PART 1

IMMUNOGEN DESIGN
1.1 Introduction

Infectious diseases are the greatest cause of morbidity and mortality worldwide; pathogenic bacteria are responsible for approximately 50% of this burden. From a public health standpoint, prevention of diseases has a greater impact and is more cost effective than treating the infection. Vaccines are the most cost-effective methods to control infectious diseases and at the same time one of the most complex products of the pharmaceutical industry. There are several infectious diseases for which traditional approaches for vaccine discovery have failed. With the advent of whole-genome sequencing and advances in bioinformatics, the vaccinology
field has radically changed, providing the opportunity for developing novel and improved vaccines. Overall, the combination of different approaches ("-omics" approaches)—genomics, transcriptomics, metabolomics, structural genomics, proteomics, and immunomics—are being exploited to design new vaccines.

1.2 HISTORICAL VIEW OF "CLASSICAL" VACCINOLOGY

The history of vaccination is traditionally dated to the publication, in 1798, of Edward Jenner’s landmark experiments with cowpox in which he inoculated a neighbor’s boy with purulent material from a milkmaid’s hand lesion in the United Kingdom. The boy, 8 years old, was subsequently shown to be protected against a smallpox challenge. For more than 80 years, little more was done with respect to immunization, until Louis Pasteur discovered the attenuating effect of exposing pathogens to air or to chemicals. This discovery was achieved as the result of leaving cultures on the laboratory bench during a summer holiday. Thus, Pasteur developed the first vaccine made in the laboratory and also founded the terminology of vaccination [1, 2]. Since the time of Pasteur until recently, there have been two paths of vaccine development: attenuation or inactivation and the production of recombinant subunits. With regard to attenuation, heat, oxygenation, chemical agents, or aging were the first methods used, notably by Pasteur for rabies and anthrax vaccines. Passage in an animal host, such as the embryonated hen’s egg, was the next method, as practiced by Theiler for the yellow fever vaccine. After the development of in vitro cell culture in the 1940s, attenuation was accomplished by a variety of means, including selection of random mutants, adaptation to growth at low temperatures, chemical mutation to induce inability to grow at high temperature (temperature sensitivity), or induction of auxotrophy in bacteria. The second set of strategies are represented by the inactivation of the microorganism or by purifying small subunits derived from the pathogen of interest.

Late in the nineteenth century, Theobald Smith in the United States and Pasteur’s colleagues independently showed that whole organisms could be killed without losing immunogenicity. This new strategy soon became the basis of vaccines for typhoid and cholera and later for pertussis, influenza, and hepatitis A. Other approaches consisted in isolation of virulence factors from the microorganisms, such as toxins or capsular polysaccharides. In the 1920s, the exotoxins of Corynebacterium diphtheriae and Clostridium tetani were inactivated by formalin, to provide antigens for immunization against diphtheria and tetanus [1]. Extracted type-b polysaccharide capsule of Haemophilus influenzae was shown attractive as a vaccine antigen since the invasive disease was almost exclusively restricted to type-b organisms, and antipolysaccharide antibodies had an important role in natural immunity. However, early observations with Hib demonstrated the limitations of plain polysaccharide as a vaccine antigen. When given, during the first 2 years of life, purified polysaccharide induced relatively low levels of serum antibodies, typically insufficient to protect against invasive disease. Following further studies with a variety of bacterial polysaccharides, and in the light of the limitations of plain polysaccharide as vaccine antigens, the Hib polysaccharide
was shown to be more immunogenic when covalently linked to a protein carrier, giving additionally boosted responses characteristic of T-dependent memory [3, 4]. Overall, with the classical vaccinology approaches many infectious diseases can be prevented. Table 1.1 reports a list of vaccines licensed for immunization in the United States Food and Drug Administration (FDA) [5].

However, the above-mentioned approaches have several limitations, including the fact that, in some cases, pathogenic microorganisms are difficult to culture in vitro and, therefore, production of live attenuated, inactivated, or subunit vaccines becomes impractical and time consuming. As a result, there are several infectious diseases for which these traditional approaches have failed and for which vaccines have not yet been developed.

### TABLE 1.1. Vaccines Licensed for Immunization and in the United States approved by FDA

<table>
<thead>
<tr>
<th><strong>Subunit Vaccines</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Anthrax</td>
</tr>
<tr>
<td>Tetanus</td>
</tr>
<tr>
<td>Diphtheria</td>
</tr>
<tr>
<td>Pertussis</td>
</tr>
<tr>
<td><em>Haemophilus influenza</em></td>
</tr>
<tr>
<td>Hepatitis B</td>
</tr>
<tr>
<td>Human papillomavirus</td>
</tr>
<tr>
<td>Meningococcus (groups A, C, Y, and W-135)</td>
</tr>
<tr>
<td>Pneumococcal</td>
</tr>
<tr>
<td>Tetanus</td>
</tr>
<tr>
<td>Typhoid&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Live Vaccines</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Tuberculosis</td>
</tr>
<tr>
<td>Measles</td>
</tr>
<tr>
<td>Mumps</td>
</tr>
<tr>
<td>Rotavirus</td>
</tr>
<tr>
<td>Varicella</td>
</tr>
<tr>
<td>Yellow fever</td>
</tr>
<tr>
<td>Herpes zoster</td>
</tr>
<tr>
<td>Smallpox</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Inactivated Vaccines</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Hepatitis A</td>
</tr>
<tr>
<td>Influenza&lt;sup&gt;u&lt;/sup&gt;</td>
</tr>
<tr>
<td>Japanese encephalitis</td>
</tr>
<tr>
<td>Plague</td>
</tr>
<tr>
<td>Polio</td>
</tr>
<tr>
<td>Rabies</td>
</tr>
</tbody>
</table>

<sup>a</sup>Additional live vaccine.
1.3 REVERSE VACCINOLOGY

1.3.1 Classical Reverse Vaccinology: The MenB Story

Conventional approaches to develop vaccines, as described above, are based on the cultivation of the microorganisms in vitro, and only abundant components can be isolated by using biochemical and microbiological methods. Although successful in many cases, these approaches have failed to provide vaccines against pathogens that did not have obvious immunodominant protective antigens [6]. With the advent of whole-genome sequencing and advances in bioinformatics, the vaccinology field has radically changed, providing the opportunity for developing novel and improved vaccines. The availability of the complete genome sequence of a free-living organism (*H. influenzae*) in 1995 [7] marked the beginning of a “genomic era,” which allowed scientists to use new approaches for vaccine design and for the treatment of bacterial infections. With this powerful sequencing technology, a new approach to identify vaccine candidates was proposed on the basis of the genomic information. This approach was called reverse vaccinology (Fig. 1.1). The novelty of reverse vaccinology was not based on growing microorganisms but on running algorithms to mine DNA (deoxyribonucleic acid) sequence information contained in the blueprint of the bacterium [8]. The first step of this *in silico* analysis is the appropriate combination of algorithms and the critical evaluation of the coding capacity. The predicted open reading frames (ORFs) are used for homology searches against a database with BLASTX, BLASTN, and TBLASTX programs to identify DNA segments with potential coding regions. Since secreted or extracellular proteins are more accessible to antibodies than are intracellular proteins, they represent ideal vaccine candidates and therefore the surface localization criterion is applied. The *in silico* approach results in the identification of

![Figure 1.1. Schematic representation and time lines of classical vaccinology in comparison to the reverse vaccinology approach. (See insert for color representation of this figure.)](image)
a large number of genes. It is, therefore, necessary to use simple procedures that allow a large number of target genes to be cloned and expressed. Once purified, the recombinant proteins are used to immunize mice. The postimmunization sera are analyzed to verify the computer-predicted surface localization of each polypeptide and their ability to elicit an immune response. The direct means to study the protective efficacy of candidate antigens is to test the immune sera in an animal model in which protection is dependent on the same effector mechanisms as in humans [9].

The reverse vaccinology approach has been applied for the first time to the bacterial pathogen Neisseria meningitides serogroup B. Although the use of vaccines based on the polysaccharide antigen has been successful for most of the species causing bacterial meningitis (H. influenzae type B, Streptococcus pneumoniae, and N. meningitidis serogroups A, C, Y, and W135), the same approach cannot be easily applied to meningococcus B. This is because the MenB polysaccharide is a polymer of α(2–8)-linked N-acetyl-neuraminic acid (or polisyalic acid), which is also present in glycoproteins of mammalian neural tissues. The poor immune response and the high risk of autoimmunity have hindered much of the research on the MenB polysaccharide [10]. An alternative approach to the MenB vaccine was based on the surface-exposed proteins contained in outer membrane vesicles (OMVs). These vaccines have been shown both to elicit serum bactericidal antibody responses and to protect against developing meningococcal disease in clinical trials [11–13]. Although these vaccines provided evidence of efficacy against the homologous strain, they show sequence and antigenic variability in their major components [14].

Thus, the genomic approach for target antigen identification was directed to develop a vaccine against serogroup B N. meningitidis [15]. Neisseria meningitidis is a Gram-negative diplococcus and an obligate human pathogen that colonizes asymptotically the upper nasopharynx tract of about 5–15% of the human population. Five serogroups (A, B, C, W135, and Y) account for virtually all of the cases of meningococcal disease [16]. N. meningitidis is the major cause of meningitis and sepsis, two devastating diseases that can kill children and young adults. Invasive meningococcal disease causes a significant public health burden worldwide, with approximately 500,000 cases and >50,000 deaths reported annually [6, 16]. MenB represents the first example of the application of reverse vaccinology and the demonstration of the power of genomic approaches for novel antigen identification. In 2000, an invasive isolate of N. meningitidis (MC58) was sequenced and analyzed to identify suitable vaccine candidates with the in silico approach described above. After discarding cytoplasmic proteins and known Neisseria antigens, 570 genes predicted to code for surface-exposed or membrane-associated proteins were identified. Successful cloning and expression was achieved for 350 proteins, which were then purified and tested for localization, immunogenicity, and protective efficacy. Of the 91 proteins found to be surface exposed, 28 were able to induce complement-mediated bactericidal antibody response, providing a strong indication of the proteins capability of inducing protective immunity [17]. Additionally, in order to test the suitability of these antigens for conferring protection against heterologous strains, the proteins were evaluated for gene presence, phase variation, and sequence conservation in a panel of genetically diverse MenB strains representative of the global diversity of the natural
N. meningitidis population [9]. This analysis yielded a handful of antigens, which were both conserved in sequence and able to elicit a cross-bactericidal antibody response against all of the strains in the panel, demonstrating that they could confer general protection against the meningococcus. To strengthen the protective activity of the single-protein antigens and to increase strain coverage, the final vaccine formulation comprises a “cocktail” of the selected antigens. This vaccine is currently in phase III clinical trials [17, 18].

1.3.2 Reverse Vaccinology Applied to the Pan-Genome Concept

In the last decade, microbial genomic sequencing has experienced an exponential growth. Sequencing of 1129 bacterial genomes have been completed and 2893 are currently in progress [19]. All of the genomic sequences are available in public databases, and they cover hundreds of species, as well as multiple pathogenic and commensal strains of the same species. Recently, an analysis of the Streptococcus agalactiae genome has led to the pan-genome definition for this pathogen [20]. It was suggested that the genomic sequence of a single strain is not genetically representative of an entire species, due to the surprising intraspecies diversity. Subsequently, in order to develop a universal vaccine with a broad range of coverage against the major circulating strains, a combination of antigens representative of different strains of the same pathogen should be included. Therefore, the classical reverse vaccinology approach has now been extended to a wider number of genomes belonging to the same species, as performed for S. agalactiae.

S. agalactiae (commonly referred to as group B Streptococcus or GBS) is an encapsulated Gram-positive coccus. GBS strains are classified into 9 serotypes according to immunogenic characteristics of the capsular polysaccharides (Ia, Ib, II, III, IV, V, VI, VII, and VIII), and approximately 10% of serotypes are nontypeable [21]. GBS was originally isolated in 1938 from animals [22], and is the main cause of bovine mastitis, an economically important problem in dairy cattle throughout the world [23]. However, invasive group B streptococcal disease emerged in the 1970s as a leading cause of neonatal morbidity and mortality in the United States [24] and represents the most common etiological agent of invasive bacterial infections (pneumonia, sepsis, and meningitis) in human neonates [25, 26]. The need for a vaccine against GBS is supported by the observation that the risk of neonatal infection is inversely proportional to the maternal antibody response specific for the GBS capsular polysaccharide. In fact, the transplacental transfer of maternal IgG antibodies protects infants from invasive group B streptococcal infection [27]. As a first approach to vaccine development, capsular polysaccharide (CPS)—tetanus toxoid conjugates against all 9 GBS serotypes were shown to induce CPS-specific IgG that is functionally active in opsonization against GBS of the homologous serotype. Clinical phase 1 and phase 2 trials of conjugate vaccines prepared with CPS from GBS types Ia, Ib, II, III, and V revealed that these preparations are safe and immunogenic in healthy adults. Although these vaccines are likely to provide coverage against the majority of GBS serotypes that cause disease in the United States, they do not offer protection against pathogenic serotypes that are more prevalent in other parts of the world (e.g., serotypes
VI and VIII, which predominate among GBS isolates from Japanese women). Hence, a universal protein-based vaccine against GBS is desirable [28]. In 2002, the whole genome sequences of 2 strains of GBS were assembled and released in the public domain. Serotype V strain 2603V/R is representative of an emerging serotype accounting for approximately 18% of early-onset cases, 14% for late-onset cases, and 31% for nonpregnant pediatric and adult cases [29, 30]. The serotype III strain (NEM316 strain) is the most common clinical serotype encountered in neonatal meningitis cases and was isolated from a fatal case of septicemia [31]. Using the sequenced strain 2603V/R as a reference, comparative genomic hybridization (CGH) was performed to circumvent the need for sequencing closely related genomes. Genomic comparisons reveal regions of loss and/or retention with respect to the reference strain. It was found that approximately 18% of the genes encoded in the sequenced strain are absent in at least one of the other 19 *S. agalactiae* strains. However, CGH experiments are limited in that they only identify the portion of the sequenced genome that is shared with the other test strains and are not able to detect genes that are absent in the reference genome. Following this work, the sequences of 6 additional strains of *S. agalactiae* were determined [31, 32]. Genome analysis of 8 clinical isolates of GBS, which are representative of the serogroup diversity and responsible for >90% of human infections in the United States, revealed the global gene repertoire of the GBS bacterial species to consist of 1806 genes present in every strain (core genome), plus 907 dispensable genes that are present in one or more but not all strains (dispensable genome) (Fig. 1.2a) [33]. In general, the core genome includes all genes responsible for the survival of the bacterium and its major phenotypic traits. By contrast, dispensable genes contribute to the species diversity and may encode supplementary biochemical pathways and functions that are not always essential for bacterial growth but that confer selective advantages, such as adaptation to different niches, antibiotic resistance, or colonization of a new host. Such genes are generally clustered on large genomic islands that are flanked by short repeated DNA sequences and are characterized by an abnormal G + C content. Investigation and functional annotation of dispensable genes reveals that hypothetical and phage- and transposon-related genes account for the vast majority of these findings (Figs. 1.2b and 1.2c). Moreover, computational predictions suggest the more genomes are sequenced, the more new genes will be found that belong to the pan-genome [20]. In fact, the pan-genome is predicted to grow about 33 new genes every time a new strain is sequenced. This profile is different from that observed for other microorganisms, such as *Bacillus anthracis*. Eight genome sequences were determined for this microorganism, but after the fourth it was verified that the number of new genes added to the pan-genome rapidly converged to zero [20, 34].

In order to identify protective antigens for GBS, computer algorithms were used to select the genes encoding putative surface-associated and secreted proteins. Among the predicted surface-exposed proteins, 396 were core genes and 193 were variable genes. Of these 589 proteins, 312 were successfully cloned and expressed in *Escherichia coli*. Then each antigen was tested in a mouse maternal immunization assay to evaluate the capacity to confer protection. No individual core proteins or a combination thereof provided high levels of immunity against a panel of GBS isolates. The best candidates, namely GBS67, GBS80, GBS104, and GBS322, identified through the reverse
Figure 1.2. Pan-genome structure. (a) Schematic representation of the pan-genome comparative analysis performed on a set of three different strain genomes. Venn diagram overlapping circles represents core, dispensable genome, and strain-specific genes. (b–c) Gene classification of a new sequenced genome according to the pan-genome structure. (See insert for color representation of this figure.)
vaccinology combined with the multigenome comparison approach, belonged to the dispensable genome with the exception of GBS322, which belongs to the core genome. The most significant protection results were achieved when the four best candidates were combined together. The four-protein cocktail conferred 59–100% protection against a panel of 12 GBS isolates, including the major serotypes, as well as 2 strains from the less common serotype VIII (81 and 94% protection) [28]. Multigenome sequencing from different streptococcal species (GBS, group A Streptococcus, and Pneumococcus) has revealed to be an important approach, not only for antigen selection but also because it has led to the discovery of important virulence factors such as pili [35–37].

1.3.3 Subtractive Reverse Vaccinology: The ExPEC Experience

Many bacterial species exist in the human gut flora, both as commensal and pathogenic strains. By comparing the genomic sequences of the two types, one can identify pathogenic specific traits. In this regard, a new concept of subtractive reverse vaccinology was applied for the first time to the pathogenic E. coli species. E. coli is a commensal member of the gastrointestinal flora of most mammals. Additionally, several facultative and obligate pathogenic variants exist that cause various types of intestinal or extraintestinal infections in humans and animals. Extraintestinal pathogenic E. coli (ExPEC) is a facultative pathogen known to cause infections such as urinary tract infection (UTI), newborn meningitis, and sepsis [38]. Thus, the development of an efficacious ExPEC vaccine would have a significant public health and economic impact, considering the increasing antibiotic resistance among ExPEC strains and the associated mortality and morbidity. Conventional vaccinology approaches (whole cell, single antigens, or polysaccharide based) have not been successful in providing a highly immunogenic, safe, and cross-protective vaccine against ExPEC strains. Recently, Moriel et al. [39] described an innovative approach based on reverse vaccinology to develop a vaccine against the extraintestinal E. coli infections.

As a first step, the neonatal meningitis-associated E. coli K1 strain IHE3034 genome was sequenced. This sequence was then compared to that of the nonpathogenic E. coli MG1655, DH10B, and W3110 strains since a vaccine directed against commensal or nonpathogenic strains could have potential implications for the equilibrium of the normal human intestinal flora. In particular, the comparison between IHE3034 and MG1655 revealed the presence of 19 genomic islands absent in the MG1655 genome. Many of these genomic islands contained the typical features of the pathogenicity islands (PAI), such as the presence of an integrase and transfer ribonucleic acid (tRNA) genes and a different G+C% content compared to the core genome. Subsequently, an in silico analysis was performed comparing IHE3034 with two other ExPEC strains, namely 536 and CFT073.

By in silico analysis, specific ExPEC-shared antigens predicted to be surface associated or secreted were selected, whereas the presence and the level of similarity of these antigens in the nonpathogenic strains mentioned before were used as exclusion criteria (hence the term subtractive). Among the predicted surface-exposed proteins, nine vaccine candidates were identified to confer protection. Interestingly, the gene
Microbial Vaccine Design: The Reverse Vaccinology Approach

Encoding for the most promising antigen ECOK1_3385 was found both in pathogenic and nonpathogenic strains, and located downstream to a type two secretion system (T2SS); in the nonpathogenic strains this T2SS is truncated, and consequently the antigen is expressed intracellularly but not secreted. By contrast, in pathogenic strains, the protein encoded by ECOK1_3385 is secreted into the supernatant, suggesting that it may play a role in virulence. Additionally, studies on the variability of ECOK1_3385 in a panel of 96 ExPEC strains showed a significant level of sequence conservation ranging from 86 to 100% identity. Furthermore, gene distribution analysis in 573 isolates representing diverse E. coli phylogeny revealed a higher proportion of ECOK1_3385 in extraintestinal and intestinal strains (67%) compared to commensal strains (30%).

Therefore, the subtractive reverse vaccinology approach used in this work has lead to the identification of nine potential antigens present in ExPEC and intestinal E. coli, suggesting that they may be useful for a broadly cross-reactive E. coli vaccine.

1.4 Vaccine Design: From Conventional Vaccinology to the Postgenomic Era Through Reverse Vaccinology

Despite antimicrobial therapies, infectious diseases remain the leading cause of death worldwide [40]. Vaccination is a powerful public health tool, with the global eradication of diseases such as smallpox and poliomyelitis (in the Western Hemisphere). Most of the current vaccines have been developed by adaptation of living organisms to growth conditions that attenuate their virulence, by preparation of suspensions of killed microbes or through the concentration and purification of proteins or polysaccharides from pathogens [41]. With the development of DNA sequencing technology, in the late 1990s the way for vaccine candidate identification has radically changed. The traditional reverse vaccinology approach has been integrated and flanked by the multigenomes comparison and after that by postgenomic technologies, which comprise the combination of transcriptomics, proteomics, and immunomics. Postgenomic strategies are referred to as functional genomics, complementing the in silico antigen discovery, not only by directly examining the genetic content but also the transcription and expression profiles of pathogens.

In particular, the study of global changes in bacterial gene expression is essential for understanding pathogenesis and survival in the host. For vaccine antigen discovery, it is important to know which genes are regulated in vivo, during infection, because they could represent protective vaccine candidates. The availability of complete genome sequences stimulated the application of complementary DNA (cDNA) microarrays in identifying the genes involved in microbial virulence, pathogenesis, and therefore useful for vaccine design. cDNA microarrays are used to study the gene expression profile of tens of thousands of genes in a single experiment. The first example where microarray-based transcriptional profiling was successfully used to identify potential vaccine candidates was in the case of MenB [42]. For the identification of new vaccine targets for MenB, transcriptional profiling studies were performed under experimental conditions mimicking certain aspects of host–pathogen interactions, such as adherence to host epithelial cells and exposure to human serum. In this
study, RNA was prepared from adherent and nonadherent bacteria, and the two RNA preparations were comparatively analyzed by DNA microarrays bearing the complete complement of polymerase chain reaction (PCR)-amplified MenB genes [42, 43]. Twelve proteins, whose transcription was found to be particularly activated during adhesion, were expressed in *E. coli*, purified, and used to produce antisera in mice. Five sera showed bactericidal activity against different strains.

While the transcriptomics approach gives a global expression profile of the gene up- and down-regulated during a particular growth condition, the proteomics approach refers to a high-throughput analysis of the complete complement of proteins, allowing the selection of a small group of antigens expressed at high levels and surface exposed [44]. In this approach, following separation by two-dimensional (2D) gel electrophoresis, the proteins are represented as fine spots on a gel that are isolated for further analysis. Mass spectrometric techniques like matrix-assisted laser desorption ionization, time of flight (MALDI–TOF) and tandem mass spectrometry (MS/MS) are used for peptide mass and sequence analysis of a given protein spot. Advancements in proteomic approaches have enabled researchers to explore this novel strategy in order to identify vaccine targets and proteins of therapeutic interest [43]. A powerful and attractive application of proteomics to identify vaccine candidates was recently reported by Grandi and colleagues, who analyzed the surface proteome of *Streptococcus pyogenes* (mainly referred to as group A *Streptococcus*, GAS) [45]. Whole cells were treated with proteases to selectively digest cell surface proteins. The resultant peptides were subsequently identified by mass spectrometry. Application of this technique to the sequenced *S. pyogenes* strain M1_SF370 resulted in the identification of 68 predicted surface-associated proteins. The approach proved to be rapid and highly selective in that the large majority (>90%) of the identified proteins fell into the categories of cell wall proteins, lipoproteins, membrane proteins, and secreted proteins. Furthermore, the method allowed for a semiquantitative evaluation of protein expression. Indeed, the number of peptides identified from a given protein correlated with the extent of its recognition by specific antibodies, as judged by fluorescence-activated cell sorting analysis. Interestingly, the list of surface-associated proteins included most of the published GAS protective antigens, as well as new protective components such as the cell envelope proteinase Spy0416, a new attracting protein for its important role in pathogenesis [45]. Recently, a similar approach has been used to identify new protective antigens in *S. agalactiae*. The authors showed that on the surface of the hypervirulent GBS COH1 strain are present 43 major proteins belonging to the families of cell wall proteins, lipoproteins, and membrane proteins. As in the case of GAS, the proteins identified comprise all protective antigens so far described in the literature as well as a new antigen, SAN_1485, which appears to be highly protective in the active maternal immunization mouse model. These data confirm the effectiveness of protease digestion coupled to mass spectrometry for the identification of surface-exposed antigens in Gram-positive bacteria and demonstrate the power of this technology for the rapid discovery of new vaccines [44]. Therefore, a combination of proteomic-based approaches and serological analysis can identify potential vaccine candidates and provide effective validation of these candidates. A number of methods have been developed to enable the high-throughput display of
the proteome of a pathogen to the host immune system. Immunomics [also known as serological proteome analysis (SERPA)] combines proteomic-based approaches with serological analysis, and it has been widely applied for antigen discovery and vaccine development. Applied to Staphylococcus aureus, this approach led to the identification of 15 highly immunogenic proteins, including known and novel vaccine candidates [46].

A number of infectious diseases are still waiting for an effective vaccine and new approaches are required to identify new promising vaccine candidates. Structure determination by high-throughput methods such as X-ray crystallography, nuclear magnetic resonance (NMR) spectroscopy, or molecular electron microscopy together with the availability of a huge number of genomic sequence data has led to the growth of the field of structural genomics. The primary aim of structural genomics is to determine the three-dimensional profile of all important biological macromolecules, with a primary focus on proteins. The secondary goal is to decrease the average cost of structure determination through high-throughput methods for protein production and structure determination [47]. A complete view of protein folds may help in assigning function for hypothetical proteins, with the prediction of the protein structure by computational approaches such as homology modeling, founded on similarities with known protein structures for constructing atomic-resolution models from amino acid sequences. Structural genomics was used in the past two decades in the rational design of important chemotherapeutics, such as the human immuno deficiency virus (HIV) [48] or influenza drugs [49], on the basis of the analysis of the complex structures from the target protein. To be effective, a vaccine must induce a strong protective immune response from B and T cells. Since antibodies, by specific recognition of antigen epitopes, are an effective line of defense in preventing infectious diseases, understanding the antibody–epitope interaction establishes a basis for the rational design of vaccines, leading to the development of epitope-driven vaccines, containing only selected epitopes that have been already described, for example, in the case of cancer [50]. Overall, a complete understanding of protein structure and function and the study of functional complexes between a given macromolecule and its effectors in the host will facilitate the rational design of vaccines.

1.5 CONCLUSIONS

We have come a long way since Edward Jenner’s landmark experiment with cowpox, an event that over the past two centuries has developed into the field of vaccinology. The current “-omics” revolution has provided researchers with advanced technologies that are decreasing the time it takes to identify and design target antigens. These high-throughput technologies have proved to be powerful in the discovery of vaccine candidates for group B meningococcus and now more recently many other bacterial pathogens. However, the decision of which of these large number of candidates to take forward in clinical trials still remains a challenge [43]. The future success of vaccines will not only depend on the advancement of scientific platforms but also the interdisciplinary involvement of researchers in different fields such as structural
biology, physical chemistry, epidemiology, and molecular immunology. The success of vaccines will be possible only through innovative ideas that will lead to fundamental breakthroughs.

REFERENCES


