11 build 407), pump P-900, Frac-920 fraction collector and manual injection system using Hamilton syringe with 250 µL capacity).
7. Microcentrifuge.
8. Eppendorf lo-bind tubes (0.5 mL).

2.2 Sample Concentration and Buffer Exchange

1. 1 M triethylammonium bicarbonate (TEAB).
2. Urea.
3. Sodium dodecyl sulfate(SDS).
4. iTRAQ LB (labeling buffer). For 100 mL of LB add 90 mL H_2O + 6 g Urea (i.e., 1 M Urea) + 2.5 mL of 1 M TEAB (i.e., 25 mM TEAB), put on shaker until completely dissolved, add 1 mL of 10 % SDS (i.e., 0.1 % SDS), adjust pH to 8.5, top-up water to 100 mL (*see* **Note 2**).
5. Spin concentrators 5 kDa MWCO.
6. Multifuge microcentrifuge.

2.3 Sample Concentration Measurement

1. Pierce BCA protein assay reducing agent compatible: Protein assay reagents A and B, compatibility reagent, reconstitution buffer (*see* **Note 3**).
2. Bovine serum albumin (BSA).
3. 10 and 350 µL-capacity multichannel pipette.
4. Multichannel solution reservoirs/basins.
5. Flat-bottom 96-well plate(s).
6. Shaker (Heidolph Titramax 100).
7. Orbital incubator set at 37 °C.
8. Microplate reader set at 560 nm.

2.4 Sample Digestion

1. Tris-(2-carboxyethyl) phosphine (TCEP) (vial with yellow lid).
2. Methyl methanethiosulfonate (MMTS) (vial with purple lid).
3. Trypsin.
4. Shaker.
5. Waterbath set at 37 °C.
6. Heating block.
7. Freeze dryer and glass flasks.

2.5 iTRAQ Labeling

1. Speedvac.
2. 8-plex iTRAQ reagents.
3. 1 M TEAB.
4. 100 % isopropanol (iTRAQ reagents).
5. Shaker.
6. Hydrochloric acid (HCl)/sodium hydroxide (NaOH) for pH adjustment.
7. pH indicator strips.

2.6 Strong Cation Exchange (SCX)

1. Buffer A for SCX (1 L): 10 mM KH_2PO_4 (1.36 g), 20 % ACN (200 mL 100 % ACN, 800 mL H_2O), adjusted to pH 2.7 with 500 mM H_3PO_4.
2. Buffer B for SCX (1 L): 10 mM KH_2PO_4 (1.36 g), 1 M KCl_2 (74.55 g). 20 % ACN (200 mL 100 % ACN, 800 mL H_2O), adjusted to pH 2.7 with 500 mM H_3PO_4.
3. 1 M TEAB.
4. 5 mg/mL BSA (in 50 mM TEAB); i.e., 5 mg BSA, 50 µL of 1 M TEAB, 950 µL of H_2O.
5. 100 % acetonitrile (ACN).
6. pH indicator strips (pH 2.0–9.0).
7. Trypsin.
8. HPLC thermostat temperature controlled column compartment (DionexTCC-100) and gradient pump.
9. Fraction collector.
10. Ultrasonic equipment.
11. SCX column (PolyLCPolySulfoethyl A 4.6×100 mm, 5 µm, 300 Å)

2.7 MS/MS Run

1. Quadrupole time-of-flight premier mass spectrometer (Waters Premier).
2. nanoUltra performance liquid chromatography system (Waters 10 kps/in anoAcquity).
3. 0.1 % Formic acid (FA); i.e., 100 µL of 100 % FA in 100 mL of H_2O (HPLC grade).

2.8 Data Analysis

1. iTRAQ quantitation software (e.g., i-Tracker [38], Proteome Discoverer [39], Protein Pilot, Mascot Distiller).
2. ProteinLynx Global SERVER (Waters).
3. Mascot Server; Mascot Daemon (Matrix Science).
4. Human proteome FASTA sequence database (UniProt).
5. Multiplexed dataset statistical analysis script. The R scripts used to process and analyze data from iTRAQ® 8-plex are

available from our anonymous FTP server by following the relevant link from the Links section of our website (http://www.biot.cam.ac.uk/sb/).

3 Methods

3.1 Sample Organization

1. Generate list of samples, noting clearly provider, initial amount provided, state of source, and demographics.
2. Randomize original Sample identifications in blind manner so each sample is given a distinct Experiment ID.

3.2 Serum Depletion

1. Dilute the plasma/serum samples four times with buffer solution A in 0.5 mL Eppendorf lo-bind tubes (e.g., add 40 µL of plasma/serum to 120 µL of buffer solution A) (*see* **Note 4**).
2. Wash the 0.22 µm filter centrifuge tubes by adding 500 µL of buffer solution A, centrifuge, and discard the filtrate.
3. Transfer the entire sample (160 + 10 = 170 µL, *see* **Note 4**) into the filter tubes, centrifuge at 16,000 × *g* for 1 min or until the samples have passed through the filter (*see* **Note 5**).
4. Set up buffer solution A and buffer solution B as the only mobile phases of the HPLC system as per manufacturer's instructions? (*See* **Note 6**.)
5. Purge HPLC lines with buffer solution A and buffer solution B at a flow rate of 1 mL/min for 10 min.
6. Set up the LC method following the manufacturer's instructions and run a blank method by injecting 200 µL of buffer solution A without a column. Ensure proper sample loop size (*see* **Note 7**).
7. Connect the Agilent MARS Hu14 column as per usual with the HPLC system. Make sure to do a drop-to-drop connection in order to avoid air bubbles in the system.
8. Equilibrate the column with buffer solution A at a flow rate of 1 mL/min for 4 min. Prior to injection of the first sample, run a blank method by injecting 200 µL of buffer solution A.
9. Inject the diluted, filtered serum into the HPLC sample loop.
10. Start the depletion method. Collect the fractions as they elute from the column (*see* **Note 8**). The HAP targeted for depletion will bind to the column, whereas other proteins flow through within the first 5 mL/20 min.
11. After a blank run, disconnect and store the depletion column at 4 °C in buffer solution A.
12. Flush the HPLC system with water for 30 min at 1 mL/min, then with 20 % ethanol or isopropanol in water for 30 min at 1 mL/min.

3.3 Sample Concentration and Buffer Exchange

1. Make up fresh iTRAQ LB (amount depends on size of pellet to be resuspended).
2. Pre-rinse a 5 kDa MWCO concentration tube (Agilent) by adding 2 mL of freshly made LB, spin it down at 4,000 × *g* for 5 min (or until spun down to <100 μL). Each concentration tube is to be used for one sample only.
3. Discard the filtrate and the concentrate in the concentration tube using a pipette. Add the sample into the now washed concentration tube and spin at 4,000 × *g* for about 20 min or until the volume is 200 μL.
4. Add 2 mL of LB. Pipette the sample up and down for mixing and to prevent proteins from stacking onto the filter. Spin again at 4,000 × *g* for ~20 min or until the volume is 200 μL.
5. Repeat the washing procedure twice for the high-abundant fraction, as the elution buffer (about 6 M Urea) is much more concentrated than the loading buffer.
6. Draw out 200 μL of the sample from the filter pocket and transfer to an Eppendorf lo-bind tube (0.5 mL).
7. This procedure should buffer exchange the sample, as well as reduce the volume down to 200 μL (duration ~2 h or longer for 16 samples). The final concentration will be 1–2 mg/mL (low-abundant) and up to 20 times the concentration for the high-abundant fraction.

3.4 Sample Concentration Measurement

1. Make up eight dilutions of fresh albumin standards ranging from 1:1 to 1:8(Pierce, 2 mg/mL) in LB.
2. Dilute the Reconstitution Buffer (RB) 1:1with ultrapure H_2O (i.e., 100 μL of RB + 100 μL of H_2O).
3. Puncture the foil covering on the Compatibility Reagent (CR) with an empty pipette tip. Add 100 μL of diluted RB and dissolve by stirring at the bottom of the tube and pipette up and down 15–20 times.

   Total volume of CR required = (*n* standards + *n* samples) × (*n* replicates) × (4 μL of CR per well).
4. Pipette 9 μL of standard/sample in duplicates into 96-well plate.
5. Add 4 μL Compatibility Reagent to the standard/sample in each well.
6. Cover plate and mix on Heidolph Titramax 100 at medium speed for 1 min.
7. Incubate plate at 37 °C for 15 min using orbital incubator.
8. Prepare Working Reagent (WR) by mixing BCA Reagent A (big bottle) and Reagent B (small, blue bottle) in a solution basin at a 49:1 dilution (e.g., 5 mL Reagent A + 100 μL Reagent B).

Total volume of Working Reagent (WR) required = (*n* standards + *n* samples) × (*n* replicates) × (260 µL WR).

9. Add 260 µL of Working Reagent to each well and shake plate for 1 min on Heidolph Titramax 100.
10. Incubate covered microplate at 37 °C for 30 min without further shaking using orbital incubator SI 50.
11. Cool plate to room temperature (RT).
12. Measure protein concentration immediately at 560 nm using Bio-Rad Model 680 Series Microplate reader (*see* **Note 9**).
13. Perform further calculations in MS Excel (*see* **Note 10**).

3.5 Sample Digestion

1. For sample reduction add 50 µg of sample in an Eppendorf lo-bind tube (0.5 mL).
2. If the volume after protein estimation is <20 µL, top-up to 20 µL with LB.
3. Add 50 mM TCEP (vial with yellow lid);if volume is <40 µL, use 2 µL of 50 mM TCEP, if volume is >40 µL, use 4 µL of 50 mM TCEP. Vortex.
4. Incubate at RT for 1 h on shaker.
5. For cysteine blocking use MMTS (vial with purple lid). If volume is <40 µL, use 1 µL of MMTS, if volume is > 40 µL, use 2 µL of MMTS. Vortex.
6. Incubate at RT for 10 min on shaker.
7. For digestion, dilute sample 1:2.5 with 50 mM TEAB (e.g., 20 µL + 50 µL).
8. Add 100 µL of 50 mM TEAB (50 mM TEAB = 1:20 1 M TEAB Solution) to 1 trypsin vial (=20 µg).
9. Add 12.5 µL (=2.5 µg trypsin) of resuspended trypsin to sample. Vortex.
10. Incubate for 1 h at 37 °C.
11. Repeat **steps 9** and **10** twice.
12. Incubate overnight at 37 °C.

3.6 iTRAQ Labeling

1. Transfer the samples to a −80 °C freezer until all samples are frozen completely.
2. Prepare the Micromodulyo freeze dryer (Thermo Electron) by turning on the freeze dryer unit and the pump connected to it at least 2 h before freeze drying the samples to ensure the "coil-in-chamber" ice condenser reaches −50 °C.
3. Transfer the frozen samples to a stainless steel rack fitting inside the 1 L glass flask provided with the freeze dryer (Thermo Electron) and place the flask on dry ice immediately where it should be kept throughout the duration of the lyophilization.

4. Connect the flask to the freeze dryer via the 8-port manifold and open the respective tube to create vacuum inside the flask. On average, a 1.5 mL sample would take about 12 h to freeze dry.
5. Resuspend pellet in 25 µL of 1 M TEAB and 60 µL of isopropanol (*see* **Note 11**).
6. Adjust pH to 8.5.Vortex for 1 min.
7. Centrifuge iTRAQ-tag vials briefly to ensure that the whole material is gathered at the bottom.
8. Pipette all 85 µL of one sample into 1 labeling vial and note the iTRAQ-tag/sample relationship. Vortex for 1 min. Incubate at RT for 2 h on shaker.
9. Add 100 µL of H_2O to each vial. Incubate at RT for 2 h on shaker.
10. Pool samples in an Eppendorf lo-bind tube (1.5 mL).Vortex.
11. Freeze-dry to dryness as described in Subheading 3.6; **steps 1–4**.
12. Proceed to SCX; otherwise store samples at −20 °C.

3.7 Strong Cation Exchange (SCX)

1. Degas solvents using ultrasonic water bath.
2. Install SCX column onto the Dionex-B HPLC.
3. Go to FLOW ON; MANUAL ACQUISITION; check UV spectrum for pressure changes (e.g., abrupt spikes in pressure measurement or unstable signal) indicating the presence of air bubbles.
4. Connect SCX Buffers to Dionex-B taking care not to introduce air bubbles into the system.
5. Prime Dionex-B with Buffers A and B, wash syringe with dH_2O.
6. Load 3 mL SCX Buffer A to syringe and inject into Dionex-B.
7. Run blank using preset program.
8. Repeat above three steps.
9. Add 120 µL of 5 mg/mL BSA (in 50 mM TEAB) to 1 vial of trypsin. Vortex.
10. Add 15 µL of ACN. Vortex. Incubate overnight at 37 °C.
11. Resuspend 40 µL of digested BSA in 3 mL SCX Buffer A.
12. Adjust pH to 2.7.
13. Load all 3 mL to syringe and inject into Dionex-B.
14. Load Eppendorf lo-bind tubes (2 mL) to Foxy fraction collector.
15. Set flow rate to 0.2 mL/min and fraction collection at 2-min intervals (i.e., 0.4 mL fraction volume).
16. Perform SCX runs for two digested BSA samples using the following gradient: min 0—100 % Buffer A/0 % Buffer B; min

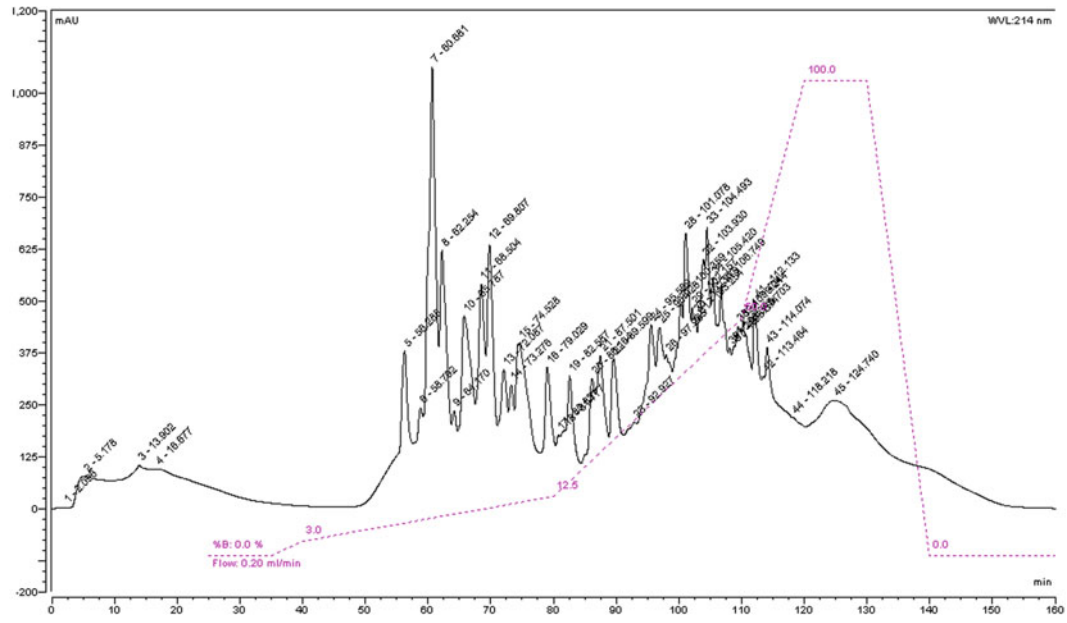


Fig. 2 Typical SCX chromatogram (*black solid line*) of human serum sample performed with described methods and gradient (*dotted line*)

35—100 % Buffer A/0 % Buffer B; min 40—97 % Buffer A/3 % Buffer B; min 80—87.5 % Buffer A/12.5 % Buffer B; min 110—50 % Buffer A/50 % Buffer B; min 120—0 % Buffer A/100 % Buffer B; min 130—0 % Buffer A/100 % Buffer B; min 140—100 % Buffer A/0 % Buffer B (*see* Fig. 2).

17. Collect the tubes from the fraction collector.
18. Resuspend sample in 3 mL of SCX Buffer A.
19. Adjust pH to 2.7.
20. Load all 3 mL to syringe and inject into Dionex-B.
21. Load Eppendorf lo-bind tubes (2 mL) to Foxy fraction collector.
22. Run sample using appropriate software program (*see* Fig. 3).
23. Collect the tubes from the fraction collector.
24. Wash system by running blank between sample runs.
25. Freeze-dry to dryness as described in Subheading 3.6; **steps 1–4**.

3.8 MS/MS Run

1. Pool SCX fractions where necessary to account for fractions that contain low peptide content in order to even out approximate total protein content among samples.
2. Resuspend each sample in 32 µL of 0.1 % FA, i.e., enough for two runs accounting for duplicate injections + 6 L residual volume (*see* **Note 12**).

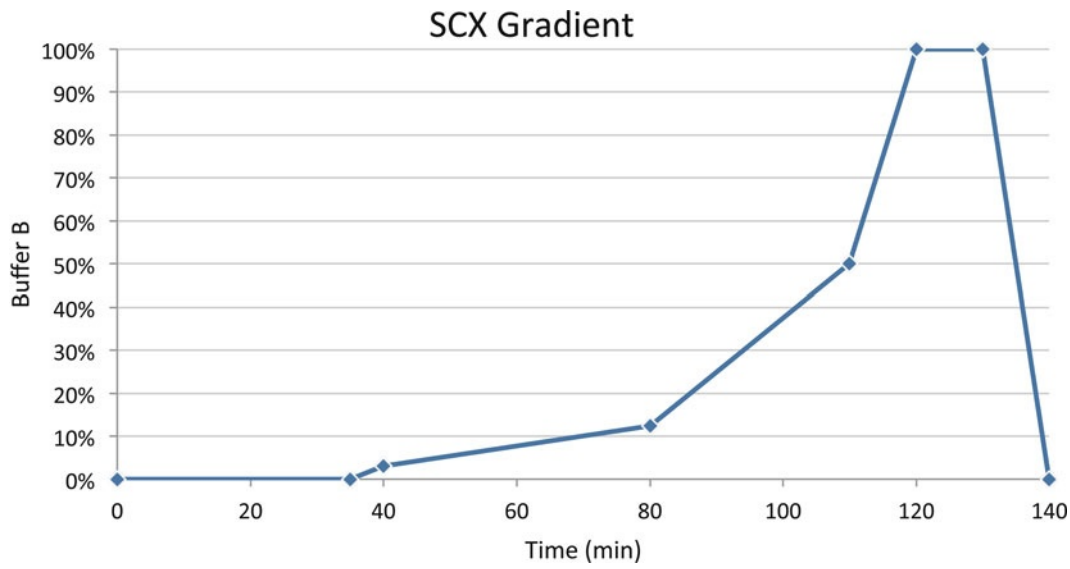


Fig. 3 Graph showing the continuous gradient used to elute peptides from the SCX column. The *Y*-axis indicates the buffer B percentage and the *X*-axis indicates the elution time.

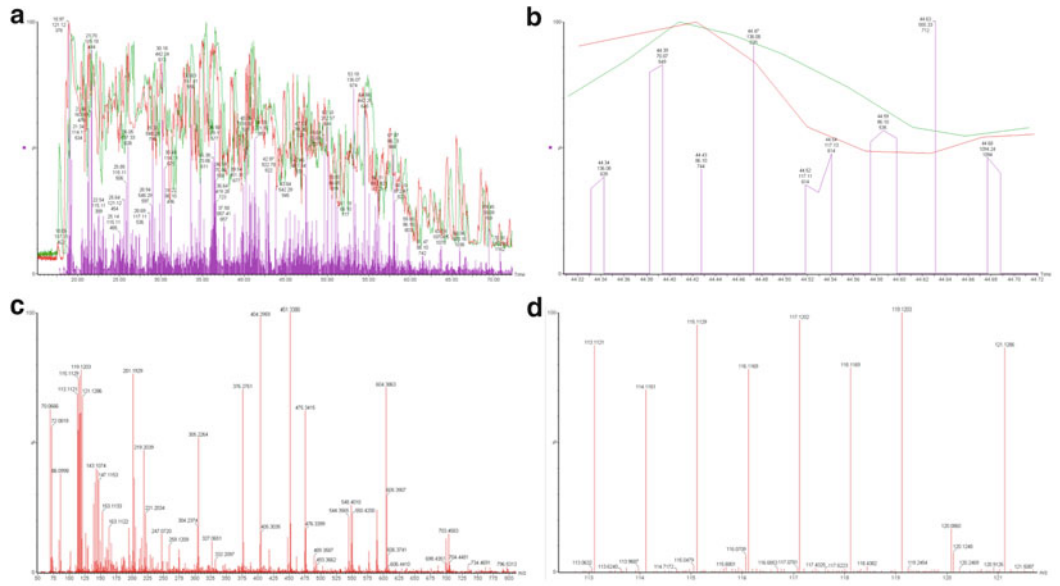


Fig. 4 Example Q-TOF MS/MS chromatograms and spectrums of iTRAQ-labeled peptides: (**a**) Base Peak Intensity (BPI) chromatograms of duplicate run (*red* and *green*) and MSMS 2+ ions chromatogram (*purple*); (**b**) zoomed-in region from (**a**); (**c**) MS/MS spectrum between 44.52 and 44.54 min; and (**d**) zoomed-in region from (**c**) showing signals of 8-plex tags (Color figure online)

3. Load 16 μL of each of the pooled fractions into MS vials and load into NanoLC.
4. Compile sample list in MassLynx (*see* **Note 13**) and perform MS run (*see* Fig. 4).

3.9 Data Analysis

1. Convert the raw data to mzXML or mzML files (*see* **Note 14**). Converters are currently available from the ProteoWizard (http://proteowizard.sourceforge.net/) and Trans-Proteomic Pipeline (TPP) (http://tools.proteomecenter.org/TPP.php) projects.
2. Generate MS2 peak list.
3. Perform MASCOT protein identification searches (*see* **Note 15**).
4. Perform iTRAQ quantitation and statistical analysis. While we are not in position to comment on comparability among the different software packages that are available, the following are capable of performing iTRAQ quantitation: i-Tracker, Proteome Discoverer, Protein Pilot, Mascot Distiller, ProQuant, jTraqX, Scaffold Q+.

4 Notes

1. The exact composition of these buffer solutions is unknown to the authors as their composition is proprietary. The buffers can be ordered separately in any desired quantity (bottles of 1 L) or as part of a kit containing both buffers and filter/concentration centrifuge tubes.
2. Do not heat samples above 37 °C for extended periods of time to minimize carbamylation of peptides by Urea (present in the buffer).
3. The Bio-Rad DC assay is not compatible with TEAB, which is a required buffer for iTRAQ labeling. The Pierce BCA assay provides compatibility for samples that contain disulfide reducing agents such as TEAB, dithiothreitol (DTT), 2-mercaptoethanol (BME), and TCEP.
4. Add an extra 10 µL of buffer A to enable the uptake of a full 160 µL later, or dilute a higher sample volume (e.g., 45 µL of sample with 135 µL of buffer solution A) and then inject 160 µL only.
5. Prepare the samples immediately before the depletion procedure or keep the diluted and filtered samples on ice until use.
6. The buffer solutions do not need any filtering or degassing. Purge the HPLC system with water first if it has been stored in 20 % ethanol/isopropanol in water, then with the buffer solutions. The direct change from an alcohol solution to the buffer solutions might cause salt precipitation within the system.
7. The sample loop volume should be at least twice the volume of the injected sample as usual when using partial loop loading. We use a 500 µL loop for a sample volume of 160 µL.
8. Use low-protein-binding tubes for fraction collection that do not contaminate the sample. Tubes made of polyethylene

glycol (PEG) can release material into the sample, which leads to a strong interfering signal during MS analysis. We use rimless polypropylene (PP) tubes with a capacity of 10 mL and, in combination with Eppendorf lo-bind tubes (0.5 mL, also made of PP), we do not experience such contamination.

9. BCA Assay does not reach a true end point, color development will continue. The absorbance increases at a rate of 0.25 % per minute.
10. To estimate the protein concentration of the test sample with Excel, interpolate its recorded absorbance in a standard curve (scatter chart in Excel) plotted with the absorbances (Y-axis) of the known concentrations (X-axis) of the multiple dilutions of the standard sample.
11. Ensure that the exact amount of isopropanol is used as this has a dehydration function and aids solubilization of the 8 plex tags. Presence of excess H_2O will result in the rapid hydrolysis of iTRAQ-tags, leading to incomplete and inconsistent labeling reactions.
12. The following MS parameters are recommended: Acquisition: 0–120 min; Source: ES (positive polarity, V analyzer mode, normal dynamic range); MS survey range: 300–1,800 Da; Switch to MS/MS acquisition when TIC rises above: 5 counts/s; Scan time: 0.48 s; Inter-scan delay: 0.02 s; Data format: Continuum; MS/MS range: 100–1,500 Da; 1 MS/MS ions maximum from single MS survey scan; Scan time: 0.31 s; Inter-scan delay: 0.02 s; Data format: Continuum; Peak detection window: 1 Da; MS cone voltage: 35 V; Collision energies 2+ ions: 300, 24.2; 566, 31.9; 830, 33; 955, 44; 1,200, 60.5; Collision energies 3+ ions: 400, 25.3; 653, 26.4; 740, 28.6; 820, 33; 1,200, 58.3; 2,000, 80.3; Collision energies 4+ ions: 435, 15.4; 547, 19.8; 605, 25.3; 1,000, 31.9; 2,000, 63.8; Reference scan time: 0.48 s; Reference scan frequency: 60 s; Reference sampling cone: 25 V; Reference collision energy: 24 V; Mass window ±: 0.5 Da; Scans to average: 3; DXC temperature correction: Off; Trapping (A = water, B = acetonitrile): 4 min; 20 μL/min; 100 % water/0 % acetonitrile; 0–8,000 psi; Gradient (A = water, B = acetonitrile; 0–10,000 psi, 3 min seal wash, 122 min run time): min 0—95 % A/5 % B; min 1—95 % A/5 % B; min 90—45 % A/55 % B; min 110—15 % A/85 % B; min 110.1—95 % A/5 % B; min 120—95 % A/5 % B (*see* Fig. 5).
13. Multiple ways exist to carry out quantitation. If an open source script is to be used, then the raw data will have to be converted to an open standard such as mzXML or mzML.
14. The following Mascot parameters are recommended: Taxonomy: Metazoa…homo sapiens(human); Database: UniProt Human (EBI integr8 project release 120, built from

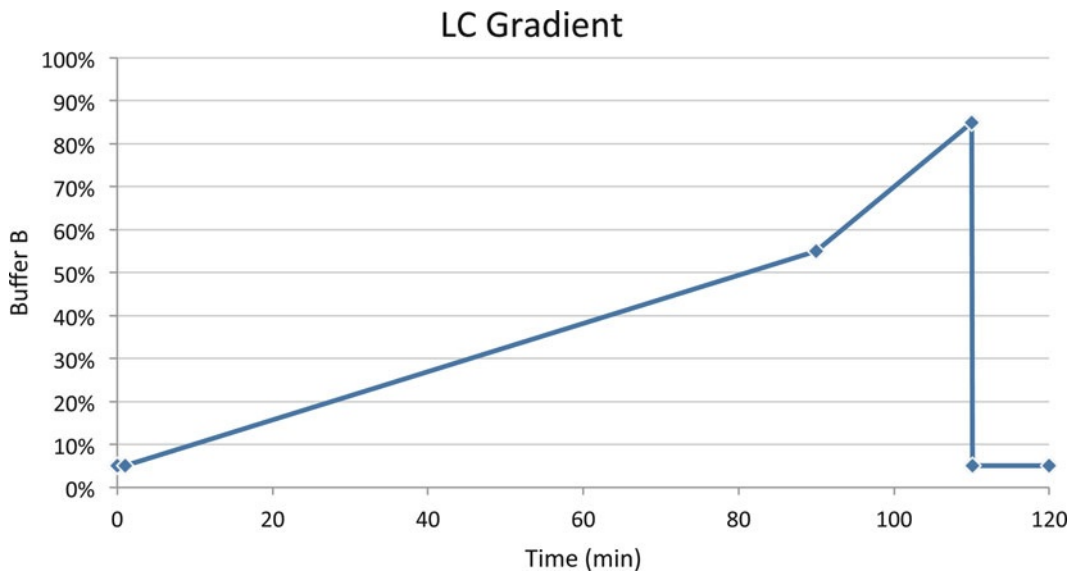


Fig. 5 Graph showing the continuous gradient used to elute peptides from the LC column. The *Y*-axis indicates the buffer B percentage and the *X*-axis indicates the elution time

UniProt 2011_06, 74,393 sequences); Decoy database: YES; Monositopic/Average: Monoisotopic; Fixed modifications: iTRAQ4plex(K)/iTRAQ8plex(K), iTRAQ4plex(N-term)/iTRAQ8plex(N-term), Methylthio (C); Variable modifications: iTRAQ4plex(Y)/iTRAQ8plex(Y), Oxidation(M); Report top: AUTO; Protein mass: (leave blank); Enzyme: Trypsin; Max. Missed cleavages: 2; Peptide charge: 2+, 3+ (and 4+); Peptide tol.: 1.2 Da; #13C: 0; MS/MS ions search: YES; Error tolerant search: (leave unselected); Data format: Mascot Generic; MS/MS tolerance.: 0.8 Da; Quantification: (leave blank); Instrument: ESI-QUAD-TOF.

15. The efficiency of iTRAQ labeling is variable. In a given iTRAQ study, the number of identified proteins may be more than five times higher than the number of quantifiable proteins [40, 41]. By using this protocol in a study on 49 samples it was possible to quantify 122 proteins (quantifiable in at least 75 % of the analyzed samples) with an average number of identified peptides/protein of 5.4. The average sequence coverage was 15.1 %.

References

1. Liotta LA, Ferrari M, Petricoin E (2003) Clinical proteomics: written in blood. Nature 425:905
2. Anderson NL, Anderson NG (2002) The human plasma proteome: history, character, and diagnostic prospects. Mol Cell Proteomics 1:845–867
3. Ray S, Reddy PJ, Jain R, Gollapalli K, Moiyadi A, Srivastava S (2011) Proteomic technologies for the identification of disease biomarkers in serum: advances and challenges ahead. Proteomics 11:2139–2161
4. Blonder J, Issaq HJ, Veenstra TD (2011) Proteomic biomarker discovery: it's more than

just mass spectrometry. Electrophoresis 32: 1541–1548

5. Zhang Q, Faca V, Hanash S (2011) Mining the plasma proteome for disease applications across seven logs of protein abundance. J Proteome Res 10:46–50
6. Liumbruno G, D'Alessandro A, Grazzini G, Zolla L (2010) Blood-related proteomics. J Proteomics 73:483–507
7. Polaskova V, Kapur A, Khan A, Molloy MP, Baker MS (2010) High-abundance protein depletion: comparison of methods for human plasma biomarker discovery. Electrophoresis 31:471–482
8. Koutroukides TA, Guest PC, Leweke FM, Bailey DMD, Rahmoune H, Bahn S, Martins-de-Souza D (2011) Characterization of the human serum depletome by label-free shotgun proteomics. J Sep Sci 34:1621–1626
9. Mayilyan KR, Weinberger DR, Sim RB (2008) The complement system in schizophrenia. Drug News Perspect 21:200–210
10. Robotti A, Natale M, Albo AG, Lis K, Perga S, Marnetto F, Gilli F, Bertolotto A (2010) Acute-phase proteins investigation based on lectins affinity capture prior to 2-DE separation: application to serum from multiple sclerosis patients. Electrophoresis 31:2882–2893
11. Ren F, Chen Y, Wang Y, Yan Y, Zhao J, Ding M, Zhang J, Jiang Y, Zhai Y, Duan Z (2010) Comparative serum proteomic analysis of patients with acute-on-chronic liver failure: alpha-1-acid glycoprotein maybe a candidate marker for prognosis of hepatitis B virus infection. J Viral Hepat 17:816–824
12. Pitarch A, Nombela C, Gil C (2011) Prediction of the clinical outcome in invasive candidiasis patients based on molecular fingerprints of five anti-Candida antibodies in serum. Mol Cell Proteomics 10:M110
13. Solassol J, Du-Thanh A, Maudelonde T, Guillot B (2011) Serum proteomic profiling reveals potential biomarkers for cutaneous malignant melanoma. Int J Biol Markers 26: 82–8710
14. Muller N, Schwarz M (2006) Schizophrenia as an inflammation-mediated dysbalance of glutamatergic neurotransmission. Neurotox Res 10: 131–148
15. Schwarz E, Guest PC, Rahmoune H, Harris LW, Wang L, Leweke FM, Rothermundt M, Bogerts B, Koethe D, Kranaster L, Ohrmann P, Suslow T, McAllister G, Spain M, Barnes A, van Beveren NJM, Baron-Cohen S, Steiner J, Torrey FE, Yolken RH, Bahn S (2012) Identification of a biological signature for schizophrenia in serum. Mol Psychiatry 17:494–502
16. Casoli T, Di Stefano G, Balietti M, Solazzi M, Giorgetti B, Fattoretti P (2010) Peripheral inflammatory biomarkers of Alzheimer's disease: the role of platelets. Biogerontology 11: 627–633
17. Jiang B, Liao R (2010) The paradoxical role of inflammation in cardiac repair and regeneration. J Cardiovasc Transl Res 3:410–416
18. Aebersold R, Mann M (2003) Mass spectrometry-based proteomics. Nature 422: 198–207
19. Martins-de-Souza D, Guest PC, Vanattou-Saifoudine N, Harris LW, Bahn S (2011) Proteomic technologies for biomarker studies in psychiatry: advances and needs. Int Rev Neurobiol 101:65–94
20. Neilson KA, Ali NA, Muralidharan S, Mirzaei M, Mariani M, Assadourian G, Lee A, van Sluyter SC, Haynes PA (2011) Less label, more free: approaches in label-free quantitative mass spectrometry. Proteomics 11:535–553
21. Olsen JV, Andersen JR, Nielsen PA, Nielsen ML, Figeys D, Mann M, Wisniewski JR (2004) HysTag – a novel proteomic quantification tool applied to differential display analysis of membrane proteins from distinct areas of mouse brain. Mol Cell Proteomics 3:82–92
22. Whetstone PA, Butlin NG, Corneillie TM, Meares CF (2004) Element-coded affinity tags for peptides and proteins. Bioconjug Chem 15:3–6
23. Zhang L, Guo YL, Liu HQ (2004) A novel class of chemically modified iodo-containing resins: design, synthesis and application to mass spectrometry-based proteome analysis. J Mass Spectrom 39:447–457
24. Chen P, Nie S, Mi W, Wang XC, Liang SP (2004) De novo sequencing of tryptic peptides sulfonated by 4-sulfophenyl isothiocyanate for unambiguous protein identification using post-source decay matrix-assisted laser desorption/ionization mass spectrometry. Rapid Commun Mass Spectrom 18:191–198
25. Hellman U, Bhikhabhai R (2002) Easy amino acid sequencing of sulfonated peptides using post-source decay on a matrix-assisted laser desorption/ionization time-of-flight mass spectrometer equipped with a variable voltage reflector. Rapid Commun Mass Spectrom 16: 1851–1859
26. Keough T, Youngquist RS, Lacey MP (2003) Sulfonic acid derivatives for peptide sequencing by MALDI MS. Anal Chem 75:156A–165A
27. Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, Aebersold R (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. Nat Biotechnol 17:994–999

28. Ross PL, Huang YN, Marchese JN, Williamson B, Parker K, Hattan S, Khainovski N, Pillai S, Dey S, Daniels S, Purkayastha S, Juhasz P, Martin S, Bartlet-Jones M, He F, Jacobson A, Pappin DJ (2004) Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. Mol Cell Proteomics 3:1154–1169
29. Schmidt A, Kellermann J, Lottspeich F (2005) A novel strategy for quantitative proteomics using isotope-coded protein labels. Proteomics 5:4–15
30. Dayon L, Hainard A, Licker V, Turck N, Kuhn K, Hochstrasser DF, Burkhard PR, Sanchez JC (2008) Relative quantification of proteins in human cerebrospinal fluids by MS/MS using 6-plex isobaric tags. Anal Chem 80: 2921–2931
31. Levin Y, Bahn S (2010) Quantification of proteins by label-free LC-MS/MS. Methods Mol Biol 658:217–231
32. Old WM, Meyer-Arendt K, Aveline-Wolf L, Pierce KG, Mendoza A, Sevinsky JR, Resing KA, Ahn NG (2005) Comparison of label-free methods for quantifying human proteins by shotgun proteomics. Mol Cell Proteomics 4: 1487–1502
33. Li GZ, Vissers JP, Silva JC, Golick D, Gorenstein MV, Geromanos SJ (2009) Database searching and accounting of multiplexed precursor and product ion spectra from the data independent analysis of simple and complex peptide mixtures. Proteomics 9: 1696–1719
34. Bateman RH, Carruthers R, Hoyes JB, Jones C, Langridge JI, Millar A, Vissers JP (2002) A novel precursor ion discovery method on a hybrid quadrupole orthogonal acceleration time-of-flight (Q-TOF) mass spectrometer for studying protein phosphorylation. J Am Soc Mass Spectrom 13:792–803
35. Wang J, Leung D (2007) Analyses of macrolide antibiotic residues in eggs, raw milk, and honey using both ultra-performance liquid chromatography/quadrupole time-of-flight mass spectrometry and high-performance liquid chromatography/tandem mass spectrometry. Rapid Commun Mass Spectrom 21:3213–3222
36. Levin Y, Wang L, Schwarz E, Koethe D, Leweke FM, Bahn S (2010) Global proteomic profiling reveals altered proteomic signature in schizophrenia serum. Mol Psychiatry 15: 1088–1100
37. Herberth M, Koethe D, Levin Y, Schwarz E, Krzyszton ND, Schoeffmann S, Ruh H, Rahmoune H, Kranaster L, Schoenborn T, Leweke MF, Guest PC, Bahn S (2011) Peripheral profiling analysis for bipolar disorder reveals markers associated with reduced cell survival. Proteomics 11:94–105
38. Shadforth IP, Dunkley TP, Lilley KS, Bessant C (2005) i-Tracker: for quantitative proteomics using iTRAQ. BMC Genomics 6:145
39. Colaert N, Barsnes H, Vaudel M, Helsens K, Timmerman E, Sickmann A, Gevaert K, Martens L (2011) Thermo-msf-parser: an open source Java library to parse and visualize thermo proteome discoverer msf files. J Proteome Res 10:3840–3843
40. Wang H, Alvarez S, Hicks LM (2012) Comprehensive comparison of iTRAQ and label-free LC-based quantitative proteomics approaches using two chlamydomonas reinhardtii strains of interest for biofuels engineering. J Proteome Res 11:487–501
41. Martins-de-Souza D, Maccarrone G, Wobrock T, Zerr I, Gormanns P, Reckow S, Falkai P, Schmitt A, Turck CW (2010) Proteome analysis of the thalamus and cerebrospinal fluid reveals glycolysis dysfunction and potential biomarkers candidates for schizophrenia. J Psychiatr Res 44:1176–1189

Chapter 19

Genome-Based Bioinformatic Prediction of Major Histocompatibility Complex (MHC) Epitopes

Simon J. Foote

Abstract

Over the last 12 years, a large amount of knowledge has been accumulated on various aspects of Major Histocompatibility Complex (MHC) molecules. In conjunction, numerous algorithms and tools have been developed to screen protein molecules for these MHC receptor sites. By combining these computational tools and databases with genomic sequence information that is now widely available for a vast range of organisms, it is possible to screen whole genomes for MHC epitopes. By prescreening these genomes, it allows the researcher to narrow down possible protein targets for further analysis by traditional tools such as gene knockouts and animal efficacy studies.

Key words MHC epitope prediction, MHC ligand, T-cell, Antigen

1 Introduction

Antigenic epitopes play a key role in eliciting a T cell response against intracellular and extracellular pathogens including bacteria, viruses, and parasites. T cells contain two receptors, a T cell receptor (TCR) which recognizes antigens and a co-receptor which binds to Major Histocompatibility Complex (MHC) molecules. There are two main types of T cells, cytotoxic T cells and helper T cells. The cytotoxic T cell's co-receptor (CD8) recognizes an MHC Class I:antigen complex and in combination with the TCR binding to the antigen leads to T cell activation and eventual cell death. Helper T cells control the immune response by directing other cells to fight infected cells. The helper T cell's co-receptor (CD4) recognizes an MHC Class II-antigen complex which activates the T cell to release cytokines that activate other cell types such as B cells. The antigen presented on the MHC molecules is typically a peptide fragment and the resulting complex is referred to as a T cell epitope. The antigen peptides are generated via one of two pathways depending upon whether the pathogen is intra or

Kelly M. Fulton and Susan M. Twine (eds.), *Immunoproteomics: Methods and Protocols*, Methods in Molecular Biology, vol. 1061,
DOI 10.1007/978-1-62703-589-7_19,

extracellular. In both pathways, specific proteases are responsible for cleaving the proteins into smaller peptides which then bind to MHC class I or class II molecules. MHC class I molecules accommodate peptides 8–10 amino acids in length and usually contain an allele specific motif sequence with conserved N-terminal and C-terminal residues [1]. MHC class II molecules bind peptides 11–25 amino acids in length. They are more variable in size, but usually contain a 9–10 residue core region [2]. By using this gathered epitope information, multiple algorithms, software tools and databases have been developed to search for MHC ligands.

The aim of this chapter will be to describe a selection of these tools and how they can be used together to search for MHC epitopes in whole genomes of bacteria. As an example of this genome-based analysis, the genome of *Francisella tularensis* subsp. *tularensis* (strain SCHU S4) which leads to typhoidal tularemia upon inhalation, will be analyzed for probable peptides that could bind to MHC class I or class II molecules. It is known that SCHU S4 express antigens unique to *subsp. tularensis* and these might contribute to protective immunity [3]. The research involved the use of deletion mutants of SCHU S4 which showed in mice that a live vaccine strain (LVS) elicited immunity to challenge with subsp. *tularensis* is known to be dependent on the actions of interferon gamma (IFNγ) and CD4+ and CD8+ T cells. Either CD4+ T cells or CD8+ T cells appear to be sufficient for survival, as mice deficient in either T cell subset are capable of resolving both primary and secondary LVS infections. Although this has been shown, the actual epitopes that induce this MHC response have not been identified.

2 Materials

For additional software and databases, *see* Tables 1 and 2.

2.1 MHC Prediction Software

1. Rankpep [4]—predicts peptide binders to MHC-I and MHC-II molecules from protein sequence/s or sequence alignments using Position Specific Scoring Matrices (PSSMs), webform (multi-sequence input), http://imed.med.ucm.es/Tools/rankpep.html.
2. IEDB Analysis Resource [5]—multi-algorithmic depending upon chosen alleles (matrices, artificial neural network, average relative binding), MHC-I or MCH-II, webform and local installation options, http://tools.immuneepitope.org/main/html/tcell_tools.html.

2.2 Databases

1. IEDB—Immune Epitope Database is a database containing curated immunological data for antibody, B and T cell epitopes derived from humans, nonhuman primates, rodents, and other animal species, http://www.immuneepitope.org.

Table 1
A selection of alternative software tools for MHC prediction (*see* Note 1)

Name	Description	Link
ProPred-I [6]	Matrix-based, employs matrices from 47 MHC-I alleles, webform (single sequence)	hhttp://www.imtech.res.in/raghava/propred1
NetMHCpan-2.4 [7]	Web-server predicts binding to any known MHC-I molecule using artificial neural networks, multi-sequence input, also downloadable version	http://www.cbs.dtu.dk/services/NetMHCpan
NetMHCIIpan-2.1 [8]	Web-server predicts binding of peptides to more than 500 HLA-DR human MHC-II alleles using artificial neural networks (ANNs)	http://www.cbs.dtu.dk/services/NetMHCIIpan
TEPITOPEpan [9]	MHC-II matrix-based, human allele specific, webform	http://www.biokdd.fudan.edu.cn/Service/TEPITOPEpan
SYFPEITHI [10]	Matrix-based prediction of MHC-I and MHC-II epitopes, Web-based and downloadable versions	http://www.syfpeithi.de

Table 2
A selection of databases containing MHC ligand and peptide information

Name	Description	Link
MHCBN [11]	Curated database consisting of detailed information about MHC binding and non-binding peptides and T-cell epitopes	http://www.imtech.res.in/raghava/mhcbn
MHCPEP [12]	Curated database of peptides known to bind to MHC molecules	http://bio.dfci.harvard.edu/DFRMLI/HTML/MHCBindingPeptides.php
SYFPEITHI	Database of published MHC ligands and peptide motifs	http://www.syfpeithi.de

2.3 Protein Localization Software

1. PSortB v3.0.2 [13]—bacterial localization prediction, http://www.psort.org/psortb.
2. SignalP v4.0 [14]—predicts presence and location of signal peptide cleavage sites, http://www.cbs.dtu.dk/services/SignalP.
3. SecretomeP v2.0 [15]—predicts nonclassical protein secretion, http://www.cbs.dtu.dk/services/SecretomeP.
4. LipoP v1.0 [16]—predicts lipoproteins and discriminates between lipoprotein signal peptides, other signal peptides and n-terminal membrane helices, http://www.cbs.dtu.dk/services/LipoP.

3 Methods

Whole genome-based analysis for MHC epitopes can be accomplished for any completed genome available from the National Center for Biotechnology Information (NCBI). Files containing the protein information for such genomes can be downloaded from NCBI via ftp. (ftp://ftp.ncbi.nlm.nih.gov/genomes) In this example, the *Francisella tularensis* subsp. *tularensis SCHU S4* genome will be analyzed (NCBI RefSeq = NC_006570).

3.1 Data Gathering and Target Selection

1. Download protein sequences for *Francisella SCHU S4* genome (1,604 proteins) from NCBI and save as a text file (ftp.ncbi.nih.gov/genomes/Bacteria/Francisella_tularensis_SCHU_S4_uid57589/NC_006570.faa) (*see* **Note 2**).
2. Identify MHC alleles to target (*see* **Note 3**).
   For MHC I, use the BALB/c mouse alleles: H-2-Db, H-2-Dd, H-Kb, H-Kd.
   For MHC II, use the mouse alleles: H2-IAd, H2-IEd.

3.2 MHCI and MHCII Analysis with RankPep

1. Fasta files containing up to 100 proteins are used for input (17 files total) (*see* **Note 4**).
2. Set parameters on the webform (http://bio.dfci.harvard.edu/Tools/rankpep.html) as indicated.
   - Select PSSM = chose alleles depending upon whether it is an MHC I or MHC II analysis.
   - Input type = fasta sequences, upload sequences by choosing an input file from **step 1** above.
   - Allowed peptide lengths = All.
   - Binding threshold = 2 %.
   - Proteasome cleavage = on.
3. Save results as html file.
4. Convert the html file to text and filter for top-scoring peptides by using a cutoff of IC50 nM ≤ 65 on the top ranking algorithm for MHC I and IC50 nM ≤ 50 for MHC II (*see* **Note 5**) (*see* Tables 3 and 4 for MHCI and MHCII results respectively).
5. Extract the identified proteins into separate fasta files for further analysis.

3.3 Protein Localization Analysis

For effective epitope presentation, it is important to verify whether top scoring proteins from the above analyses are presented outside the cell, thus being accessible to MHC binding proteins.

1. Generate a fasta file with the sequences from those proteins identified in Subheading 3.2.

Table 3
MHCI peptides identified by RankPep with binding thresholds ≤65 nM along with the corresponding protein localization information as identified by PSortB, SecretomeP, LipoP and SignalP

	RankPep				PSortB	SecretomeP	LipoP	SignalP
Accession	PSSM	Consensus peptide	Optimal score	Bind threshold	Location	Secretion possible	Class	Sig. pep pres
YP_169146.1	10mer_H2_Kd.p.mtx	TYIDGGKGVL	29.131	65.79	Cytoplasmic	No	CYT	No
YP_169174.1	10mer_H2_Kd.p.mtx	TYIGVGILLL	29.978	67.7	Cytoplasmic Membrane	No	TMH	No
YP_169175.1	10mer_H2_Dd.p.mtx	AGGYINYFFL	35.566	67.01	Cytoplasmic Membrane	No	TMH	No
YP_169197.1	9mer_H2_Kd.p.mtx	KYIKENTGL	27.098	66.34	Cytoplasmic	No	CYT	No
YP_169451.1	10mer_H2_Dd.p.mtx	SGPYSLLFIF	36.737	69.22	Cytoplasmic	No	CYT	No
YP_169464.1	8mer_H2_Kd.p.mtx	YYKNYQNL	31.016	72.2	Cytoplasmic	No	CYT	No
YP_169465.1	8mer_H2_Kd.p.mtx	RIPNYAKL	28.29	65.85	Cytoplasmic	No	CYT	No
YP_169519.1	11mer_H2_Kd.p.mtx	LYQPNAVTQNK	31.314	68.02	Periplasmic	No	SpI	Yes
YP_169540.1	10mer_H2_Kd.p.mtx	KYPSFVIQNI	29.332	66.24	Cytoplasmic Membrane	No	TMH	No
YP_169618.1	9mer_H2_Kd.p.mtx	KYICIVCGL	30.341	74.28	Cytoplasmic	No	CYT	No
YP_169630.1	8mer_H2_Kd.p.mtx	YYQFYQDL	32.137	74.81	Cytoplasmic	No	CYT	No
YP_169635.1	9mer_H2_Kd.p.mtx	FYICNNTTL	30.375	74.36	CytoplasmicMembrane	No	TMH	No
YP_169699.1	8mer_H2_Kd.p.mtx	YYHTFTTL	28.129	65.48	CytoplasmicMembrane	No	CYT	No
YP_169751.1	10mer_H2_Kd.p.mtx	SYVDFGITLL	29.636	66.93	OuterMembrane	Yes	CYT	No
YP_169784.1	9mer_H2_Dd.p.mtx	VGPKRRKAL	32.458	77.25	Cytoplasmic	No	CYT	No
YP_169855.1	9mer_H2_Kd.p.mtx	RYLCIPTLI	29.102	71.24	CytoplasmicMembrane	No	TMH	No

(continued)

Table 3
(continued)

	RankPep				PSortB	SecretomeP	LipoP	SignalP
Accession	PSSM	Consensus peptide	Optimal score	Bind threshold	Location	Secretion possible	Class	Sig. pep pres
YP_169902.1	8mer_H2_Kd.p.mtx	YYIFYTVL	31.481	73.28	CytoplasmicMembrane	No	TMH	No
YP_169903.1	10mer_H2_Kd.p.mtx	PYPPFITSTL	29.545	66.72	Cytoplasmic	No	CYT	No
YP_169913.1	8mer_H2_Kd.p.mtx	YILFFEPL	30.741	71.56	Cytoplasmic	No	CYT	No
YP_169938.1	8mer_H2_Kd.p.mtx	YQPDFCLI	29.774	69.3	Cytoplasmic	No	CYT	No
YP_169943.1	9mer_H2_Kd.p.mtx	LYQQMPTSL	28.979	70.94	CytoplasmicMembrane	No	TMH	No
YP_169947.1	8mer_H2_Kd.p.mtx	YIYVFSGL	30.86	71.83	Cytoplasmic	No	CYT	No
YP_169966.1	8mer_H2_Kd.p.mtx	FYIVFFKL	29.282	68.16	CytoplasmicMembrane	No	CYT	No
YP_169988.1	9mer_H2_Kd.p.mtx	RYAKNNEEL	27.72	67.86	Cytoplasmic	No	CYT	No
YP_170040.1	9mer_H2_Kd.p.mtx	LYNQNIELV	26.913	65.89	Cytoplasmic	No	CYT	No
YP_170050.1	8mer_H2_Kd.p.mtx	YYIDFESL	34.041	79.24	Cytoplasmic	No	CYT	No
YP_170132.1	8mer_H2_Kd.p.mtx	YYISYKDI	28.071	65.34	Cytoplasmic	Yes	CYT	No
YP_170216.1	9mer_H2_Kd.p.mtx	KYAQNQHKL	26.662	65.27	OuterMembrane	Yes	SpII	Yes
YP_170227.1	10mer_H2_Kd.p.mtx	EYKPIGKYLL	31.005	70.02	Unknown	Yes	SpII	No
YP_170296.1	8mer_H2_Kd.p.mtx	YIINYKLL	28.236	65.72	Unknown	No	CYT	No

YP_170390.1	10mer_H2_Kd.p.mtx	WYIYPSKIGL	33.593	75.86	CytoplasmicMembrane	No	TMH	No
YP_170399.1	9mer_H2_Kd.p.mtx	KYYKINTQL	28.272	69.21	Unknown	No	SpI	No
YP_170490.1	10mer_H2_Kd.p.mtx	KYPNAIIEGI	28.817	65.08	Cytoplasmic	No	CYT	No
YP_170502.1	8mer_H2_Kd.p.mtx	SIIFFTPL	30.095	70.05	CytoplasmicMembrane	No	CYT	No
YP_170513.1	8mer_H2_Kd.p.mtx	YYPVTAEL	29.661	69.04	Unknown	Yes	CYT	No
YP_170604.1	8mer_H2_Kd.p.mtx	YIINYKLL	28.236	65.72	Unknown	No	CYT	No
YP_170662.1	10mer_H2_Kd.p.mtx	KYPHIPIGLL	29.381	66.35	Cytoplasmic	No	CYT	No
YP_170683.1	9mer_H2_Kd.p.mtx	KYIQLFTQL	28.308	69.3	Cytoplasmic	No	CYT	No

Those highlighted are good candidate epitopes
CYT = cytoplasmic, TMH = transmembrane helix, SpI = signal peptide (signal peptidase I), SpII = lipoprotein signal peptide (signal peptidase II)

Table 4
MHCII peptides identified by RankPep with binding thresholds ≤50 nM along with the corresponding protein localization information as identified by PSortB, SecretomeP, LipoP, and SignalP

	RankPep				PSortB	SecretomeP	LipoP	SignalP
Accession	PSSM	Consensus peptide	Optimal score	Binding threshold	Location	Secretion possible	Class	Signal peptide present
YP_003097805.1	I_Ed.p.mtx	KYFKKQALK	37.937	52.68	Cytoplasmic	No	CYT	No
YP_169178.1	I_Ed.p.mtx	KDKLRNTLK	39.793	55.26	Cytoplasmic	No	CYT	No
YP_169210.1	I_Ad.p.mtx	QMVSAAAAL	32.402	60.97	Cytoplasmic	No	CYT	No
YP_169235.1	I_Ed.p.mtx	KQLDKNTLK	36.584	50.8	Cytoplasmic	No	CYT	No
YP_169256.1	I_Ad.p.mtx	TALHLAAAE	28.526	53.68	Cytoplasmic	No	CYT	No
YP_169363.1	I_Ad.p.mtx	EAVKAAHAA	29.06	54.68	Cytoplasmic	No	CYT	No
YP_169420.1	I_Ad.p.mtx	NAVHGSDAE	27.17	51.12	Extracellular	No	CYT	No
YP_169432.1	I_Ad.p.mtx	KGVSIADPN	28.337	53.32	Cytoplasmic	No	CYT	No
YP_169500.1	I_Ad.p.mtx	TAVSHAQAG	27.347	51.46	Cytoplasmic	No	CYT	No
YP_169504.1	I_Ad.p.mtx	EMISYFQAN	26.843	50.51	Cytoplasmic	No	CYT	No
YP_169572.1	I_Ed.p.mtx	YYVKDNGKK	39.95	55.47	Unknown	No	SpI	No
YP_169652.1	I_Ad.p.mtx	EAFHATLEE	33.037	62.16	Cytoplasmic	No	CYT	No
YP_169653.1	I_Ed.p.mtx	FEQWQNRLK	38.432	53.37	Cytoplasmic	No	CYT	No
YP_169665.1	I_Ed.p.mtx	EYHKQITLE	36.367	50.5	Unknown	No	CYT	No
YP_169711.1	I_Ad.p.mtx	EIVHASFAS	26.6	50.05	Cytoplasmic	No	CYT	No
YP_169903.1	I_Ad.p.mtx	EAIRVTAAN	28.095	52.86	Cytoplasmic	No	CYT	NO
YP_169914.1	I_Ed.p.mtx	KYVFLNRLH	37.497	52.07	Cytoplasmic	No	CYT	No

YP_169960.1	I_Ad.p.mtx	NQLHAAVVE	26.772	50.38	Cytoplasmic	No	CYT	No
YP_169973.1	I_Ed.p.mtx	KVKWQNHLE	38.751	53.81	Unknown	No	CYT	No
YP_169977.1	I_Ed.p.mtx	YVKKDITLK	36.744	51.02	Cytoplasmic	No	SpI	No
YP_170041.1	I_Ad.p.mtx	KAVQATQAN	33.449	62.94	Cytoplasmic	No	CYT	No
YP_170069.1	I_Ed.p.mtx	LRKWQITLK	42.771	59.39	CytoplasmicMembrane	No	TMH	No
YP_170075.1	I_Ad.p.mtx	KAQQVTHAK	28.251	53.16	Cytoplasmic	No	CYT	No
YP_170087.1	I_Ed.p.mtx	HVHKQNTIK	41.962	58.27	Unknown	Yes	SpI	Yes
YP_170105.1	I_Ad.p.mtx	RTIHVAIAN	27.163	51.11	Cytoplasmic	No	CYT	No
YP_170109.1	I_Ed.p.mtx	KLRWQPTIK	36.323	50.44	Cytoplasmic	No	CYT	No
YP_170136.1	I_Ed.p.mtx	KRVKQQRLK	47.847	66.44	Cytoplasmic	Yes	CYT	No
YP_170144.1	I_Ed.p.mtx	KYAKQIAQK	45.31	62.92	Periplasmic	Yes	SpI	No
YP_170148.1	I_Ad.p.mtx	SHVHIKHGE	27.093	50.98	Unknown	No	CYT	No
YP_170159.1	I_Ad.p.mtx	HHISIAQAN	26.595	50.04	Cytoplasmic	No	CYT	No
YP_170185.1	I_Ed.p.mtx	KEMKQNQAK	38.308	53.19	Cytoplasmic	Yes	CYT	No
YP_170193.1	I_Ad.p.mtx	CAVIATKAG	26.686	50.21	Cytoplasmic	No	CYT	No
YP_170212.1	I_Ad.p.mtx	EAVSCHDVN	26.777	50.38	Cytoplasmic	No	CYT	No
YP_170216.1	I_Ed.p.mtx	IVHFQNTLK	46.033	63.92	OuterMembrane	Yes	SpII	Yes
YP_170224.1	I_Ed.p.mtx	KVKKQKTLK	44.177	61.34	Cytoplasmic	No	CYT	No
YP_170268.1	I_Ad.p.mtx	QALHVVDPE	30.635	57.64	Cytoplasmic	No	CYT	No
YP_170278.1	I_Ad.p.mtx	QWIHKALAN	33.97	63.92	Cytoplasmic	No	CYT	No
YP_170369.1	I_Ed.p.mtx	KQHKQPTQK	37.068	51.47	CytoplasmicMembrane	No	CYT	No
YP_170434.1	I_Ed.p.mtx	EYVRRQTLK	40.786	56.64	Cytoplasmic	No	CYT	No

(continued)

Table 4
(continued)

	RankPep				PSortB	SecretomeP	LipoP	SignalP
Accession	PSSM	Consensus peptide	Optimal score	Binding threshold	Location	Secretion possible	Class	Signal peptide present
YP_170467.1	I_Ed.p.mtx	EYRKQNTQN	38.61	53.61	Unknown	Yes	SpI	Yes
YP_170508.1	I_Ed.p.mtx	QHAKQNKLK	38.031	52.81	Cytoplasmic	No	CYT	No
YP_170509.1	I_Ad.p.mtx	QGVSYLPSY	27.372	51.5	OuterMembrane	Yes	SpII	Yes
YP_170651.1	I_Ad.p.mtx	SGVHATLAG	30.03	56.51	CytoplasmicMembrane	No	TMH	No

Those highlighted are good candidate epitopes

2. Use the Web sites described in Subheading 2.3, to analyze the proteins for their localization.
3. Combine these results with those of the RankPep analysis (*see* Table 3 (MHCI) and Table 4 (MHCII)) in order to identify target proteins for further analysis.

3.4 IEDB Tools Analysis for MHCI and MHCII Proteins

This tool can identify further peptides within a given protein that could also be candidate MHC epitopes.

1. For general analysis, go to the Web site: http://tools.immuneepitope.org/main/html/tcell_tools.html and click either Peptide binding to MHC class I or Peptide binding to MHC class II molecules(*see* **Note 6**).
2. Fill in the webform with the required information:
   - A protein sequence as identified above.
   - Prediction method: IEDB_recommended.
   - Specify MHC source organism, alleles and peptide lengths to search.
   - Show output: Percent below (cutoff) = 2.
3. Run analysis to output possible MHC peptides ranked by consensus result with the lower the percentile rank the better the binder (*see* **Note 7**).
4. Repeat **step 2** for each protein that was identified as a possible candidate.

4 Notes

1. There are currently more than 30 prediction tools available for screening peptides for their ability to bind to MHC I and II molecules. These tools make use of a variety of algorithms including binding matrices, artificial neural networks, support vector machines and partial least square function. The majority of the tools classifies the peptides into binders and non-binders and also predicts the binding affinity of the theoretical binders. All of these tools are described more thoroughly elsewhere [17].
2. The downloaded file will contain all the genome's translated sequences in fasta format.
3. Mouse models using BALB/c mice are the standard way to test the immune response to LVS challenges, the identified MHC I alleles for this strain of mice are H-2-Db, H-2-Dd, H-Kb, H-Kd and for MHC II are I_Ad and I_Ed.
4. As for most of the available tools either online or standalone, RankPep has a limit on the number of sequences (maximum

Table 5
Example result from IEDB MHCI analysis of YP_170216.1 (*see* Note 8)

Allele	#	Start	End	Pep Len	Sequence	Method used	% Rank	ANN IC50 (nM)	ANN rank	SMM IC50 (nM)	SMM rank	CombLib. Sidney2008 score	ComLib. Sidney2008 ank
H-2-Kd	1	214	222	9	KYAQNQHKL	Consensus (ANN, SMM, CombLib_Sidney2008)	0.2	38.7	0.2	46.6	0.2	0.0	12.1
H-2-Kd	1	214	222	9	KYAQNQHKL	Consensus (ANN, SMM, CombLib_Sidney2008)	0.2	16.7	0.1	41.8	0.2	2e-05	0.5
H-2-Kd	1	374	382	9	FYQAYYDYI	Consensus (ANN, SMM, CombLib_Sidney2008)	0.2	38.7	0.2	46.6	0.2	0.0	12.1
H-2-Kd	1	374	382	9	FYQAYYDYI	Consensus (ANN, SMM, CombLib_Sidney2008)	0.2	16.7	0.1	41.8	0.2	2e-05	0.5
H-2-Kd	1	81	89	9	KLQVSIGNI	Consensus (ANN, SMM, CombLib_Sidney2008)	0.7	599.2	0.7	160.3	0.6	0.0	4.8
H-2-Kd	1	477	486	10	NYVDPHYDNP	SMM	0.7	0.0	0.0	2,381.3	0.7	0.0	0
H-2-Ld	1	30	38	9	RPKIKAPAL	Consensus (ANN, SMM)	0.8	2,190.8	1.2	133.9	0.4	0.0	0
H-2-Kd	1	188	196	9	AYFSLITAK	Consensus (ANN, SMM, CombLib_Sidney2008)	0.9	1,083.1	0.9	32.7	0.2	3e-05	0.9
H-2-Ld	1	280	288	9	ANIKVPINM	Consensus (ANN, SMM)	0.9	1,651.6	1.1	209.7	0.7	0.0	0
H-2-Kd	1	37	46	10	ALWNNQDKNI	SMM	1	0.0	0.0	2,877.8	1.0	0.0	0
H-2-Kd	1	465	473	9	LYQALGSGY	Consensus (ANN, SMM, CombLib_Sidney2008)	1.1	1,723.4	1.1	277.7	1.1	2e-05	0.6
H-2-Kd	1	377	386	10	AYYDYINTLQ	SMM	1.1	0.0	0.0	3,167.3	1.1	0.0	0
H-2-Kd	1	188	197	10	AYFSLITAKE	SMM	1.2	0.0	0.0	3,292.7	1.2	0.0	0
H-2-Kd	1	188	197	10	AYFSLITAKE	SMM	1.2	0.0	0.0	3,247.8	1.2	0.0	0
H-2-Kd	1	214	223	10	KYAQNQHKLG	SMM	1.2	0.0	0.0	3,292.7	1.2	0.0	0

100 per search) it can analyze at once. Therefore, the complete fasta file will have to be broken up into smaller chunks or individual sequences will have to be used. The simplest way to perform this is to cut and paste the required number into a new text file.

5. The html files require a tool to parse the search results into a tab-delimited format that can be easily filtered. One such online tool can be found here: http://www.webtoolhub.com/tn561393-html-to-text-converter.aspx
6. The Web site only allows one protein at a time to be analyzed. For power users, download the software from http://tools.immuneepitope.org/analyze/html_mhcibinding20090901B/download_mhc_I_binding.html and follow provided installation instructions. Each allele and each length must be searched separately, so by writing a simple script that can loop through each allele and each length available for that allele, batch analysis is possible.
7. To see the full overview of the results, check the expand results box. This shows the IC50 values for each algorithm. For baseline cutoffs, binding affinities (IC50) $\leq$ 50 nM are considered strong binders and $\leq$500 nM medium binders.
8. To download the results, click the download link at the bottom of the results page. The table is truncated due to its length (*see* Table 5).

References

1. Falk K, Rötzschke O, Stevanović S, Jung G, Rammensee HG (1991) Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules. Nature 23(1991): 290–296
2. Chicz RM, Urban RG, Gorga JC et al (1993) Specificity and promiscuity among naturally processed peptides bound to HLA-DR alleles. J Exp Med 178:27–47
3. Hua S, Harris G, Chen W et al (2010) Molecular immune responses to aerosol challenge with francisella tularensis in mice inoculated with live vaccine candidates of varying efficacy. PLoS One 5:e13349
4. Reche PA, Glutting JP, Zhang H, Reinherz EL (2004) Enhancement to the RANKPEP resource for the prediction of peptide binding to MHC molecules using profiles. Immunogenetics 56:405–419
5. Vita R, Zarebski L, Greenbaum JA et al (2010) The immune epitope database 2.0. Nucleic Acids Res 38:D854–D862
6. Singh H, Raghava GPS (2003) ProPred1: prediction of promiscuous MHC class-I binding sites. Bioinformatics 19(8):1009–1014
7. Hoof I, Peters B, Sidney J et al (2009) NetMHCpan, a method for MHC class I binding prediction beyond humans. Immunogenetics 61:1–13
8. Nielsen M, Justesen S, Lund O et al (2010) NetMHCIIpan-2.0 - improved pan-specific HLA-DR predictions using a novel concurrent alignment and weight optimization training procedure. Immunome Res 6:9
9. Zhang L, Chen Y, Wong H et al (2012) TEPITOPEpan: extending TEPITOPE for peptide binding prediction covering over 700 HLA-DR molecules. PLoS One 7:e30483
10. Rammensee J, Bachmann N, Emmerich P et al (1999) SYFPEITHI: database for MHC ligands and peptide motifs. Immunogenetics 50:213–219
11. Bhasin M, Singh J, Raghava GPS (2003) MHCBN: a comprehensive database of MHC

binding and non-binding peptides. Bioinformaatics 19:665–666
12. Brusic V, Rudy G, Kyne AP et al (1996) MHCPEP—a database of MHC-binding peptides: update 1995. Nucleic Acids Res 24:242–244
13. Gardy JL, Spencer C, Wang K et al (2003) PSORT-B: improving protein subcellular localization prediction for Gram-negative bacteria. Nucleic Acids Res 31: 3613–3617
14. Petersen TN, Brunak S, von Heijne G, Nielsen H (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods 8:785–786
15. Bendtsen JD, Kiemer L, Fausbøll A, Brunak S (2005) Non-classical protein secretion in bacteria. BMC Microbiol 5:58
16. Juncker AS, Willenbrock H, von Heijne G (2003) Prediction of lipoprotein signal peptides in Gram-negative bacteria. Protein Sci 12:1652–1662
17. Lin HH, Ray S, Tongchusak S et al (2008) Evaluation of MHC class I peptide binding prediction servers: applications for vaccine research. BMC Immunol 9:8

Chapter 20

Structure-Based Prediction of Major Histocompatibility Complex (MHC) Epitopes

Andrew J. Bordner

Abstract

Because of the enormous diversity of both MHC proteins and peptide epitopes, computational epitope prediction methods are needed in order to supplement limited experimental data. These prediction methods are useful for guiding experiments and have many potential biomedical applications. Unlike popular sequence-based methods, structure-based epitope prediction methods can predict epitopes for multiple MHC types with highly distinct peptide binding propensities. In this chapter, we describe in detail our previously developed structure-based epitope prediction methods for both class I and class II MHC proteins. We also discuss the relative advantages and disadvantages of sequence-based versus structure-based methods and how to evaluate prediction performance.

Key words Peptide docking, Molecular mechanics, Machine learning, Random Forest, Binding affinity

1 Introduction

1.1 Motivation for Computational Models

The binding of small peptide fragments of antigens to class I or class II MHC proteins is an essential step in generating a cytotoxic or helper T cell mediated immune response, respectively. MHC genes, which are usually referred to as HLA (for Human Leukocyte Antigen) in humans, are highly polymorphic, with hundreds to thousands of variants at each locus. Also MHC is polygenic; in humans there are three class I HLA genes (HLA-A,B,C) and six class II genes (α and β chains for each of HLA-DP,DQ,DR), since the peptide-binding domain of class II MHC is composed of two protein subunits. Furthermore, MHC genes manifest co-dominance, in which both the maternal and paternal alleles are equally expressed. Thus, for example, an individual typically has six different class I MHC types expressed. In addition, the peptide fragments bound to the MHC proteins are also extremely diverse since, in principle, there are 20^N possible peptides of length N. The number of possible peptide–MHC combinations emerging

Kelly M. Fulton and Susan M. Twine (eds.), *Immunoproteomics: Methods and Protocols*, Methods in Molecular Biology, vol. 1061, DOI 10.1007/978-1-62703-589-7_20, © Springer Science+Business Media, LLC 2013

from the diversity in both MHC proteins and peptides renders comprehensive experimental measurement infeasible. Computational methods can address this challenge by predicting peptide–MHC binding affinities for as yet uncharacterized combinations. It should be emphasized, however, that almost all epitope prediction methods rely on the limited existing experimental data in order to fit model parameters. Thus, more specifically, computational methods are able to generalize from experimental data and thereby make predictions for new peptides and/or MHC types. Finally, computational methods can guide experiments by suggesting strong binders for an under- or uncharacterized MHC type. Therefore, by closing the loop from computation to experiment and back to computation, an iterative procedure can be employed to continually improve computational models for epitope prediction.

There are many potential biomedical applications of computational T cell epitope prediction. One application is vaccine design, in which epitope prediction can be used to develop a vaccine with broad protective immunity by discovering peptide epitopes from pathogen antigens that bind to multiple common MHC types present in a population. Also, specific class II MHC types have been either positively or negatively associated with different autoimmune diseases [1–9]. Epitope prediction can be applied to discovering possible autoantigens for risk-associated alleles and aid in discovering therapeutics that inhibit peptide–MHC binding. Finally, computational methods can be used to discover epitopes for use in peptide-based allergy immunotherapies, which appear promising [10–16]. These represent just a few examples of possible applications for computational epitope prediction.

1.2 Relative Advantages and Disadvantages of Structure-Based Versus Sequence-Based Approaches

MHC epitope prediction methods can be generally classified as either structure-based or sequence-based. Structure-based methods rely on some type of molecular modeling of the peptide–MHC complex in order to predict strongly binding peptides, or epitopes. This can be divided into two steps: (1) predicting the atomic structure of the peptide–MHC complex and (2) using this structure to predict the peptide binding affinity. On the other hand, sequence-based methods use only the linear amino acid sequence of the peptide to predict its binding affinity to a particular MHC type. This article focuses on structure-based approaches; however, it is useful to first consider the complementary advantages and disadvantages of each approach in order to determine which to use for a given problem. These are summarized in Table 1.

One advantage of sequence-based methods is that they are usually considerably faster than structure-based methods, typically taking only a fraction of a second per prediction while structure-based approaches generally require computationally costly peptide docking. Peptide docking is the procedure for predicting the structure of a peptide–MHC complex using extensive sampling of

Table 1
A comparison of the relative advantages and disadvantages of sequence-based and structure-based approaches to MHC epitope prediction

	Advantages	Disadvantages
Sequence-based approaches	Fast Only require peptide sequences and experimental data	Cannot generalize to dissimilar MHC types
Structure-based approaches	Can generalize to dissimilar MHC types Can potentially generalize to nonstandard peptides	Slow Require prediction of peptide–MHC complex structures

possible conformations for the bound peptide (described in Subheading 3.1). Because of their speed, many online prediction servers implementing sequence-based epitope prediction methods are available (for example the collection of Immune Epitope Database (IEDB) analysis tools [17, 18]). Another advantage of sequence-based approaches is that they do not require any available X-ray structures of the particular MHC type of interest. This is not a significant advantage because representative X-ray structures exist for all MHC loci and all MHC proteins share a common fold so that comparative modeling can be used with structure-based methods to predict epitopes for MHC types without available structures (see discussion below). The primary advantage of structure-based approaches is that they can make predictions for uncharacterized MHC types. This is because the peptide binding affinity prediction depends only on the calculated interaction energy, which is valid for any MHC type. In other words, the laws of physics apply to all MHC types. In contrast, sequence-based approaches work by discovering peptide sequence patterns that correlate with binding affinity and therefore require sufficient binding affinity data (typically ≥1,000 nonredundant peptides) for the MHC type of interest, or a highly similar type for fitting the prediction model. Furthermore, because structure-based methods only rely on the interaction energy they can potentially be adapted to discovering nonstandard epitopes or immune modulators, for example post-translationally modified peptides, peptides with nonnatural residues (like D-amino acids), or even small molecules. This would be impossible for sequence-based methods, which require a standard amino acid sequence as input.

In consideration of these differences between the two approaches, one should use sequence-based methods, preferably via one of the online servers, when epitope predictions are needed for one of the well-characterized MHC types. However, a structure-based approach is the only available option for the many uncharacterized MHC types that are dissimilar enough to have distinct and

unknown peptide binding preferences. For the remainder of this chapter, we will focus on the structure-based approach to MHC epitope prediction.

1.3 Previous Work on Structure-Based Epitope Prediction

Below we will describe our structure-based methods for class I and class II epitope prediction in detail. A number of other structure-based prediction methods have been described in the literature and provide alternative approaches. A non-comprehensive list of these methods is given in Table 2. The reader is referred to the original references for details.

2 Materials

The ICM molecular mechanics program (Molsoft LLC) was used for all peptide docking simulations as well as for viewing the peptide–MHC structures. Because a large number of docking simulations are required for structure-based epitope prediction, a computer cluster that can run at least 100 simulations simultaneously is needed in order to complete this task within a reasonable time frame.

3 Methods

3.1 Predicting the Atomic Structure of the Peptide–MHC Complex

3.1.1 Flexible Peptide Docking

We have adopted a flexible docking approach for predicting the structures of peptide–MHC complexes in which limited readjustments of the peptide backbone is allowed during the molecular mechanics simulation. While a threading approach would be considerably faster, it would also fail to account for these backbone changes. This could reduce structural accuracy, particularly for class I MHC, in which the central backbone conformation is variable. As can be seen in Fig. 2, even though the MHC-contacting portion of the peptide backbone has approximately the same extended conformation there are still small differences in the bound peptides for class II MHC as well. These backbone conformational differences result from different peptide side chains contacting the MHC protein. Therefore one expects that backbone shifts of a similar magnitude will be necessary in order to accurately predict the structures of diverse peptides bound to an MHC protein.

3.1.2 Molecular Mechanics Using the ICM Program

The ICM program (whose name is an acronym for Internal Coordinate Mechanics) performs global optimization of a physical energy function in torsion angle space (i.e., with bond lengths and angles fixed) using an efficient biased probability Monte Carlo search algorithm. In the algorithm, stochastic sampling of torsion angles is biased to favor values observed in X-ray structures and each Monte Carlo move is followed by local optimization. Also, a

Table 2
List of representative structure-based epitope prediction methods for class I and class II MHC

Reference	MHC types examined	Structure prediction method	Scoring method
Class I MHC			
Rognan et al. [41]	HLA-A*0204 and H2-K^k	Homology modeling	Weighted linear combination of interaction energy components
Schueler-Furman et al. [42]	Multiple human and murine class I types	Threading	Empirical residue contact potential
Liu et al. [43]	HLA-A*0201	Monte Carlo simulated annealing	Distance-dependent empirical potential
Antes et al. [44]	HLA-A*0201	Molecular dynamics	Position specific scoring matrix (PSSM) derived from total interaction energies
Knapp et al. [45]	HLA-A*0201	Threading	Three different scoring functions
Yanover and Bradley [46]	Multiple HLA-A and HLA-B types	Loop modeling and all-atom refinement with Rosetta program	PSSMs based on binding energies
Class II MHC			
Davies et al. [47]	Four HLA-DR types	Simulated annealing with AMBER force field using MHC homology models	Overall interaction energy
Schafroth and Floudas [48]	HLA-DRB1*0101	Docked individual side chains into MHC pockets using constrained global optimization	Overall interaction energy
Tong et al. [49]	HLA-DQ1*0301/ DQB1*0302	Four-step docking procedure	Weighted linear combination of interaction energy components
Zhang et al. [50]	HLA-DRB1*0101	Compared three different methods: (1) complex structure prediction using MODELLER, (2) molecular mechanics with explicit water, and (3) no structure prediction but derives scores from available X-ray structures.	Method 1 was scored with a statistical residue pair potential while methods 2 and 3 were used to derive PSSMs for sequence-based prediction
Patronov et al. [51]	HLA-DPA1*0103/ DPB1*0201	Autodock	PSSM derived from the Autodock scoring function
Doytchinova et al. [52]	HLA-DPA1*0103/ DPB1*0201	Molecular dynamics	PSSM derived from interaction energy components

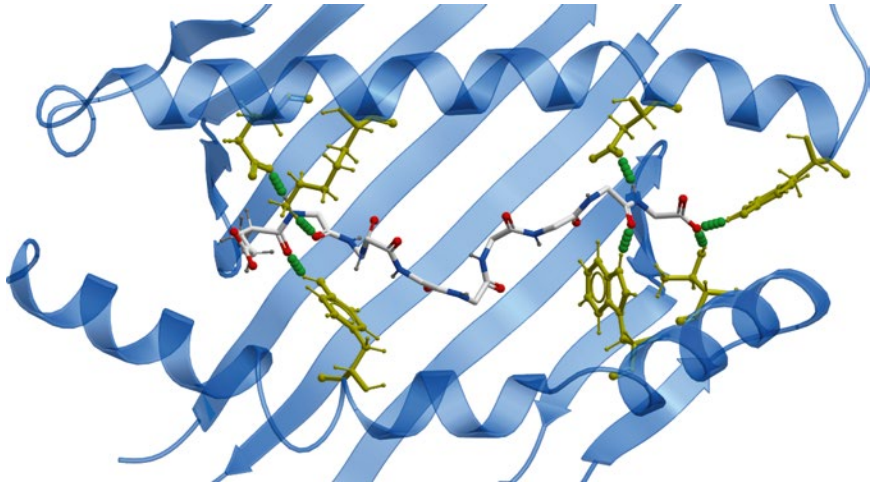


Fig. 1 Conserved hydrogen bonds between the peptide backbone N- and C-termini and residues in the α-helices flanking the peptide-binding cleft of a class I MHC protein (HLA-A*0201, PDB entry 1JF1). These hydrogen bonds allow a particular class I MHC protein to bind to a large repertoire of different peptides and are imposed in the docking procedure described here. The MHC is shown in *blue ribbon* representation and the MHC residues participating in hydrogen bonds in *yellow*. The peptide backbone extends from *left* (N-terminus) to *right* (C-terminus)

nonredundant list of conformations encountered during the simulations are stored in a conformational stack and used to guide the sampling away from unproductive regions of conformational space. The optimization method is described in detail in Abagyan and Totrov 1994 [19].

3.1.3 Structural Features of Peptide Binding to Class I and Class II MHC

Next, we summarize the differences between peptide binding to class I and class II MHC, which are an important consideration in determining which distance restraints are imposed during peptide docking, as explained in the next section. Class I MHC binds short 8–11 residue peptides within a binding cleft that is closed at both ends. The N- and C-terminal segments of the peptide are held in place by conserved hydrogen bonds between the MHC protein and the peptide backbone (*see* Fig. 1) [20]. Because the peptide termini are held in approximately the same positions, longer peptides have a central backbone bulge and consequently more structural variability. In contrast, class II MHC binds longer 15–25 residue peptides within an open binding cleft that allows the peptide termini to extend beyond the MHC protein. Furthermore, the central portion of the peptide contacting the MHC, typically defined to be nine residues in length, assumes a conserved extended polyproline II backbone conformation (*see* Fig. 2) [21]. Interestingly, although class I MHC is one protein subunit while class II MHC is comprised of two subunits (α and β), the structures of their respective peptide-binding domains are remarkably similar.

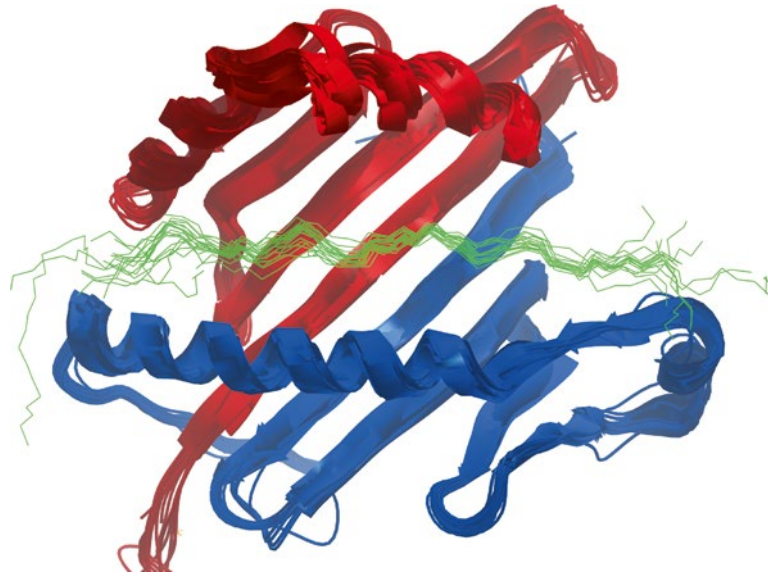


Fig. 2 Structures of 19 human class II HLA proteins complexes for 11 different allotypes (PDB entries taken from table 1 of ref. 23). The protein backbones of the HLA peptide domains were first aligned. The HLA α and β chains are displayed in *blue* and *red*, respectively while the backbones of the bound peptides are *green*. This figure shows the common extended backbone structure for the central portion of each peptide contacting the HLA protein. The peptide backbone structures diverge as they extend from the binding cleft

3.1.4 *Distance Restraints*

As in all molecular docking tasks, it is crucial to allow just enough conformational freedom in the ligand sampling in order to attain the native conformation while limiting the sampling space. Not constraining the ligand enough can result in it wandering away from the binding site. It can also allow the ligand to find an incorrect binding pose that has lower energy than the native one due to inevitable inaccuracies in the energy function used for docking. Obviously, constraining the ligand too much so that it cannot assume that native conformation is a more serious error since it guarantees that the docking simulation will fail.

Because of the differences in peptide binding to class I and class II MHC proteins, outlined in the previous section, we have employed distinct docking constraints that reflect these differences. These constraints are imposed by adding a quadratic restraint penalty term with energy $E_{\text{restraint}} = k\sum_i r_i^2$, in which k is a constant and r_i are the distances between corresponding backbone atoms in the docked peptide and the peptide in the original MHC structure. In other words, we dock the peptide into a receptor structure taken from an X-ray peptide–MHC complex structure and restrain particular segments of the peptide backbone to be similar to those in

the X-ray complex structure throughout the simulation. The sum of the physical energy and restraint energy is minimized during the molecular mechanics simulation.

For class I MHC, we impose constraints with $k = 10$ kcal/(mol Å^2) for nitrogen atoms in residues 1, 2, and M and carbonyl carbon and oxygen atoms in residues 1,2, M-1, and M, in which M is the peptide length [22]. This ensures that the peptide N- and C-terminal backbone segments maintain conserved hydrogen bonds with the MHC protein. For peptide docking to class II MHC, we imposed a weaker quadratic restraint term with $k = 1$ kcal/(mol Å^2) but between all peptide backbone atoms from residue P-1 to residue P9 [23]. This keeps the peptide near the observed consensus backbone conformation observed in X-ray structures of peptide–class II MHC complexes during the Monte Carlo optimization.

3.1.5 Two-Step Docking Procedure

Because of the large number of degrees of freedom (typically 4–7 per residue) in the peptide that must be sampled we employed a two-step docking procedure for computational efficiency. In the first step, an all-atom model of the peptide is docked into a grid potential representation of the MHC peptide-binding domain. Briefly, a grid potential is first precalculated as the interaction energy between a probe atom and the receptor (MHC protein in this case) at all points in a rectilinear three-dimensional grid of points covering the binding pocket. A point spacing of 0.5 Å provided adequate sampling. Five different grid potentials were generated for non-hydrogen atom van der Waals ($E_{\text{C-vdW}}$), hydrogen atom van der Waals ($E_{\text{H-vdW}}$), hydrogen bond (E_{hb}), electrostatics (E_{el}), and hydrophobic solvation (E_{hb}) energy components. Details are given in ref. 24. We found that scaling the hydrogen bond, electrostatic, and hydrophobic grid potentials by multiplicative factors improved docking accuracy. We have reoptimized these weights for the class II MHC docking and recommend these more recent values for both class I and class II MHC docking. The total energy for grid docking was then calculated as sum of the all-atom energy of the peptide, calculated using the ECEPP/3 force field, and the interaction energy of the peptide with the MHC calculated using the scaled grid potentials, or

$$E_{\text{total}} = E_{\text{C-vdw}} + E_{\text{H-vdw}} + 0.5E_{\text{hb}} + 5.0E_{\text{el}} + 2.0E_{\text{hp}} + E_{\text{peptide}} + E_{\text{restraint}}. \quad (1)$$

Importantly, in order to avoid numerical instabilities and reduce steric clashes, a smooth van der Waals energy term that approaches a finite energy at zero atomic separation (7 kcal/mol for our simulations) was used. The grid potential docking simulation was then performed by first minimizing only the restraint potential in order to approximately position the peptide backbone following by ICM Monte Carlo sampling for a total of 5×10^7 energy evaluations using a temperature of 700 K. Example ICM scripts for calculating the grid potentials and performing the grid-based docking are given in **Notes 1** and **2**, respectively.

The second step of the docking procedure is all-atom refinement and reranking of the lowest energy potential docking solutions from the first step using a more accurate energy function. We chose to rerank the top 50 solutions from the first step, which were saved in the ICM conformational stack. This was accomplished by replacing the grid potentials by an all-atom model of the MHC and then performing local optimization of the total energy for the peptide–MHC complex. A more accurate generalized Born or Poisson-Boltzmann electrostatics term combined with a nonpolar solvation term proportional to the solvent-accessible surface area (SASA) with a constant of 12 kcal/(mol Å^2) was used. This resulted in only a small change in the structure of the docked complex that mostly relieved small steric clashes between overlapping atoms. The other main purpose of this step was to calculate interaction energy components using the most accurate methods available in the modeling program for later use in binding affinity prediction. **Note 3** shows the ICM commands used for this step of the docking procedure.

Convergence is an important consideration in the initial grid docking step. Running the docking simulations too short will yield unconverged and inaccurate results while running them too long will waste computational resources and take longer to complete. The optimal number of Monte Carlo iterations required for convergence may be determined by running multiple independent long runs for the same complex and then finding the minimum number of iterations at which the lowest energies attained up to that point in all runs are within a narrow range (typically ~0.5 kcal/mol). Because changing the docking protocol may affect convergence, the optimal simulation length should be reestimated after any major revisions.

3.1.6 Alternative Receptor Conformations

A key challenge in any docking simulation, whether they are small molecule drugs or peptides, is accounting for protein flexibility at the binding site due to induced fit. The grid potential representation of the MHC peptide-binding site described above is static. Although it does implicitly allow limited flexibility because it is smoother than the all-atom potential, particularly for van der Waals interactions, it cannot account for significant conformational changes in the binding site. However, because grid potentials are precalculated, they have a significant advantage over an all-atom energy function because they are much faster to evaluate. One technique to account for receptor flexibility while maintaining the computational efficiency of grid potentials is to dock a peptide to multiple conformations of the receptor and then select the lowest energy docking solution. While this has not been widely tested for peptide–MHC interactions we did find that it yielded good results for peptide docking to HLA-A*0201 [22]. Structural superposition of the MHC peptide-binding domain in all available peptide-HLA-A*0201 complex X-ray structures showed that the binding cleft

residue conformations fell into two different groups characterized by alternate rotamer conformations of residues R97 and Y116. We performed the peptide docking procedure described in the previous section but docked each peptide into representative structures from each of the conformational groups, namely Protein Data Bank (PDB) [25] entries 1JF1 and 1I7U, and then selected the lowest energy docking solution. This approach can be applied to other MHC types for which multiple experimental structures with different bound peptides are available.

3.1.7 Assessing the Accuracy of the Predicted Complex Structures Through Cross-Docking

The peptide–MHC binding prediction consists of two steps: (1) predicting the structure of the peptide–MHC complex and (2) calculating the binding affinity from the predicted structure. Accurate prediction of the complex structure is crucial for success because an incorrect structure will yield an incorrect binding affinity result, no matter how accurate step 2 is. Therefore we recommend first thoroughly validating the peptide docking procedure. The most straightforward method is by cross-docking, in which a peptide from a given peptide–MHC complex structure (structure A) is docked into an MHC structure from another complex with a different peptide bound (structure B). The final docking solution is then compared with the original experimental structure (structure A). This procedure accounts for peptide binding induced changes in the contacting MHC residues and thus is a more realistic estimate of actual prediction performance than self-docking, in which the peptide is redocked into a structure of an MHC bound to the same peptide. In self-docking, the MHC side chains are already preorganized for binding that particular peptide, which is not the case in actual practice. The quality of the cross-docking solutions can then be determined by root-mean-square deviation (RMSD) or fraction of native atomic contacts. The all-atom RMSD should be calculated only for the peptide residues contacting the MHC since the outward-facing residues are relatively unconstrained. Also, because the N- and C-terminal peptide segments are restrained in docking to class I MHC, the backbone RMSD is most meaningful for the central unrestrained segment. Besides evaluating docking performance, statistics from cross-docking results for many structures can be used to optimize the docking protocol.

One caveat of assessing docking results by comparison with X-ray structures of complexes is that the experimental structures may deviate from the structure of an isolated peptide–MHC complex. First, X-ray structures with resolutions >2.0 Å may contain local errors [26]. When multiple structures of the same complex are available, it is usually best to select the one with the highest resolution. Also, some side chains may have alternate conformations in the structure, which are often included in the PDB file. Additionally, interactions with other proteins, such as T cell receptors or symmetry-related chains, can alter the bound peptide conformation.

Finally, as mentioned above, peptide side chains facing away from the MHC protein are particularly flexible and their conformations in the X-ray structure are likely to be influenced by interactions with surrounding molecules in the crystal. They should therefore be omitted from RMSD calculations between predicted and experimental structures.

3.1.8 Homology Modeling of MHC Proteins

Despite the fact that there are thousands of different HLA types, X-ray structures are currently available for only 22 class I and 13 class II HLA types. Fortunately, HLA proteins share high sequence similarity within each locus. Most of these differences occur within the peptide-binding pocket, which is the only portion relevant for epitope prediction. Furthermore, X-ray structures are available for all class I and class II loci (HLA-A,B,C and HLA-DP,DQ,DR). This enables the construction of homology models based on confident sequence alignments for most HLA types using available experimental structures as templates. Peptide docking into homology models has not been thoroughly explored and therefore remains a topic of active research. Previously, we examined docking accuracy using HLA homology models by building models for five different HLA-B types and comparing the results with X-ray structures. The results from that limited study showed rather uneven performance with docking to some models achieving good accuracy, for example <1 Å average central backbone RMSD for two types, while other types had considerably less accurate results. It is likely that implementing some method to account for binding site flexibility [27] would improve the reliability of docking to HLA homology models.

3.2 Predicting the Peptide–MHC Binding Affinity

3.2.1 Machine Learning Methods for Predicting the Binding Affinity

Once the structure of the peptide–MHC complex has been predicted, the final step in epitope prediction is to utilize this structure to predict the binding affinity, which is sometimes referred to as scoring. Because the main goal is usually to differentiate peptides into binders and non-binders rather than predicting numerical binding affinity values it is recommended to use binary prediction (binder versus non-binder) rather than regression to predict binding affinities. Of course, a scoring method optimized for binary prediction generally performs better at this task than binary predictions made from a regression method using a cutoff to separate the two classes. This is also the approach taken by many sequence-based epitope prediction methods. Hereafter, by binding affinity prediction we are referring this binary prediction.

Because the main advantage of structure-based epitope prediction is the universal applicability of a prediction model to any MHC type, the input data for the binding affinity prediction should represent physical interactions contributing to binding and not peptide sequence patterns, which are only relevant for a particular MHC type. The input data we use includes the interaction

Table 3
Features calculated from the peptide docking solution that are used as input data for predicting peptide binding affinity

Input data feature
van der Waals energy ("vw,14")
Hydrogen bond energy ("hb")
Electrostatics energy ("el")
Nonpolar solvation energy ("sf")
Side chain entropy ("en")
Empirical potentials
20 residue type counts

The first five quantities are interaction energy components calculated in ICM, with the term name given in *parentheses*

energy components calculated from the peptide–MHC docking solution listed in Table 3. These energies are calculated after relaxing the complex structure through local geometry optimization in torsion angle space using ICM. We also found it useful to include empirical potentials, such as the Betancourt–Thirumalai contact potential [28] and DFIRE-SCM potential [29], since these may capture energetic contributions to binding that are not well reproduced by the force field terms. We also include 20 residue counts, which are the number of each residue type appearing in the peptide. This is physically motivated by the contribution of the unbound peptide to the total binding affinity via a random coil model in which each residue makes an additive contribution based on its type.

We have chosen to use supervised machine learning in order to predict peptide–MHC binding affinities. Supervised machine learning is able to fit a nonlinear prediction model based on a limited set of training examples for both classes, which in this case are binding and non-binding peptides for a particular MHC type. We have chosen an IC_{50} cutoff of 500 nM to classify peptides as either binders or non-binders. There are many different machine learning algorithms that could be applied to this problem including Support Vector Machines (SVMs), Random Forests, and Artificial Neural Networks (ANNs). Recently we have applied Random Forests [30] to class II MHC epitope prediction [23] because it generally yields prediction performance comparable to competing methods without requiring as much parameter tuning (there is essentially only one parameter to adjust: the number of variables in each classification tree) or data normalization. Furthermore, the contribution of each feature to the overall

prediction accuracy can be readily estimated from the input data [30]. We used the convenient implementation of Random Forests in the randomForest package in the R statistical language [31] for this purpose, which is available from the CRAN Web site (http://cran.r-project.org). Training data consists of a nonredundant set of peptides with experimentally measured binding affinities to a particular MHC type. We chose the types with the most available binding data, HLA-A*0201 for class I MHC and HLA-DRB1*0101 for class II MHC. Because there are many peptides in these sets that differ at only one or two amino acid positions, it is a good idea to remove redundant sequences by, for example, clustering by sequence similarity and selecting one representative sequence from each cluster. Furthermore, the training data sets should contain similar numbers of binder and non-binder examples since machine learning methods generally perform best on a balanced training data set.

Class II epitope prediction has an additional complication compared with class I epitope prediction, the peptide segment binding to the MHC protein is *a priori* unknown. This is easy to account for in the prediction by docking all possible 9-mer peptide segments to the MHC, scoring each one, and predicting the peptide as a binder if any of its 9-mer segments is predicted to be a binder and as a non-binder otherwise. Rather the difficulty comes in training the model because the experimental data only represent the binding affinity of each complete peptide and not which portion of it predominantly binds to the MHC protein. We have adopted a bootstrap approach in which sequence-based predictions for 9-mer peptide fragments are used to train the machine learning classifier. First, binding affinity predictions are made for all possible 9-mer fragments of peptides in the nonredundant set using the same IC_{50} cutoff as for complete peptides (500 nM). Because there are considerably more non-binding fragments, a balanced training set is then compiled by adding an equal number of randomly selected non-binding fragments to all predicted binding fragments. The machine learning classifier, e.g., Random Forest, is next trained on the properties listed in Table 3, which were calculated from docking solution for the peptide fragments.

3.2.2 Assessing Prediction Performance

Experimental data for training and validation can be obtained from a number of sources. One of the largest databases of peptide binding affinity data for both class I and class II MHC types is the Immune Epitope Database [32] (http://www.immuneepitope.org/). MHCBN [33] and AntiJen [34] are two other online databases of peptide–MHC binding affinities. Smaller data sets comprised of overlapping peptides from a small set of antigens are arguably less biased and thus particularly suited for testing. For example, such data sets for multiple alleles are available from ref. 35 or the online Dana-Farber repository [36] (http://bio.dfci.harvard.edu/DFRMLI/).

The prediction performance for epitope prediction can be estimated using cross-validation. This involves dividing the available data set (peptides with known binding affinities for a particular MHC type) into approximately equal size subsets (typically 5–10) and making predictions for each subset in turn using a prediction model trained on the remaining subsets. Thus at the end of this procedure, predictions have been made for all of the peptides in the data set using models trained on a nonoverlapping set of peptides. Because this mimics an actual prediction scenario, in which predictions are made for novel peptide sequences not used to train the model, it provides a realistic estimate of how the prediction method will perform in practice. Test statistics give a useful summary of prediction performance on the test set. The area under the ROC curve (AUC) is a common choice because it is independent of the inevitable tradeoff between specificity and sensitivity due to, for example, varying the score cutoff.

One key concern for all computational epitope prediction algorithms that are applied to the same MHC type as used for training is redundancy in the training data [37]. Even a set of unique peptide sequences taken directly from a database typically have many highly similar peptides that share all but one or two residues, which could lead to overestimation of prediction performance assessed by, for example, cross-validation. In order to avoid overestimation of prediction accuracy, peptide sequence similarity between the training and test sets must be reduced. Similarity within either the training or test set would not bias prediction performance estimates as much except to reduce the effective size of the data sets. Redundancy is particularly a problem with class II MHC data since similar epitopes may be located at different positions within peptides having lower overall sequence identity. A simple solution is to use one of the redundancy-reduced cross-validation sets compiled for testing various sequence-based prediction methods [38–40]. Finally, redundancy is not a problem for structure-based methods that are trained on data for an MHC type at a different locus than that used for testing, for example making predictions for an HLA-DQ type using a model trained on HLA-DR1*0101.

One completely independent but small data set for validation of class II MHC epitope prediction are the available X-ray structures of peptide–MHC complexes in the PDB. The 9-mer fragment contacting the MHC, which presumably constitutes the primary epitope, can be readily observed in the experimental structures. This information can then be used to evaluate the performance for the independent prediction task of finding which 9-mer fragment of the peptide binds most strongly to the canonical binding site in the MHC protein. One such data set is given Table 1 of ref. 23.

3.3 Conclusions and Future Directions

We have described our computational approach for structure-based epitope prediction for both class I and class II MHC proteins. The computational procedure is comprised of two steps: constrained molecular mechanics docking of the peptide to the MHC structure followed by machine learning based prediction of binding affinity. Although this is a structure-based method, it still exploits binding data for a particular MHC type (the one with most available binding data) in order to train the machine learning algorithm used in step 2.

There are several possible directions for extension or improvement of the described method. One possibility is to use structure-based predictions for a series of peptides containing all possible residue types at each binding pocket in order to fit a PSSM, as was done in several of the published methods listed in Table 2. This would dramatically reduce the time required for prediction since docking simulations would only be needed to precalculate the PSSM. It is unclear, however, how much the approximation that each peptide residue makes an additive contribution to the overall binding affinity, as is implicitly assumed with a PSSM, would adversely affect prediction performance. Another important area for further study, which was discussed above, is to optimize the docking procedure for homology models since most MHC types must be modeled on existing X-ray structures. Finally, further testing, comparison, and optimization of structure-based methods using common data sets compiled from the continually growing amount of peptide–MHC binding data is expected to lead to improved accuracy for these methods in the future.

4 Notes

1. Sample ICM script to calculate the grid potentials representing the MHC peptide-binding cleft for class II MHC epitope prediction

```
# Prepare the X-ray structure of the peptide-MHC (commands not shown)
# Align peptide-MHC complex so that binding cleft is aligned along coordinate
# axes.
# This insures that the box used to generate the grid potentials is as small as
# possible.
# If needed, rename the chains in the PDB file so that α and β MHC chains are
# "A" and "B", respectively and the peptide is chain "C".
# Also, need to renumber the peptide residues to conform to standard
# values, i.e. residue #1 for P1
# Use the convertObject macro to convert the structure into an ICM object
# with internal coordinates, which is required for docking.

# Read the precalculated object prepared as described above.
read object "complex.ob"
```

```
# Calculate a box that encloses the peptide-binding cleft in the MHC.
# This box should be as small as possible to just contain the entire
# flexible peptide.
dockBox=Box(Sphere(a_.c//* a_.a,b//* 7.0),7.0)
adj_dist=(40.0-Abs(dockBox[6]-dockBox[3]))/2.0
dockBox=dockBox+(0.//0.//-adj_dist//0.//0.//adj_dist)

# Calculate the grid potential maps within the box using a 0.5 Å spacing.
make map potential "gh,gc,ge,gb,gs,sf" dockBox 0.5

# Write the hydrogen van der Waals (gh), non-hydrogen van der Waals (gc),
# hydrogen bond (gb), electrostatics (ge), hydrophobic (gs) grid potentials
# to files for later use in docking.
write m_gh gh_map_file delete
write m_gc gc_map_file delete
write m_gb gb_map_file delete
write m_ge ge_map_file delete
write m_gs gs_map_file delete
```

2. Sample ICM script to dock the peptide into the MHC grid potentials for class II MHC epitope prediction

```
# Read the ICM format structure of the peptide-MHC complex
read object "complex.ob" name="complex"
# Create a peptide structure with the amino acid sequence specified by
# "peptide_seq"
build string IcmSequence(peptide_seq,"nterm","coo-") name="pep"
# Assign standard residue numbers to the peptide (-pep_index is the first
# residue number)
align number a_pep. -pep_index
set object a_pep.
# Set quadratic constraints
set tether a_pep./-1:9/n,ca,c,o a_complex.c/-1:9/n,ca,c,o
free_backbone_vars=v_pep./-999:-2/phi,PSI  | v_pep./10:999/phi,PSI
# Read precalculated grid potentials for peptide-binding cleft of MHC
read map gh_map_file name="m_gh"
read map gc_map_file name="m_gc"
read map gb_map_file name="m_gb"
read map ge_map_file name="m_ge"
read map gs_map_file name="m_gs"
# Multiply grid potentials by weights given in Equation 1
m_gb=0.5*m_gb
m_ge=5.0*m_ge
m_gs=2.0*m_gs

# set quadratic restraint weight k=1.0 kcal/(mol Å^2)
tzWeight=1.0
# Use smoothed vdW potential with maximum value of 7 kcal/mol at zero
# separation.
vwMethod="soft"
vwSoftMaxEnergy=7.0
# Include all intramolecular energy terms for the peptide, peptide-MHC
# interaction energy grid potentials, and the quadratic restraint energy
# for the peptide backbone
set terms only "gb,gh,gc,ge,gs,en,vw,14,hb,to,el,tz"
```

```
mncallsMC=50000000  # total number of function calls=5×10^7
mncalls=2000  # up to 2000 steps for local optimization
mnconf=200  # number of conformations in the conformational stack
temperature=700  # Monte Carlo sampling temperature=700K
# If the Monte Carlo procedure rejects>10 consecutive trial moves then
# randomly shift all peptide torsion variables by up to 30°.
mnreject=10
rejectAction="random"
set symmetry exact  # needed when comparing by torsion angles
vicinity=30.0  # torsion angle cutoff for assigning conformations to the
# stack
set vrestraint a_pep./* # use biased rotamer angle sampling for efficiency
# Perform Monte Carlo sampling of all peptide side chains and unrestrained
# backbone angles
MC_vars=v_pep.//xi* | free_backbone_vars
# Perform local optimization of all peptide torsion angles after each
# Monte Carlo move
local_min_vars=v_pep.//*
# Perform Monte Carlo sampling
montecarlo MC_vars local_min_vars
store stack a_pep.  # Put the conformational stack into the ICM object for
# the peptide
# Write the peptide structure along with the conformational stack to a file
write object "grid_docking_results.ob"
```

3. Sample ICM script to perform all-atom refinement of grid potential docking solutions and calculation of interaction energy components for class II epitope prediction

```
# Read the ICM format structure of the peptide-MHC complex
read object "complex.ob" name="complex"
# Find MHC residues in the α and β chains contacting the peptide
orig_interface_res=Res(Sphere(a_complex.c/-1:9 a_$pdb_name.a,b 4.0))
interface_a_res_nums=Split(String(orig_interface_res & a_complex.a),"/")[2]
interface_b_res_nums=Split(String(orig_interface_res & a_complex.b),"/")[2]

# Read the structure containing the conformational stack from the grid
# potential docking
read object "grid_docking_results.ob" name="pep"
load stack a_pep.

# Calculate residue type counts for the peptide
aa_str="ACDEFGHIKLMNPQRSTVWY"
pep_seq=String(Sequence(a_pep.))
residue_counts=Iarray(20)
for i=1, Nof(pep_seq)
 residue_counts[i]=Nof(pep_seq,aa_str[i])
endfor
print "Residue type counts=",residue_counts

# Assign the same quadratic restraints used for grid potential docking
set tether a_pep./-1:9/n,ca,c,o a_complex.c/-1:9/n,ca,c,o
free_backbone_vars=v_pep./-999:-2/phi,PSI  | v_pep./10:999/phi,PSI

tzWeight=1.0
vwMethod="soft"
vwSoftMaxEnergy=7.0
set terms only "en,vw,14,hb,to,el,tz,sf"
```

```
mncalls=2000
# Use generalized Born method for electrostatics energy
electroMethod="generalized Born"
# Use hydrophobic solvation energy proportional to the SASA with
# constant=0.012 kcal/(mol Å2)
surfaceMethod="constant tension"
surfaceTension=0.012
for i=1,n_confs  # Iterate over all conformations in the stack
 load conf i
 set symmetry exact  # needed when comparing by torsion angles
 set vrestraint a_pep./*  # use biased rotamer angle sampling for efficiency
 interface_res=a_pep.a/$interface_a_res_nums | a_pep.b/$interface_b_res_nums |
a_pep.c
 local_min_vars=(v_//* & interface_res) & !(v_//phi,psi,omg & a_pep.a,b)
 make born  # calculate Born radii
 # Perform 10 iterations of local optimization to ensure convergence
 for j=1,10
  minimize local_min_vars
 endfor
 # Calculate physical interaction energy terms
 show energy a_pep.* mute
 make born  # Calculate Born radii for peptide bound to MHC
 E_all_tot=Energy("vw,14,hb,el,to,en,sf")  # all terms except tether (tz)
 printf "conf=%d, total energy=%f\n",i,E_all_tot
 E_vw_tot=Energy("vw,14")
 E_hb_tot=Energy("hb")
 E_el_tot=Energy("el")
 E_to_tot=Energy("to")
 E_tz_tot=Energy("tz")
 E_en_tot=Energy("en")
 E_sf_tot=Energy("sf")
 make born a_pep.c a_pep.c # Calculate Born radii for peptide only
 show energy a_pep.c a_pep.c mute
 E_all_pep=Energy("vw,14,hb,el,to,en,sf")
 E_vw_pep=Energy("vw,14")
 E_hb_pep=Energy("hb")
 E_el_pep=Energy("el")
 E_to_pep=Energy("to")
 E_tz_pep=Energy("tz")
 E_en_pep=Energy("en")
 E_sf_pep=Energy("sf")
 make born a_pep.a,b a_pep.a,b # Calculate Born radii for MHC only
 show energy a_pep.a,b a_pep.a,b mute
 E_all_MHC=Energy("vw,14,hb,el,to,en,sf")
 E_vw_MHC=Energy("vw,14")
 E_hb_MHC=Energy("hb")
 E_el_MHC=Energy("el")
 E_to_MHC=Energy("to")
 E_tz_MHC=Energy("tz")
 E_en_MHC=Energy("en")
 E_sf_MHC=Energy("sf")
 E_all_diff=E_all_tot - E_all_pep - E_all_MHC
 E_vw_diff=E_vw_tot - E_vw_pep - E_vw_MHC
```

```
E_hb_diff=E_hb_tot - E_hb_pep - E_hb_MHC
E_el_diff=E_el_tot - E_el_pep - E_el_MHC
E_to_diff=E_to_tot - E_to_pep - E_to_MHC
E_en_diff=E_en_tot - E_en_pep - E_en_MHC
E_sf_diff=E_sf_tot - E_sf_pep - E_sf_MHC
# Calculate Betancourt-Thirumalai and DFIRE empirical potentials: E_BT
# and E_DFIRE (commands not shown)
# Print out interaction energies
# Select conformation with lowest total interaction energy (E_all_diff)
print i,E_all_diff,E_vw_diff,E_hb_diff,E_el_diff,E_to_diff,E_en_diff, E_sf_
diff,E_BT,E_DFIRE
endfor  # for i=1,n_confs
```

Acknowledgment

This work was supported by the Mayo Clinic.

References

1. Todd JA, Bell JI, McDevitt HO (1987) HLA-DQ beta gene contributes to susceptibility and resistance to insulin-dependent diabetes mellitus. Nature 329:599–604
2. Todd JA, Wicker LS (2001) Genetic protection from the inflammatory disease type 1 diabetes in humans and animal models. Immunity 15:387–395
3. Baisch JM, Weeks T, Giles R, Hoover M, Stastny P, Capra JD (1990) Analysis of HLA-DQ genotypes and susceptibility in insulin-dependent diabetes mellitus. N Engl J Med 322:1836–1841
4. Wordsworth BP, Lanchbury JS, Sakkas LI, Welsh KI, Panayi GS, Bell JI (1989) HLA-DR4 subtype frequencies in rheumatoid arthritis indicate that DRB1 is the major susceptibility locus within the HLA class II region. Proc Natl Acad Sci U S A 86:10049–10053
5. Fogdell A, Hillert J, Sachs C, Olerup O (1995) The multiple sclerosis- and narcolepsy-associated HLA class II haplotype includes the DRB5*0101 allele. Tissue Antigens 46: 333–336
6. Oksenberg JR, Barcellos LF, Cree BA, Baranzini SE, Bugawan TL, Khan O, Lincoln RR, Swerdlin A, Mignot E, Lin L, Goodin D, Erlich HA, Schmidt S, Thomson G, Reich DE, Pericak-Vance MA, Haines JL, Hauser SL (2004) Mapping multiple sclerosis susceptibility to the HLA-DR locus in African Americans. Am J Hum Genet 74:160–167
7. Sollid LM, Markussen G, Ek J, Gjerde H, Vartdal F, Thorsby E (1989) Evidence for a primary association of celiac disease to a particular HLA-DQ alpha/beta heterodimer. J Exp Med 169:345–350
8. Matsuki K, Grumet FC, Lin X, Gelb M, Guilleminault C, Dement WC, Mignot E (1992) DQ (rather than DR) gene marks susceptibility to narcolepsy. Lancet 339:1052
9. Mignot E, Lin L, Rogers W, Honda Y, Qiu X, Lin X, Okun M, Hohjoh H, Miki T, Hsu S, Leffell M, Grumet F, Fernandez-Vina M, Honda M, Risch N (2001) Complex HLA-DR and -DQ interactions confer risk of narcolepsy-cataplexy in three ethnic groups. Am J Hum Genet 68:686–699
10. Muller U, Akdis CA, Fricker M, Akdis M, Blesken T, Bettens F, Blaser K (1998) Successful immunotherapy with T-cell epitope peptides of bee venom phospholipase A2 induces specific T-cell anergy in patients allergic to bee venom. J Allergy Clin Immunol 101:747–754
11. Maverakis E, Beech J, Stevens DB, Ametani A, Brossay L, van den Elzen P, Mendoza R, Thai Q, Macias LH, Ethell D, Campagnoni CW, Campagnoni AT, Sette A, Sercarz EE (2003) Autoreactive T cells can be protected from tolerance induction through competition by flanking determinants for access to class II MHC. Proc Natl Acad Sci U S A 100: 5342–5347

12. Marcotte GV, Braun CM, Norman PS, Nicodemus CF, Kagey-Sobotka A, Lichtenstein LM, Essayan DM (1998) Effects of peptide therapy on ex vivo T-cell responses. J Allergy Clin Immunol 101:506–513
13. von Garnier C, Astori M, Kettner A, Dufour N, Heusser C, Corradin G, Spertini F (2000) Allergen-derived long peptide immunotherapy down-regulates specific IgE response and protects from anaphylaxis. Eur J Immunol 30: 1638–1645
14. Haselden BM, Kay AB, Larche M (1999) Immunoglobulin E-independent major histocompatibility complex-restricted T cell peptide epitope-induced late asthmatic reactions. J Exp Med 189:1885–1894
15. Oldfield WL, Larche M, Kay AB (2002) Effect of T-cell peptides derived from Fel d 1 on allergic reactions and cytokine production in patients sensitive to cats: a randomised controlled trial. Lancet 360:47–53
16. Larche M (2006) Immunoregulation by targeting T cells in the treatment of allergy and asthma. Curr Opin Immunol 18:745–750
17. Kim Y, Ponomarenko J, Zhu Z, Tamang D, Wang P, Greenbaum J, Lundegaard C, Sette A, Lund O, Bourne PE, Nielsen M, Peters B (2012) Immune epitope database analysis resource. Nucleic Acids Res 40:W525–W530
18. Kim Y, Ponomarenko J, Zhu Z, Tamang D, Wang P, Greenbaum J, Lundegaard C, Sette A, Lund O, Bourne PE, Nielsen M, Peters B (2012) IEDB T cell epitope prediction tools. http://tools.immuneepitope.org/main/html/tcell_tools.html.
19. Abagyan R, Totrov M (1994) Biased probability Monte Carlo conformational searches and electrostatic calculations for peptides and proteins. J Mol Biol 235:983–1002
20. Madden DR (1995) The three-dimensional structure of peptide–MHC complexes. Annu Rev Immunol 13:587–622
21. Jardetzky TS, Brown JH, Gorga JC, Stern LJ, Urban RG, Strominger JL, Wiley DC (1996) Crystallographic analysis of endogenous peptides associated with HLA-DR1 suggests a common, polyproline II-like conformation for bound peptides. Proc Natl Acad Sci U S A 93:734–738
22. Bordner AJ, Abagyan R (2006) Ab initio prediction of peptide-MHC binding geometry for diverse class I MHC allotypes. Proteins 63: 512–526
23. Bordner AJ (2010) Towards universal structure-based prediction of class II MHC epitopes for diverse allotypes. PLoS One 5:e14383
24. Fernandez-Recio J, Totrov M, Abagyan R (2002) Soft protein-protein docking in internal coordinates. Protein Sci 11:280–291
25. Rose PW, Beran B, Bi C, Bluhm WF, Dimitropoulos D, Goodsell DS, Prlic A, Quesada M, Quinn GB, Westbrook JD, Young J, Yukich B, Zardecki C, Berman HM, Bourne PE (2011) The RCSB protein data bank: redesigned web site and web services. Nucleic Acids Res 39:D392–D401
26. Acharya KR, Lloyd MD (2005) The advantages and limitations of protein crystal structures. Trends Pharmacol Sci 26:10–14
27. B-Rao C, Subramanian J, Sharma SD (2009) Managing protein flexibility in docking and its applications. Drug Discov Today 14:394–400
28. Betancourt MR, Thirumalai D (1999) Pair potentials for protein folding: choice of reference states and sensitivity of predicted native states to variations in the interaction schemes. Protein Sci 8:361–369
29. Zhang C, Liu S, Zhou H, Zhou Y (2004) An accurate, residue-level, pair potential of mean force for folding and binding based on the distance-scaled, ideal-gas reference state. Protein Sci 13:400–411
30. Breiman L (2001) Random forests. Mach Learn 45:5–32
31. R Development Core Team (2008) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna
32. Peters B, Sidney J, Bourne P, Bui HH, Buus S, Doh G, Fleri W, Kronenberg M, Kubo R, Lund O, Nemazee D, Ponomarenko JV, Sathiamurthy M, Schoenberger S, Stewart S, Surko P, Way S, Wilson S, Sette A (2005) The immune epitope database and analysis resource: from vision to blueprint. PLoS Biol 3:e91
33. Lata S, Bhasin M, Raghava GP (2009) MHCBN 4.0: a database of MHC/TAP binding peptides and T-cell epitopes. BMC Res Notes 2:61
34. Toseland CP, Clayton DJ, McSparron H, Hemsley SL, Blythe MJ, Paine K, Doytchinova IA, Guan P, Hattotuwagama CK, Flower DR (2005) AntiJen: a quantitative immunology database integrating functional, thermodynamic, kinetic, biophysical, and cellular data. Immunome Res 1:4
35. Lin HH, Zhang GL, Tongchusak S, Reinherz EL, Brusic V (2008) Evaluation of MHC-II peptide binding prediction servers: applications for vaccine research. BMC Bioinformatics 9(Suppl 12):S22
36. Zhang GL, Lin HH, Keskin DB, Reinherz EL, Brusic V (2011) Dana-Farber repository for

machine learning in immunology. J Immunol Methods 374:18–25

37. El-Manzalawy Y, Dobbs D, Honavar V (2008) On evaluating MHC-II binding peptide prediction methods. PLoS One 3:e3268
38. Bordner AJ, Mittelmann HD (2010) Prediction of the binding affinities of peptides to class II MHC using a regularized thermodynamic model. BMC Bioinformatics 11:41
39. Wang P, Sidney J, Kim Y, Sette A, Lund O, Nielsen M, Peters B (2010) Peptide binding predictions for HLA DR, DP and DQ molecules. BMC Bioinformatics 11:568
40. Nielsen M, Lund O (2009) NN-align. An artificial neural network-based alignment algorithm for MHC class II peptide binding prediction. BMC Bioinformatics 10:296
41. Rognan D, Lauemoller SL, Holm A, Buus S, Tschinke V (1999) Predicting binding affinities of protein ligands from three-dimensional models: application to peptide binding to class I major histocompatibility proteins. J Med Chem 42:4650–4658
42. Schueler-Furman O, Altuvia Y, Sette A, Margalit H (2000) Structure-based prediction of binding peptides to MHC class I molecules: application to a broad range of MHC alleles. Protein Sci 9:1838–1846
43. Liu Z, Dominy BN, Shakhnovich EI (2004) Structural mining: self-consistent design on flexible protein-peptide docking and transferable binding affinity potential. J Am Chem Soc 126:8515–8528
44. Antes I, Siu SW, Lengauer T (2006) DynaPred: a structure and sequence based method for the prediction of MHC class I binding peptide sequences and conformations. Bioinformatics 22:e16–e24
45. Knapp B, Omasits U, Frantal S, Schreiner W (2009) A critical cross-validation of high throughput structural binding prediction methods for pMHC. J Comput Aided Mol Des 23:301–307
46. Yanover C, Bradley P (2011) Large-scale characterization of peptide-MHC binding landscapes with structural simulations. Proc Natl Acad Sci U S A 108:6981–6986
47. Davies MN, Sansom CE, Beazley C, Moss DS (2003) A novel predictive technique for the MHC class II peptide-binding interaction. Mol Med 9:220–225
48. Schafroth HD, Floudas CA (2004) Predicting peptide binding to MHC pockets via molecular modeling, implicit solvation, and global optimization. Proteins 54:534–556
49. Tong JC, Zhang GL, Tan TW, August JT, Brusic V, Ranganathan S (2006) Prediction of HLA-DQ3.2beta ligands: evidence of multiple registers in class II binding peptides. Bioinformatics 22:1232–1238
50. Zhang H, Wang P, Papangelopoulos N, Xu Y, Sette A, Bourne PE, Lund O, Ponomarenko J, Nielsen M, Peters B (2010) Limitations of Ab initio predictions of peptide binding to MHC class II molecules. PLoS One 5:e9272
51. Patronov A, Dimitrov I, Flower DR, Doytchinova I (2011) Peptide binding prediction for the human class II MHC allele HLA-DP2: a molecular docking approach. BMC Struct Biol 11:32
52. Doytchinova I, Petkov P, Dimitrov I, Atanasova M, Flower DR (2011) HLA-DP2 binding prediction by molecular dynamics simulations. Protein Sci 20:1918–1928

Chapter 21

Prioritization of Therapeutic Targets of Inflammation Using Proteomics, Bioinformatics, and *In Silico* Cell–Cell Interactomics

Arsalan S. Haqqani and Danica B. Stanimirovic

Abstract

Leukocyte extravasation is a multistep process, involving the movement of leukocytes out of the circulatory system, through vascular endothelium and to the site of tissue damage or infection. Protein–protein interactions play key roles in the extravasation process and have been attractive therapeutic targets for inhibiting inflammation using blocking (or neutralizing) antibodies. These targets include protein–protein interactions between cytokines (or chemokines) and their receptors on leukocytes and between adhesions molecules involving leukocyte–endothelium contacts. A number of therapeutics against these targets are currently used in clinic for treatment of inflammatory disorders, however, they are associated with side-effects partly due to the off-target actions (i.e., nonspecific targets). There is a need for novel targets involved in the leukocyte extravasation process that are specific to leukocyte subsets or to individual inflammatory disorder, and are amenable for drug development (i.e., duggable). In this chapter, we describe a methodology to identify novel "druggable" targets involving protein–protein interactions between activated leukocytes and endothelial cells using a combination of proteomics, bioinformatics and *in silico* interactomics. The result is a prioritized list of protein–protein interactions in a network consisting of leukocyte–endothelial cell communication and contacts. These prioritized targets can be pursued for the development of therapeutics such as neutralizing antibodies and for their validation through preclinical testing. The method described here provides the workflow to identify and clinically target important cell–cell interactions that are specific/selective for particular inflammatory disorders and to improve currently available therapies.

Key words Protein–protein interactions, Intercellular, Target prioritization, Therapeutics, Inflammation, Extravasation, Druggable, Proteomics, Bioinformatics

1 Introduction

Proteomics and other "omics" technologies have been applied to profile genome-wide molecular changes in diseased states in the hope of identifying new targets for therapy and diagnosis. The main bottlenecks in translating "omics" discoveries into applications are overwhelming data sets that are difficult to rationalize and prioritize for validation and development. Validating entire "lists" of

Kelly M. Fulton and Susan M. Twine (eds.), *Immunoproteomics: Methods and Protocols*, Methods in Molecular Biology, vol. 1061,
DOI 10.1007/978-1-62703-589-7_21,

differentially expressed genes/proteins is costly and time-consuming. As a result, a single-factor/reductionist approach has often been applied to select (or "cherry pick") a couple of molecules for further evaluation, and a significant portion of disease-implicated molecules are simply overlooked. Prioritization of identified differentially expressed molecules in the entire data-set in an unbiased manner is required to identify novel and more specific targets.

Target prioritization involves narrowing down the candidate targets to those that can be exploited for the development of therapeutic ligand or in targeted diagnostic. Typically, the lead target is used to develop a series of targeting agents (TA) using combinatorial small molecule libraries or monoclonal antibody generation by hybridoma or antibody display technologies; these targeting agents are then evaluated and optimized to select a "lead drug". In most cases, the target prioritization is driven by the "druggability" of the targets. Some important criteria for target "duggability" include differential expression of the target in diseased condition(s), specificity for the disease and tissue of interest, and its structural accessibility for binding to the TA for a rapid (functional) response. The later criterion is dependent on the type of TA; whereas for small molecules "druggable" targets could be intracellular, for biologic therapeutics the accessibility of targets has been generally limited to molecules secreted into the extracellular milieu or expressed on the cell surface. A number of these "accessible" molecules have been therapeutically targeted using blocking (or neutralizing) antibodies in various disorders; examples in inflammation include Tysabri® (natalizumab), an anti-VLA4 blocking antibody [1] and Remicade® (infliximab), a TNFα-blocking antibody [2]. For both small molecule and biologic TAs, some details of the target structure and TA interactions with specific target epitopes are also desirable [3]. Due to space restriction, we have limited the method described in this chapter to prioritization of proteomics targets suitable for the development of therapeutic antibodies useful in inhibiting specific steps in the inflammation cascade.

Molecules involved in the process of leukocyte extravasation have been attractive therapeutic targets for inhibiting inflammation. Leukocyte extravasation is a multistep process, involving the movement of leukocytes out of the circulatory system, their activation, binding, and diapedesis through vascular endothelium and their migration towards the site of tissue damage or infection [4, 5]. Protein–protein interactions play key roles in each step of the process [6–8]: leukocyte activation/chemotaxis involves cytokine interactions with leukocyte-expressed-receptors (e.g., TNFα-receptor, IL23R), while leukocyte contact with and diapedesis through endothelial cells involves cell adhesion molecules (e.g., VLA4-VCAM1, LFA1-ICAM1). While *intra*cellular protein–protein interactions (e.g., in signal transduction) also play

important roles, the interactions involving "accessible" membrane and secreted molecules are considered more attractive targets for biologic therapeutics. Hence Tysabri® and Remicade® are clinically and successfully used to inhibit T-cells from entering the brain by blocking VCAM1-VLA4 interaction in multiple sclerosis, and to reduce leukocyte activation by preventing TNFα binding its receptor in rheumatoid arthritis, respectively. Despite their use in clinic, these therapeutics are not free of side-effects [6, 9, 10], which are in part due to widespread physiological functions of their target antigens. In addition, targeting these interactions alone does not provide complete protection, suggesting that other molecules may "compensate" when these interactions are blocked [11]. Thus, there is a need to identify new "druggable" protein–protein interactions involved in the leukocyte extravasation process that are specific to pathogenic cell type or to individual inflammatory disorder.

We recently proposed a methodology to identify novel set of "druggable" protein–protein interactions between Th17 and brain endothelial cells (ECs) using a combination of proteomics, bioinformatics and *in silico* interactomics [8] (Fig. 1). Here we provide protocols for this method. The method involves inflammatory activation of brain EC and T-cells followed by advanced label-free proteomics (nanoLC-MS/MS) of the cells. An initial list of accessible proteins (i.e., present on cell surface and secretome) is generated and prioritized based on bioinformatics analysis of their potential interactions using the database of protein–protein interactions, compiled from existing databases from the public domain. The result is a reconstructed brain EC-T-cell network—interactome—consisting of cell–cell communications and cell–cell contacts. The network is used to identify novel and specific protein–protein interactions that can be blocked to disrupt the extravasation process. While the method has been used to identify interactions between brain ECs and T-cells, it is also applicable to other disease-specific leukocyte–EC interactions as well as other cell–cell interactions and communications.

2 Materials

2.1 ECs and T-Cells

1. Brain ECs used were the hCMEC/D3 human brain endothelial cell line as a stable human *in vitro* model of the BBB [12]. hCMEC/D3 were activated under serum-free conditions using various inflammatory insults, including TNFα/INFγ, IL-1β or simulated ischemia/reperfusion conditions as previously described [11, 13, 14].
2. T cells with encephalo-tropism were generated by activation of lymphocytes, isolated from multiple sclerosis patients, with IL-23 as recently described [15].

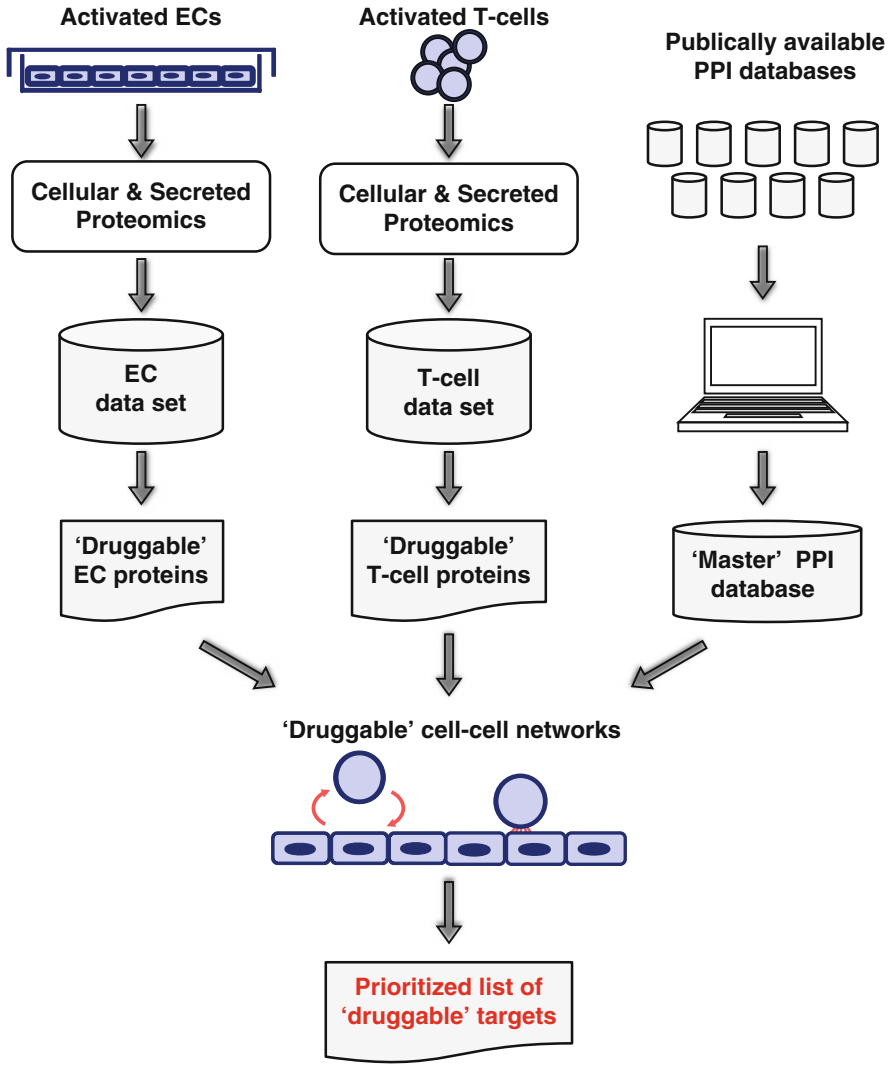


Fig. 1 Schematic representation of the experimental approach. Endothelial cells (ECs) and T-cells are activated using appropriate inflammatory conditions and their cellular and secreted proteins are analyzed by proteomics to generate a list of respective datasets. "Druggable" proteins in each database are identified using bioinformatics described in Subheading 3. In parallel, protein–protein interaction (PPI) databases are obtained from publically available sites and combined to create a "master" PPI database in house. These are used to create "druggable" cell–cell networks between EC and T-cells and produce a list of prioritized "druggable" targets

2.2 Proteomic Sample Preparation

1. At least 100 μg of protein is needed from ECs and T-cells for proteomics analysis.
2. Buffer-exchange/concentrating columns: such as Centriprep or Amicon Ultra columns (Millipore, QC, Canada) with a nominal molecular cut-off of 5,000 Da.
3. Prechill acetone by refrigerating at −20 °C for 20 min.

4. Benchtop centrifuge: Eppendorf Centrifuge 5415D (Brinkmann Instruments, Westbury, NY).
5. AMBIC solution: 50 mM ammonium bicarbonate.
6. 550 Sonic Dismembrator (Fisher Scientific, Canada).
7. Denaturing SDS buffer (DS buffer): 50 mM Tris–HCl, pH 8.5, 0.1 % SDS.
8. DTT stock: 100 mM of dithiothreitol freshly prepared in 50 mM AMBIC solution.
9. IAA stock: 250 mM of iodoacetamide freshly prepared in 50 mM AMBIC solution.
10. MilliQ water: Milli-Q® Ultrapure MS-grade water (Millipore).
11. Trypsin Gold, Mass Spectrometry Grade (Cat. # V5280, Promega, Madison, WI).
12. Pierce Strong Cation Exchange Spin Columns (Cat. # 90008, Thermo Scientific, Waltham, MA).
13. Phosphate-buffered saline (PBS): 137 mM NaCl, 2.7 mM KCl, 10 mM $Na_2HPO_4{\cdot}2H_2O$, 2.0 mM KH_2PO_4, pH 7.4
14. 0.22 µm syringe filters.
15. High pressure liquid chromatography (HPLC) grade acetonitrile.
16. HPLC grade formic acid.

2.3 Mass Spectrometry

1. An online nanoflow liquid chromatography (nanoLC) system with a reverse phase nanoLC column such as the highly reproducible nanoAcquity UPLC system (Waters, Millford, MA) with a 100 µm I.D. × 10 cm 1.7 µm BEH130C18 column (Waters).
2. A tandem MS instrument capable of performing electrospray ionization directly on the eluting peptides followed by high throughput MS/MS analysis such as the LTQ XL™, LTQ Orbitrap™ (Thermo) or the QTOF™ Ultima (Waters). The LTQ instruments are of choice due to its ability to do high-throughput sequencing. In our laboratory LTQ can obtain 7–10 times more MS/MS spectra than QTOF Ultima.

2.4 Software and Databases for Data Analysis

1. NanoLC-MS/MS data may be converted into a number of data formats that are compatible for analysis by software such as MSight, Mascot, or MatchRx. The software that comes with the MS instrument or a number of available tools from http://tools.proteomecenter.org can also be used. ReAdW from http://tools.proteomecenter.org was used here for LTQ data.
   (a) MSight software is a visualization tool [16] available free of charge from http://web.expasy.org/MSight. It allows graphical representation of the LC-MS and LC-MS/MS data. MSight version 1.0 was used here.

Table 1
Example of a list of publically available mammalian protein–protein interaction databases

Database	Web site
BioGRID	http://thebiogrid.org
Bond	http://bond.unleashedinformatics.com
GeneRif	ftp://ftp.ncbi.nlm.nih.gov/gene/GeneRIF
HCPIN	http://nesg.org:9090/HCPIN
HPID	http://wilab.inha.ac.kr/hpid
HPRD	http://www.hprd.org
HUGE	http://www.kazusa.or.jp/huge
I2D	http://ophid.utoronto.ca/ophidv2.201
KEGG	http://www.genome.jp/kegg
MINT	http://mint.bio.uniroma2.it
MPPI	http://mips.helmholtz-muenchen.de
Reactome	http://www.reactome.org

(b) Mascot® software (Matrix Science Ltd., London, UK) is a probability-based search engine for identifying peptide sequences from the nanoLC-MS/MS data using protein database searching [17]. Mascot® version 2.2.0 was used here.

(c) MatchRx software (from National Research Council (NRC)) extracts peptide abundance values from LC-MS and LC-MS/MS data and allows quantitative comparison of peptide levels in two or more samples [18]. MatchRx also overlays the quantitative differences of peptides and LC-MS/MS identification results (from Mascot) into MSight images. MatchRx version 4.0 was used in studies described here.

2. MeV or similar software for analysis, visualization and data-mining of large-scale omic data from http://www.tm4.org/mev.
3. SwissProt protein database in FASTA format: uniprot_sprot.fasta file from ftp://ftp.uniprot.org.
4. Uniprot Knowledgebase: uniprot_sprot.dat file from ftp://ftp.uniprot.org.
5. Protein–protein interaction software are listed in Table 1.
6. Cytoscape an open-source software for complex network analysis and visualization [19] from http://www.cytoscape.org.

3 Methods

3.1 Proteomic Analyses of Activated ECs and T-Cells

The first step requires generation of a list of proteins that are expressed in activated ECs and activated leukocytes. It is important that the right types of cells (e.g., brain EC, lung EC, Th17, B-cells) and cell-activation paradigms are chosen for each specific inflammatory condition. Advanced proteomics methods can be used to identify novel proteins that are expressed specifically in the type of ECs or leukocytes being examined (*see* **Note 1**). In this chapter, we provide a general protocol for isolation of proteins from cellular and secreted fractions followed by proteomics analysis (*see* **Note 2**). Four types of samples are recommended for use in subsequent interactomics analysis: (1) Cellular fractions of activated ECs; (2) Cellular fractions of activated T-cells; (3) Secreted fraction of activated ECs; (4) Secreted fraction of activated T-cells. For each sample type, use of an appropriate control is highly recommended (e.g., untreated control, unrelated ECs or T-cells) to obtain a more specific target.

3.1.1 Cell Activation and Harvest

1. Grow at least 1×10^6 cells on culture plates or flasks and activate the cells to mimic inflammatory conditions, using appropriate controls (*see* **Note 3**). Examples of EC activation include inflammatory cytokines (IL1beta or TNFalpa/IFNgamma) or simulated *in vitro* ischemia/reperfusion as previously described [11, 13, 14]. An example of activated leukocyte includes generation of encephalo-tropic T cells by activation of lymphocytes (from multiple sclerosis patients) with IL-23 to produce IL-17-producing T-helper (Th17) [15].
2. Remove the medium for isolation of secreted proteins and harvest the plated cells by scraping in PBS for isolation of cellular proteins (*see* **Note 4**).

3.1.2 Protein Isolation from Cultured Media

1. Filter the collected cultured medium through a 0.22 μm syringe filter to remove any floating cells.
2. Concentrate and desalt the medium to about less than 0.5 mL using buffer-exchange/concentrating columns using the manufacturer's protocol.
3. Precipitate the proteins by adding 10-volumes of ice-chilled acetone and incubating at −20 °C for at least 1 h.
4. Centrifuge at $1{,}000 \times g$ for 10 min to pellet the proteins.
5. Decant the supernatant.
6. Resuspend the pelleted secreted proteins in AMBIC solution and measure protein concentration.

3.1.3 Protein Isolation from Cells

1. Centrifuge the harvested cells at $500 \times g$ for 5 min and aspirate off the PBS supernatant.
2. Resuspend the pellet in 0.1 mL of DS buffer.

3. Lyse by sonication for 4–5 cycles of 15 s each at a low setting of 3, followed by 20 s on ice.
4. Centrifuge at 10,000 × *g* for 10 min at 4 °C to pellet the cell debris and nuclei.
5. Transfer supernatant to a fresh tube and measure protein concentration (for example by using a commercially available standard protein concentration kit).

3.1.4 Protein Digestion, Fractionation and NanoLC-MS/MS Analysis

1. To cellular and secreted proteins, add freshly prepared DTT stock to a final concentration of 4 mM and incubate at 95 °C for 10 min to reduce the disulfide bonds.
2. Cool at room temperature for 2 min.
3. Add freshly prepared IAA stock to a final concentration of 10 mM and incubate at room temperature for 20 min in the dark to alkylate cysteine residues.
4. Add trypsin (in Milli-Q® water) at a 1:50 (w/w) trypsin-to-protein ratio and incubate at 37 °C for 12–16 h.
5. Fractionate the tryptic peptides using strong cation exchange columns using the manufacturer's protocol.
6. Samples may be stored at –80 °C for up to 4 weeks or analyzed directly by LC-MS/MS.
7. Inject 1–5 % of each sample into a nanoLC system setup online to a tandem mass spectrometer. For example, a typical analysis would involve peptide separation on the nanoAcquity system by gradient elution (1–95 % ACN, 0.2 % formic acid) over 60 min at a flow rate of 400 nL/min and acquiring MS/MS spectra on 2+, 3+ and 4+ charged precursor ions with m/z values between 400 and 2,000 (*see* **Note 5**).
8. Search the acquired MS/MS spectra against a human SwissProt protein database using Mascot® or another search engine (*see* **Notes 6** and **7**).
9. Obtain normalized relative quantitation on peptide and protein levels in each sample using MatchRx or another label-free quantitative software (*see* **Note 8**).
10. Generate a quantitative protein list for each sample type: (1) Cellular fractions of activated ECs; (2) Cellular fractions of activated T-cells; (3) Secreted fraction of activated ECs; (4) Secreted fraction of activated T-cells. An example of the protein list is shown in Fig. 2.

3.2 List of "Druggable" Proteins in Activated ECs and T-Cells

The purpose of this step is to reprioritize the protein list by giving higher ranking to proteins that are more easily accessible to therapy (i.e., proteins on cell surface and extracellular space) and are induced under activated cell state (i.e., inflammation-responsive proteins) (*see* **Note 9**).

Acc	GN	Protein name	No of peptides	Maximum Mascot score	Normalized Intensities								Median Fold-Change: Log_2(T/C)	Rankings					Final Ranking
					Treated$_1$	Treated$_2$	Treated$_3$	...	Control$_1$	Control$_2$	Control$_3$	...		Inflammatory response	"Druggability"	Interactomics			
																Degree	Betweeness	...	
Acc$_1$	*GN$_1$*	*Protein$_1$*	*NP$_1$*	*Sc$_1$*	#	#	#	...	#	#	#	...	*MFC$_1$*	#	#	#	#	...	#
Acc$_2$	*GN$_2$*	*Protein$_2$*	*NP$_2$*	*Sc$_2$*	#	#	#	...	#	#	#	...	*MFC$_2$*	#	#	#	#	...	#
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
P19320	*VCAM1*	*Vascular cell adhesion protein 1*	*60*	*122*	#	#	#	...	#	#	#	...	*18.7*	*1*	*1*	2	5	...	2
P05362	*ICAM1*	*Intercellular adhesion molecule 1*	*50*	*105*	#	#	#	...	#	#	#	...	*7.1*	*2*	*1*	4	7	...	8
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
Acc$_N$	*GN$_N$*	*Protein$_N$*	*NP$_N$*	*Sc$_N$*	#	#	#	...	#	#	#	...	*MFC$_N$*	#	#	#	#	...	#

Fig. 2 An example of a quantitative protein list showing various types of rankings

3.2.1 Cell-Surface and Extracellular Proteins

1. Download an updated UniprotKB database or directly query www.uniprot.org (*see* **Note 6**).
2. Use the quantitative protein list for each sample type as an initial list (Fig. 2).
3. To identify if each protein is known to be present on cell surface and/or extracellular environment, query the protein accession in the database and look for the following attributes.
   (a) Under section "Comments" or "CC" and sub-section "Subcellular location", determine whether the protein is located in plasma membrane and/or is secreted.
   (b) Under section "Gene Ontology" or "DR GO" and sub-section "Cellular Component" or "C", determine whether the protein is located in plasma membrane and/or extracellular space.
   (c) Under section "Sequence Annotation (Features)" or "FT" and sub-section "Regions", determine whether the "Topological domain" is "Extracellular" and protein contains "Transmembrane" domain.
4. Add a new column for "ranking by druggablility" to each protein list and rank as follows (*see* **Note 10**):
   (a) Proteins with known extracellular domains found on the plasma membrane or secreted proteins are given a ranking of 1.
   (b) Other plasma membrane proteins are given a ranking of 2.
   (c) Protein with no known subcellular localizations are labeled as "Unknowns" and given a ranking of 5.
   (d) The remaining proteins (including nuclear and mitochondrial proteins) are either given a ranking of 1,000 or removed from the list.

Proteins	Interactants	Pubmed Reference [optional]	Evidence [optional]	known/predicted [optional]	Database [optional]					
					HPRD	APID	GeneRif	I2D	BioGrid	GeneMania
VCAM1										
	VLA4 (ITGA4/ITGB1)	12097820,19057841		known	✓	✓	✓	✓	✓	✓
	ITGAD	10438935		known	✓	✓	✓	✓	✓	✓
	ITGB7	12682249		known	✓	✓	✓	✓		
	CCL17	12193695	Reconstituted Complex	known	✓	✓	✓	✓	✓	✓
	CCL22			known	✓	✓	✓	✓	✓	✓
	EZR	12082081	Affinity Capture-Western; Reconstituted Complex	known	✓	✓	✓	✓	✓	✓
	MSN			known	✓	✓	✓	✓	✓	✓
	CTSG			known	✓	✓	✓	✓		
	ELANE			known	✓	✓	✓	✓		
	IL13			known	✓	✓	✓	✓		
	IL1B1			known	✓			✓		
	SEC61B			known			✓			
ICAM1										
	LFA1 (ITGAL/ITGB2)	10493852,11786177		known	✓					
	Mac1 (ITGAM/ITGB2)	9488691,11073102		known						
	EGFR	16799092								

Fig. 3 A simple format for protein–protein interaction database

3.2.2 Inflammation-Responsive Proteins

1. For each sample type, calculate a median fold-change value for each protein in the list (*see* **Note 11**).
2. Carry out statistical analysis (e.g., *t*- or *U*-tests) to identify whether the "control" and "treated" intensities are significantly different ($p < 0.05$) (*see* **Note 12**).
3. Proteins showing fold change ≥2 and *p*-value ≤0.05 are considered inflammation-responsive.
4. Add a new column for "ranking by inflammatory changes" to each protein list and rank as follows (*see* **Note 10**):
   (a) Proteins showing tenfold change are given a ranking of 1.
   (b) Proteins showing fourfold change are given a ranking of 2.
   (c) Proteins showing twofold change are given a ranking of 3.
   (d) Proteins showing no change are given a ranking of 1,000.

3.3 Prioritization of "Druggable" Targets Using Cell–Cell Interactomics

In this step, the list of "druggable" inflammatory proteins is further prioritized based on their potential involvement in interaction between ECs and T-cells. This requires the use of a database of known protein–protein interactions, which is first compiled from existing databases in the public domain. The database is used to reconstruct networks between ECs and T-cells consisting of cell–cell communications and cell–cell contacts (*see* **Note 9**).

3.3.1 Protein–Protein Interaction Databases

1. A number of publically available protein–protein interaction databases currently exist. Download each protein–protein interaction database listed in Table 1 to a local location (*see* **Note 13**).
2. Since each database has a different format, convert each database into a single format as shown in Fig. 3 (*see* **Note 14**).
3. Merge all the databases into one "master" database (Fig. 3).

4. It may be necessary to limit the interactions to immunoprecipitation and affinity pull-down assays and to mammalian systems to reduce the incidence of false interactions.
5. The resulting protein–protein interaction database, referred as "master" PPI database, is used in the next section.

3.3.2 In Silico Reconstruction of EC-T Cell Interactions

1. Use the list of "druggable" cell-surface proteins from activated ECs as "List 1", and the list of "druggable" cell-surface proteins from activated T-cells as "List 2".
2. Search each protein from List 1 in the "Protein" column of the "master" PPI database.
3. If the protein is found in the "Protein" column and has protein(s) in the interactant column, search each of the interactants in List 2 proteins.
4. For each interactant found, put each protein–protein interaction pair in a tab-delimited SIF file format for Cytoscape (e.g., VCAM1<tab>pp<tab>VLA4<newline>) (*see* **Note 15**).
5. Visualize all the protein–protein interaction pairs using Cytoscape.
6. Graph theories could be applied using Cytoscape tools or other methods [19, 20] to rank the protein list based on various types of measures of centrality (e.g., betweenness, degree) of the interactions in the network. Thus, for each measure, a new column is added to the list and proteins are ranked by descending measure. Proteins not involved in protein–protein interactions are either removed or given a ranking of 1,000 (*see* **Note 10**).
7. Optionally all the different columns of rankings can be combined to obtain a final ranked list (e.g., using reciprocal or other types of ranking methods [21–23]).
8. The above steps will identify interacting proteins between ECs and T-cells involved in cell–cell contacts. To identify cytokine-receptor interactions involved in cell–cell communications, the steps can be repeated using secreted proteins as one of the lists. That is, use the list of "druggable" cell-surface proteins from activated ECs as "List 1", and the list of secreted proteins from activated T-cells as "List 2", and vice versa, i.e., cell-surface T-cell proteins as "List 1", and secreted EC proteins as "List 2".

3.4 Validation Studies

A range of protocols, from nanoLC-MS/MS to cell–cell interactomics, have been described here and it is recommended to include various types of validation studies to have confidence in the results.

1. Quantitative results from nanoLC-MS/MS analysis are generated in an automated manor using label-free quantitative software. While validating all results is tedious, it is recommend that proteins showing statistically significant expression changes in

their intensities (inflammation responsive proteins) should be either manually validated in the raw nanoLC-MS/MS data or "visually" validated using the MSight software that allows reconstruction of the raw data into a 2D gel-like image.

2. Since nanoLC-MS/MS proteomics can produce a large list of inflammation-responsive proteins, validations of a few of the results using alternative methods is highly recommended. For confirmation of changes at protein level, ELISA, western blots, MRM or *in situ* immunochemistry may be used. For confirmation of changes at mRNA levels, quantitative RT-PCR, northern blots or *in situ* hybridization may be used. Additionally, confirmation through literature mining also adds key value to the results by providing further validity to the method. In brain ECs we have validated a large number of proteins responding significantly to inflammatory insults by western blotting and immunochemistry, including ICAM1, VCAM1, and integrins [8].
3. It may be necessary to experimentally validate the found protein–protein interaction pairs between EC and T-cells using molecular techniques in co-culture experiments. This requires establishing co-cultures by culturing either EC with T-cells or EC with T-cell conditioned media, whereas EC without T-cells or the conditioned media can be used as control, respectively. The co-cultures are then immunoprecipitated with first protein and the second protein is examined by using western blotting or MRM. Alternatively, immunocytochemistry for both proteins can be carried out to demonstrate co-localization [11].
4. Finally, to validate that the found protein–protein interaction pairs are functionally important for the extravasation process, interaction-blocking therapeutic could be used. This could include a commercial or a newly developed antibody, a peptide or small molecule that could be tested for its ability to inhibit leukocyte chemotaxis, adhesion and transmigration through ECs in an *in vitro* assay [11].

3.5 Conclusions and Further Steps

Proteomics and other genomics methods generate an overwhelming number of molecules and prioritizing them to identify those implicated in disease is a challenge. While bioinformatics and statistical analyses can reduce the number of proteins involved, additional methods to identify novel and more specific targets are needed. We have described here a methodology to further narrow down proteomics list of activated EC and activated T-cells to identify potential therapeutic targets of inflammation using *in silico* interactomics. These methods have been applied recently to reconstruct intercellular interactions and communications between activated human brain EC and activated human Th17 cells [8], identifying a large number of known and novel "druggable" protein–protein interaction pairs. Some of these include: VCAM1-VLA4,

ICAM1-LFA1, ICAM1-Mac1, PECAM1-PECAM1, and E-selectin-PSGL1. These interactions could be prevented or suppressed using specific antibodies against one or more interacting proteins. These antibodies could be screened in EC-T-cell co-culture experiments to demonstrate inhibition of interactions as described [11]. This method identified several known interactions, including VCAM1-VLA4, already targeted by clinically used therapeutic antibodies (e.g., Natalizumab—Tysabri), providing a proof-of-concept for potential clinical translation of the proposed approach. The method provides the opportunity to identify and target other important leukocyte–EC interactions that are more specific/selective for inflammatory cell subtypes and for various types of inflammatory diseases.

4 Notes

1. Alternatively, a protein list from published proteomics analyses of activated ECs and/or activated leukocytes can also be used. In such case, the proteomics step can be skipped. However, the type of cells, activation protocols and controls used in the publication should be noted to ensure it is appropriate for the inflammatory condition in question.
2. Subcellular proteomics can also be used to identify more specific and relevant targets such as luminal membrane proteins (blood-accessible targets), glycoproteins or lipid raft proteins [11, 24]. This would require subcellular fractionation and glycoprotein enrichment followed by the proteomics method described here.
3. Once cells have been plated and reached confluence, it is important that the cells be incubated for an appropriate period of time to allow enough proteins to accumulate in the medium for detection by proteomics. It may be necessary to measure protein content at the end of the incubation to ensure that there is enough protein (>150 μg). Usually one 10-mm plate of confluent cells gives about 100–500 μg of protein, and multiple plates may be pooled if necessary. Cells should be activated in a media that is free of any serum or other supplements. To ensure this, cells should be washed at least three times with a buffered saline solution and then incubated in serum-free medium in the presence of activating inflammatory cytokines.
4. Cell harvesting by scraping is recommended rather than by trypsin treatment to avoid digestion and “loss” of cell-surface molecules.
5. Injecting at least one blank run is recommended in-between samples to clean out the column and prevent cross-contamination between samples. If the file size is large, it may be necessary

to split the files into two *m/z* ranges 400–1,200 and 1,200–2,000.

6. An updated SwissProt database in FASTA format (uniprot_sprot.fasta) and its corresponding Uniprot knowledgebase (uniprot_sprot.dat) should be downloaded monthly from Uniprot FTP site (ftp://ftp.uniprot.org). The MS/MS data should be searched against the most up-to-date FASTA file.
7. For Mascot analysis, specify trypsin enzymatic cleavage with one possible missed cleavage; allow variable modification of oxidation (+15.99 Da) at the Met residues, allow fixed modification of carbamidomethyl (iodoacetamide derivative; +57.02 Da) at the Cys residues, set parent ion tolerance ≤0.5 Da and fragment ion tolerance ≤0.2 Da.
8. Automated steps in MatchRx software allow quantitative comparison of peptide and protein levels in multiple samples [18]. This involve peak identification and quantitation (AUC) in each LC-MS run, alignment of multiple LC-MS runs, peak normalization and identification of differentially expressed peptides and overlaying of the results on MSight images [17] for visual verification.
9. While these analyses can be done manually for a short protein list (<10), use of a computational method (e.g., macros, Perl scripts) is highly recommended for large datasets.
10. The rankings given here are only suggestive. Other scores should be tested to identify the most optimal ranking system for each paradigm.
11. For each Control-Treated pair, a fold-change ratio can be calculated as base-2 log of IntensityTreated/IntensityControl, and can be used to calculate a median fold-change ratio. Alternatively, median intensities can be calculated for each treated and control samples and then a fold-change ratio value can be calculated using these values.
12. This can be done using MeV (http://www.tm4.org/mev) or similar software. If comparing two states, a *t*-test can be used for parametric data or a Mann-Whitney *U*-test for nonparametric data. If more than two states are being compared, ANOVA followed by Bonferroni post hoc test should be used.
13. It is recommended that like protein databases, the interaction databases should be routinely updated since new interactions are continuously identified and validated as well as some old ones are occasionally refuted and deleted from the databases.
14. The format can be simple or complex, however it should include two main columns: “Proteins” and “Interactant”.
15. While SIF is a simpler format and is useful for a quick “look” at the interactions using Cytoscape, use of XML formats is

recommended since additional information such as subcellular localization, expression, rankings can be included. The later data type also gives more flexibility in Cytoscape to "design" a cell–cell interaction [19].

References

1. Brown BA (2009) Natalizumab in the treatment of multiple sclerosis. Ther Clin Risk Manag 5:585–594
2. Benucci M, Saviola G, Manfredi M, Sarzi-Puttini P, Atzeni F (2012) Tumor necrosis factors blocking agents: analogies and differences. Acta Biomed 83:72–80
3. Perot S, Sperandio O, Miteva MA, Camproux AC, Villoutreix BO (2010) Druggable pockets and binding site centric chemical space: a paradigm shift in drug discovery. Drug Discov Today 15:656–667
4. Rossi B, Constantin G (2008) Anti-selectin therapy for the treatment of inflammatory diseases. Inflamm Allergy Drug Targets 7:85–93
5. Engelhardt B, Wolburg H (2004) Mini-review: transendothelial migration of leukocytes: through the front door or around the side of the house? Eur J Immunol 34:2955–2963
6. Coisne C, Lyck R, Engelhardt B (2007) Therapeutic targeting of leukocyte trafficking across the blood–brain barrier. Inflamm Allergy Drug Targets 6:210–222
7. Wittchen ES (2009) Endothelial signaling in paracellular and transcellular leukocyte transmigration. Front Biosci 14:2522–2545
8. Haqqani AS, Stanimirovic DB (2011) Intercellular interactomics of human brain endothelial cells and th17 lymphocytes: a novel strategy for identifying therapeutic targets of CNS inflammation. Cardiovasc Psychiatry Neurol 2011:175364
9. Ransohoff RM (2005) Natalizumab and PML. Nat Neurosci 8:1275
10. Scheinfeld N (2004) A comprehensive review and evaluation of the side effects of the tumor necrosis factor alpha blockers etanercept, infliximab and adalimumab. J Dermatolog Treat 15:280–294
11. Cayrol R, Wosik K, Berard JL, Dodelet-Devillers A, Ifergan I, Kebir H, Haqqani AS, Kreymborg K, Krug S, Moumdjian R, Bouthillier A, Becher B, Arbour N, David S, Stanimirovic D, Prat A (2008) Activated leukocyte cell adhesion molecule promotes leukocyte trafficking into the central nervous system. Nat Immunol 9:137–145
12. Weksler BB, Subileau EA, Perriere N, Charneau P, Holloway K, Leveque M, Tricoire-Leignel H, Nicotra A, Bourdoulous S, Turowski P, Male DK, Roux F, Greenwood J, Romero IA, Couraud PO (2005) Blood–brain barrier-specific properties of a human adult brain endothelial cell line. FASEB J 19:1872–1874
13. Stanimirovic D, Shapiro A, Wong J, Hutchison J, Durkin J (1997) The induction of ICAM-1 in human cerebromicrovascular endothelial cells (HCEC) by ischemia-like conditions promotes enhanced neutrophil/HCEC adhesion. J Neuroimmunol 76:193–205
14. Haqqani AS, Kelly J, Baumann E, Haseloff RF, Blasig IE, Stanimirovic DB (2007) Protein markers of ischemic insult in brain endothelial cells identified using 2D gel electrophoresis and ICAT-based quantitative proteomics. J Proteome Res 6:226–239
15. Kebir H, Kreymborg K, Ifergan I, Dodelet-Devillers A, Cayrol R, Bernard M, Giuliani F, Arbour N, Becher B, Prat A (2007) Human TH17 lymphocytes promote blood-brain barrier disruption and central nervous system inflammation. Nat Med 13:1173–1175
16. Palagi PM, Walther D, Quadroni M, Catherinet S, Burgess J, Zimmermann-Ivol CG, Sanchez JC, Binz PA, Hochstrasser DF, Appel RD (2005) MSight: an image analysis software for liquid chromatography–mass spectrometry. Proteomics 5:2381–2384
17. Hirosawa M, Hoshida M, Ishikawa M, Toya T (1993) MASCOT: multiple alignment system for protein sequences based on three-way dynamic programming. Comput Appl Biosci 9:161–167
18. Haqqani AS, Kelly JF, Stanimirovic DB (2008) Quantitative protein profiling by mass spectrometry using label-free proteomics. Methods Mol Biol 439:241–256
19. Saito R, Smoot ME, Ono K, Ruscheinski J, Wang PL, Lotia S, Pico AR, Bader GD, Ideker T (2012) A travel guide to Cytoscape plugins. Nat Methods 9:1069–1076
20. Barabasi AL, Oltvai ZN (2004) Network biology: understanding the cell's functional organization. Nat Rev Genet 5:101–113
21. Assfalg A, Erdfelder E (2012) CAML—maximum likelihood consensus analysis. Behav Res Methods 44:189–201

22. bi-Haidar A, Kaur J, Maguitman A, Radivojac P, Rechtsteiner A, Verspoor K, Wang Z, Rocha LM (2008) Uncovering protein interaction in abstracts and text using a novel linear model and word proximity networks. Genome Biol 9(Suppl 2):S11
23. Melnik O, Vardi Y, Zhang CH (2004) Mixed group ranks: preference and confidence in classifier combination. IEEE Trans Pattern Anal Mach Intell 26:973–981
24. Haqqani AS, Hill JJ, Mullen J, Stanimirovic D (2011) Methods to study glycoproteins at the blood–brain barrier using mass spectrometry. In: Nag S (ed) Methods in molecular biology: the blood–brain and other neural barriers. The Humana Press, Totowa, NJ, pp 337–353

Chapter 22

Commercial Considerations for Immunoproteomics

Scott M. Ferguson

Abstract

The underlying drivers of scientific processes have been rapidly evolving, but the ever-present need for research funding is typically foremost amongst these. Successful laboratories are embracing this reality by making certain that their projects have commercial value right from the beginning of the project conception. Which factors to be considered for commercial success need to be well thought out and incorporated into a project plan with similar levels of detail as would be the technical elements. Specific examples of commercial outcomes in the field of Immunoproteomics are exemplified in this discussion.

Key words Technology transfer, Commercialization, Intellectual property management, Return on investment, Innovation exploitation

1 Introduction

The world of science has changed considerably over the last few decades with an increased focus on the needs of creating innovative new products or services and accelerating technology development for greater economic competitiveness. This trend is being driven by a number of influences. One of these comes from industry, where in order to continue to succeed in a globally competitive landscape, it is looking for innovation wherever it can be found and not just within their own walls. Additionally, for many countries national economic growth is increasingly driven by innovation and its deployment, so greater emphasis is placed on maximizing the development and deployment of commercially viable innovations.

Regardless of whether you are based in an academic setting, research institution, hospital, industrial lab, or other research center, there are increasing pressures for commercial value being realized from your research. Exemplar of this is the current regime for many grant applications. In the past, R&D grant or sponsorship applications may have been primarily curiosity driven projects with an understood requirement for publication of the findings of the research as part the final outcome. In today's realities, many of

Kelly M. Fulton and Susan M. Twine (eds.), *Immunoproteomics: Methods and Protocols*, Methods in Molecular Biology, vol. 1061,
DOI 10.1007/978-1-62703-589-7_22,

these granting agencies have added requirements to provide explicit plans for intellectual property (IP) protection and management, often with a further requirement to provide an outline for the transfer of the arising IP for commercial development.

Organizations are increasingly challenged to locate new sources of funding for their R&D activities, and the attractive concept of seeing a return on the investment (ROI) for their innovations through commercial development has led to the common practice of technology transfer from innovators to developers. The financial opportunities presented by industry and government economic growth initiatives further facilitate this reality and R&D organizations will quickly act on these opportunities for the benefit of their own financial needs.

For the bench scientist, this means that projects must be conceived with commercial objectives from the onset. If the host organization, or external funding source(s) cannot see the potential for an ROI, the chances of having a project funded become less likely in the current realities of science and innovation. The additional challenge to researchers of achieving commercial viability and success requires significant additional preparation and evaluation at the onset of any project to ascertain if a potential niche exists for products and/or services from this effort. These added requirements of course are over and above the already highly competitive environment and technological challenges of the science field.

2 Elements Required for Commercial Success

As noted, in the field of science today success is tied not only to scientific or technological prowess, but also to the ability to translate novel innovations into commercial products and services. In order to maximize the potential for commercial success, several elements should be considered when planning a project:

- Project plan initiation
- Opportunity evaluation
- Communicating
- Technology transfer and the entrepreneurial team
- Specific opportunities for immunoproteomics

2.1 Innovating: Where to Begin

One of the key challenges in launching an R&D project today is first identifying the unique opportunity to be addressed, and then clearly communicating your goals, objectives and outcomes to the key stakeholders of the proposed work. To begin, you must have the end in mind [1]. The project should address the application of the science and not just the specific scientific problem being tackled. Understanding this larger context of any project is

critical for the progressive challenges that will be encountered throughout the process of moving the technology from the bench to the market.

Industry and other stakeholders will most often ask the key question—how can this work be translated into economic benefit? Underlying this fundamental question are numerous other sub questions such as—what is the end use of the technology, what is the need being addressed, who else is addressing this need, and how will we have a competitive advantage to ensure a return on this investment? For the bench scientist, these are clearly very different questions than those encountered with respect to the scientific design and execution of a project and these will add a new level of complexity for gaining approval for a project. This is especially true when you consider that most scientists do not encounter much, or any, training towards these business objectives in the traditional science programs.

2.2 Evaluating the Commercial Opportunity

Much has already been written on the subject of strategy, technology transfer or the commercialization of R&D [2–7], so the focus here will be limited to covering the highlights of some of the key activities that should be addressed when performing the commercial evaluation of a project opportunity. Additionally, many organizations will have business or commercial development offices, or technology transfer professionals to help researchers with this evaluation and analysis.

As noted in the previous section, the first step is to understand the global application of the proposed science or technology. With this understanding, the project opportunity can be evaluated in terms of its market opportunity, e.g., what is the current market (products and sales/size), what is the competitive landscape (products on the market and in development), who are the competitors (size, strengths, weaknesses), what is the intellectual property landscape (open or segmented), who are the likely receptors of this technology (licensees, partners, buyers), what are the distribution networks (who, how will product be sold) and most importantly, what is the advantage this technology brings to this market space? These are some of the key questions to address when looking at the market opportunity. Additional details around marketing can be found in the references and literature in general.

Ideally this analysis will lead to the answer that there is potentially a good market opportunity for the project. Most often the opportunity will be either a new market entry with a clear unmet need being addressed, or an incremental improvement over existing products that also addresses a need (market pull) in the target markets. In both of these situations, it should be reasonable to predict an estimated ROI for the project.

The ROI analysis should focus on the addressable market opportunity, which is the market capture you would expect to gain

over time (price/profit, percentage of overall market sales, which territories, etc.). If potential sales profits exceed total project development and production costs within a desired time frame, then there is a case to be made for the project going forward. The greater the potential risk weighted ROI, the happier the funders/stakeholders will tend to be in seeing this project move forward.

2.3 Communicating the Project Opportunity

Many a time has a project failed because the overall advantages and/or benefits of a project were not clearly identified and communicated to stakeholders. You may have the best technical solution for a given scientific problem, but other projects move forward due to their ability to sell the project.

To be successful, you must first understand who you are communicating with [2]. For example, the research audience wants details on the scientific approach, the business manager wants to understand the market opportunity and how and when a profit will be realized, and the corporate development folks want a high level view of the overall project and how it may fit within their existing organization. Knowing your audience will allow you to target your communications with them to address their needs, and will increase your likelihood of your project going forward. When you can help solve someone else's problem, your opportunity value definitely increases.

This is true both within your organization and with potential external partners and stakeholders. Take the time to learn their needs/challenges first, and then proceed to explain how your project can address these opportunities. The successful projects are those that can clearly and succinctly communicate how they will solve these problems, and ultimately, make a financial return for the stakeholder organization. Marketed products are quite often not the best technical solutions to a given need, but were successful due to the strength of the marketing team behind them. Never underestimate the communications element of a project.

2.4 Technology Transfer Is a Team Activity

Science has often been seen as a solitary activity with high levels of competition between individual laboratories. Commercialization is more often a team activity requiring the skills of numerous parties from the bench to the market for success to be realized. For any project to successfully reach the market from the original concept, a collaborative effort should be considered. Without the broader range of skills required for commercial development and deployment, many projects result in excellent publications and numerous patents, but the commercial development objectives are not fully realized. Without the commercial success, the full value of the organizations technology investment may not be recouped.

Experience has often shown that from project initiation, successful innovation projects engage the scientific and business/technology transfer professionals to help to identify the overall

project opportunity, namely the full range of possible scientific and commercial outcomes. The project opportunity is then presented to the key stakeholders—management and/or collaborators to gain the necessary resources to deliver on the initial objectives. The scientific, business and management teams will likely engage granting agencies, industrial partners or other investors to further fund and develop the technology opportunity. With each milestone or stage of development, the team skills tend to expand and the scope of activity within the project broadens.

Effective collaborations will accelerate technology development towards market readiness. By employing the strengths of team members/collaborators, the learning curve is eliminated or reduced, and effective project plans can be quickly executed through the use of the additional expertise and resources of these team members. Understanding what skills will be necessary, how to engage them, and how to manage the larger project team, all become necessary skills for the project leaders of today.

3 Applications and Translation on Immunoproteomic Technologies

As we have seen in the earlier chapters, immunoproteomic technologies are rapidly advancing and the ability to identify truly novel biomarkers through these techniques is unprecedented. The researcher in this field must decide on a project/business model that fits the organizational goals and structure and provides for some of the commercial opportunities described previously. A review of some earlier publications discussing the commercialization of proteomics technologies [8–15] indicates that one key question still prevails today as was the case a decade ago. What is the commercial opportunity for immunoproteomic technologies?

A more recent review of the proteomics field further substantiates the concept that biomarker research has made substantial progress recently [16]. With the improvements in technology and techniques comes the reality of rapid and widespread identification of numerous novel biomarkers. It should be noted here that proteomics and other biomarker identification technologies have resulted in the discovery of thousands of such novel biomarkers. This great success ultimately means that biomarkers, in and of themselves, fall into the "dime a dozen" category from the industry or developers perspective. Industry has become inundated with opportunities to develop products around biomarkers and cannot keep up with the supply of targets. Validation of these biomarkers is therefore a necessary step to have enhanced value [17], and increase the likelihood of the project moving down the commercial development path.

For example, validated biomarkers can be further developed and evaluated for use in diagnostic applications, or potentially for

the development of therapeutic interventions. As a research tool, there are numerous additional applications, such as in the fundamental understanding of cell function, perturbation and pathogenesis. As a screening tool, drug modes of action can be elucidated and toxicology better understood. As the tools and techniques continue to improve, the possible outcomes will also evolve for the use of these biomarkers.

If we apply the principles cited earlier in this chapter in terms of the elements required for commercial success, the following questions should be addressed for your specific opportunity. First and foremost, what is the ultimate outcome for your research (in context of your organizational objectives and situation):

- Are you a service provider for parties external to your organization, providing biomarkers for their diagnostic/therapeutic needs?
- Are you part of a larger corporate team to develop diagnostic tools and/or therapeutic drugs for a specific indication?
- Do you expect to generate intellectual property uniquely, or in partnership with others?

How you answer these fundamental questions will dictate how your laboratory will operate and partner with others. Options may include the following.

3.1 Services Provider

If you have decided that you are a service provider to others, you need to address these additional fundamental questions of how you will develop your business model and generate income for your organization—Do you sell/license: Equipment time and expertise; Biomarkers for select indications; Validated biomarkers; Screening services; etc.

3.2 Integrated Team Member

If your analysis leads you to the conclusion that you will be part of an internal product development team or partnership consortia focussed towards the development of diagnostics and/or therapeutics technologies—your revenue questions are most often around licensing royalties versus outright sales revenues. With licensing revenues, income is typically spread over a longer time frame. Assignment or sales of your product/technology results in a single or several staged payments realized much sooner than would be the case for royalties.

4 Conclusions

Whether your research focus is in immunoproteomics, or for that matter in any other field, consideration should be made as to the potential outcomes of the project as early in the process as possible.

A basic understanding of the current business and economic drivers of science will be beneficial to all researchers and their managers by incorporating this understanding into project proposal plans. This may vary depending on the type of organization and the directives it embraces, but almost certainly, the enhanced value from partnered/sponsored work and/or a clear path for commercial exploitation will increase the likelihood of the project approval.

References

1. Covey SR (1989) The 7 habits of highly effective people. Simon & Schuster, New York
2. Kassicieh SK, Radosevich HR (1994) From lab to market: commercialization of public sector technology. Plenum, New York
3. Sarna L (2000) Guide to commercializing R&D. Jewel Publications Ltd., Montreal, QC
4. Jolly VK (1997) Commercializing new technologies: getting from mind to market. Harvard Business School Press, Boston, MA
5. Kaplan RS, Norton DP (2004) Strategy maps: converting intangible assets into tangible outcomes. Harvard Business School Publishing Corporation, Boston, MA
6. Tidd J, Bessant J, Pavitt K (2001) Managing innovation: integrating technological, market and organizational change. Wiley, Chichester
7. Lane RP, Daly PB (2006) Managing biotechnology risk from concept to commercialization: a strategic approach. BioExecutive International April 2006
8. Branca MA (2004) The path to market for genomics-based diagnostics. Spectrum diagnostics and related technologies. Decision Resources, Waltham, MA
9. Krul KG (1999) Pharmagogenomics: key technologies, major players, and lingering questions. Spectrum pharmaceutical industry dynamics. Decision Resources, Waltham, MA
10. Miller T (2000) Challenges in genomics-based biotechnology: Part I: building a proprietary technology portfolio. Spectrum pharmaceutical industry dynamics. Decision Resources, Waltham, MA
11. Miller T (2000) Challenges in genomics-based biotechnology: Part II: key genomics leaders. Spectrum pharmaceutical industry dynamics. Decision Resources, Waltham, MA
12. Haberman A (2001) Protein biochips: an emerging technology for proteomics. Spectrum drug discovery and design. Decision Resources, Waltham, MA
13. Lockwood D, Branca M (2001) Genomics dealmaking in 2000 and future outlook, Part II: proteomics and functional genomics. Spectrum drug discovery and design. Decision Resources, Waltham, MA
14. Krul KG (2001) Advances in proteomics: its use and commercialization. Spectrum pharmaceutical industry dynamics. Decision Resources, Waltham, MA
15. Matheson A (1998) Profiting from proteomics: strategies and technologies. Scrip Reports, BS 982, UK
16. Crowther G (2012) Biomarker discovery methods evolving. Genet Eng Biotechnol News 32(9):44
17. Mou Z, He Y, Wu Y (2009) Immunoproteomics to identify tumor-associated antigens eliciting humoral response. Cancer Lett 278:123–129

INDEX

Kelly M. Fulton and Susan M. Twine (eds.), *Immunoproteomics: Methods and Protocols*, Methods in Molecular Biology, vol. 1061, DOI 10.1007/978-1-62703-589-7, © Springer Science+Business Media, LLC 2013

L

M

P

R

S

T

W

Y