University of Florence

International Doctorate in Structural Biology
Cycle XXII (2007-2009)

A “Structural Vaccinology” approach on vaccine candidates against *Neisseria Meningitidis*

Ph.D. thesis of

Sara Dragonetti

Tutor | Coordinator

Prof. Lucia Banci | Prof. Claudio Luchinat

S.S.D. CHIM/03

This thesis has been approved by the University of Florence, the University of Frankfurt and the Utrecht University

Il vero sapiente è colui che sa di non sapere

(Socrate)

A Caterina

CONTENTS

Chapter 1

Introduction

1.1 The Post-genomic vaccine development

For over a century, vaccines were developed according to Pasteur's principles of isolating, inactivating and injecting the causative agent of an infectious disease. The availability of a complete microbial genome sequence in 1995 marked the beginning of the genomic era that allowed scientists to change the approach for vaccine development starting from genomic information; this process was named "reverse vaccinology" (Fig.1).

This can be considered as one of the most powerful examples of how genomic information can be used to develop therapeutic interventions, which were difficult or impossible to obtain with conventional approaches used in the past to produce vaccines. The latter approaches were based indeed on the cultivation of the microorganism in laboratory conditions from which single components were isolated individually by using biochemical, microbiological and serological methods. Each antigen was produced in pure form either directly from the bacterium or using the recombinant DNA technology, and finally tested for its ability to induce an immune response. However, although successful in many cases, this approach presented several limitations.

This approach needed to grow the pathogen in vitro, so not applicable to uncultivable microorganisms, and in many cases the antigens expressed during infection are not produced in laboratory conditions. Moreover, the proteins that are most abundant and easily purified are not necessarily protective antigens and, in any case, only a few molecules can be isolated and tested simultaneously. In conclusion, this method can employ many years to identify a protective and useful antigen, and has failed to provide a vaccine against those pathogens that did not have obvious immunodominant protective antigens.

As the genomic era progressed, it becomes apparent that multi-strain genome analysis is fundamental to the design of universal vaccines. The availability of the complete genome of microorganisms combined with novel

genome-based approaches (proteomic and microarray analysis) has introduced new perspectives in vaccine research (1).

For the first time, the genome sequences represent an inclusive virtual catalogue of all the potential vaccine candidates from which it is possible to select the molecules that are likely more effective, regardless of their abundance, or if they are expressed in vivo or in vitro conditions. Pan-genome approach represents an advance in the use of reverse vaccinology, since it highlights the potential of looking at more than one genome for the same bacterial species to overcome the problems represented by gene presence and variability (Fig.1).

In the post-genomic era the next challenge of the vaccines study is the merging of the vaccinology with structural biology (1).

1.1.2 The future of the vaccinology: the structural approach

Structural genomics, a high-throughput application of structural biology which utilizes X-ray crystallography or nuclear magnetic resonance (NMR) spectroscopy for protein structures determination, is increasingly being applied to vaccines and drugs development as a result of the explosion of genome and proteome data, and has provided continuing improvements in the fields of protein expression, purification, and structural determination (2).

Structural biology studies in the past few decades allowed the identification and understanding of many basic principles of proteins, nucleic acid structures, molecular machines and viruses. Besides, structural biology demonstrated to have enormous biomedical applications; indeed structure-based development of drugs is already well integrated into the pharmaceutical industry and lead to the identification of important drugs.

Structural biology can also significantly aid the rational design of future vaccines. The application of structural biology principles in the vaccinology field resulted in the creation of a powerful new approach named structural

vaccinology (Fig. 1), which gave a tremendous help in the development of vaccines against diseases that we could not defeat using other approaches.

Structural studies allow indeed the atomic resolution of the structures of potential antigens and, through their structure, the rational design of target epitopes to use as vaccine candidates will be performed.

Several examples of the application of structural biology to the design of effective vaccines are already in place against several pathogens that most heavily afflict global health, like for example meningococcus.

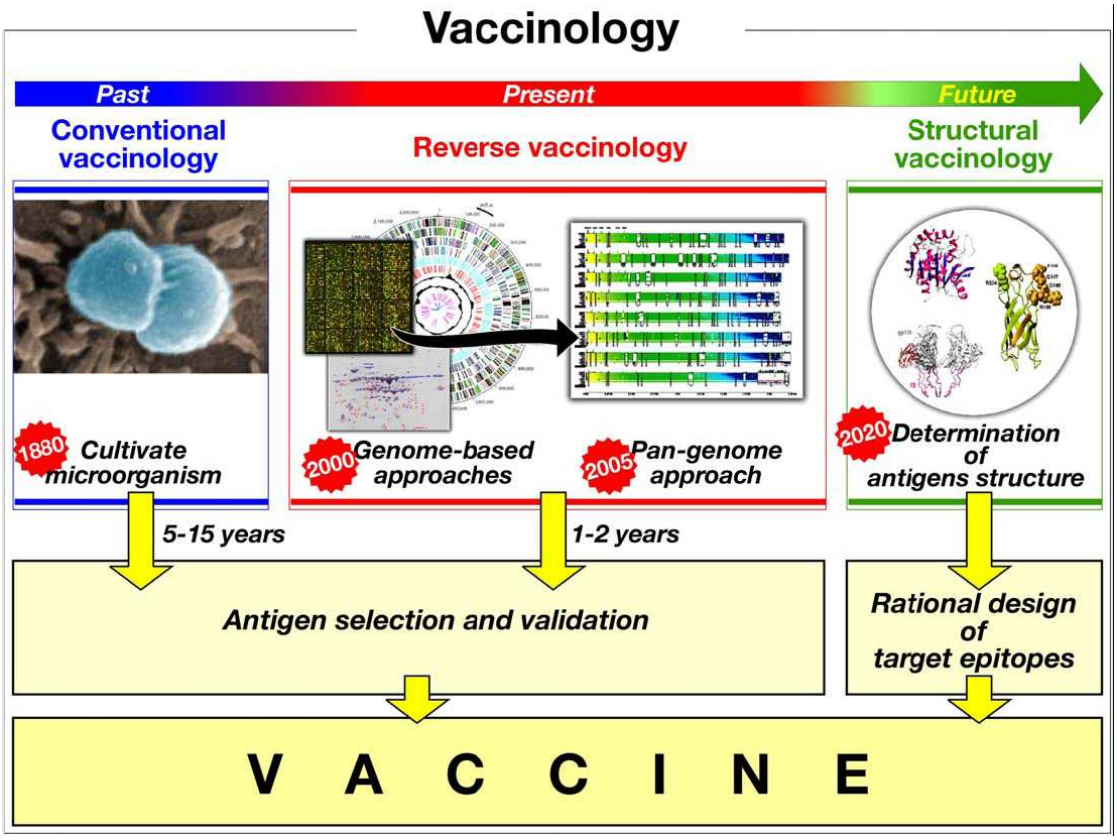


Figure 1. The vaccinology field. Past: Conventional approaches to produce vaccines. Present: The genomic era has completely changed the way to design vaccines (Reverse Vaccinology); Future: The structural biology applied to vaccinology (Structural Vaccinology) (Rappuoli, 2006).

1.2 Historical background

The discovery that various diseases are caused by microorganism occurred in twilight of nineteenth century and symbolized a milestone in the studies of infectious diseases. In 1879 Louis Pasteur demonstrated the link between microbiological organisms and puerperal sepsis.

Description of illness resembling meningococcal disease dates back to the 16th century. It was first described by Vieusseux in 1805 after an epidemic of meningitis in Geneva (3) and in 1882 the pathogen responsible for this illness was first isolated from the cerebrospinal fluid (CSF) of an infected patient by the Italian pathologists Marchiafava and Celli (4).

Anton Weichselbaum in 1887 first identified the bacterium causing meningococcal disease in the CSF of six of eight patients of bacterial meningitis and the bacterium was named *Neisseria intracellularis* (5).

In 1909, immunologically distinct serotypes of the meningococcus were identified. This established the basis for serum therapy, which was instituted in 1913 by Flexner in the USA and through which mortality was reduced from nearly 100% to 30% (6, 7).

Subsequently the discovery of sulfonamides and other antimicrobial agents dramatically reduced the incidence of meningitis and the advent of penicillin G greatly improved the treatment of meningococcal meningitis and sepsis, resulting fundamental in decreasing mortality rates.

Despite treatment with appropriate antimicrobial agents and optimal medical care, the overall case fatality rates remained relatively stable over the past 20 years, at 9 to 12%, with a rate of up to 40% among patients with meningococcal sepsis (8).

About the 20% of survivors of meningococcal disease have sequelae, such as hearing loss, neurological disability, or loss of a limb (9).

1.3 Microbiology and Epidemiology

Neisseria meningitidis is a Gram-negative, capsulated, aerobio diplococco, member of to the bacterial family *Neisseriaceae*. It is surrounded by an outer

membrane composed of lipids, outer membrane proteins (OMPs) and lipopolysaccharides. Pathogenic meningococci are enveloped by a polysaccharide capsule attached to this outer membrane. *Neisseria meningitidis* is the major cause of bacterial septicemia and the most common cause of bacterial meningitis, that can have catastrophic consequences in individuals and can become epidemic in developed and developing countries (Buysse et al 2008). Meningococci can be classified on the basis of the chemical composition of different capsular polysaccharide structures into 13 serogroups: A, B, C, H, I, K, L, M, W135, X, Y, Z and 29E (serogroup D is no longer recognized). Among them, only five (A, B, C, Y and W135) have been associated with the disease and so considered as pathogenic serogroups.
Meningococci are further classified into 20 serotypes and 10 subtypes (on the basis on the antigencity the outer membrane porins PorA and PorB) and 13 immunotypes (on the basis of lipooligosaccharide structure) (10, 11, 12).
Meningitis is an inflammatory disease of the leptomeninges, the tissues surrounding the brain and spinal cord. The meninges consist of three parts: the pia, arachnoid, and dura matres. Meningitis reflects inflammation of the arachnoid mater and the cerebrospinal fluid (CSF) in both the subarachnoid space and in the cerebral ventricles.

Epidemic rates of meningococcal disease varies from <1-3/100,000 in many developed nations to 10-25/100,000 in some developing countries. This difference in attack rates reflects the difference in pathogenic properties of *N. meningitidis* strains and differences in socioeconomic and environmental conditions. The proportion of cases caused by each serogroup varies by age group; more than half of cases among infants aged <1 year are caused by serogroup B, for which no vaccine is available (13, 14).
Of the five common pathogenic serogroups (A, B, C, Y and W135) responsible for about 90% of infections caused by *N. meningitidis*, serogroups A, B and C account for most cases of meningococcal disease throughout the world, with

serogroups A and C predominating in Asia and Africa and serogroups B and C responsible for the majority of cases in Europe and the Americas (15, 16, 17, 18, 19) (figure 2).

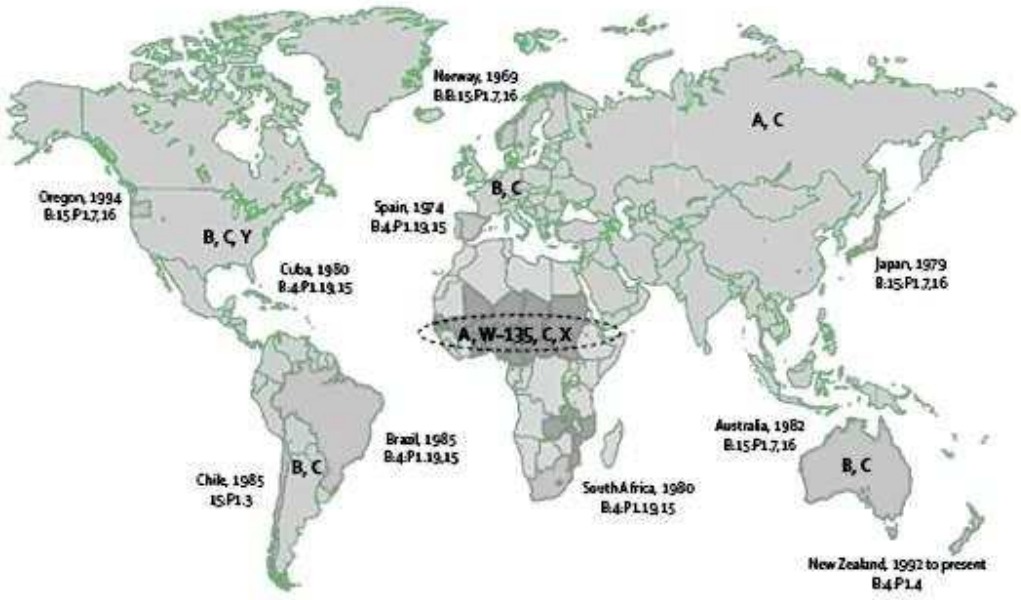


Figure 2.

Worldwide distribution of major meningococcal serogroups (Sthephens, 2007).

Rates of meningococcal disease are highest for young children at 6 to 24 months of age, when maternal antibody has disappeared and increase for adolescents and young adults. In endemic situations serogroup B is most common in infants, serogroup C in adolescents and serogroup B and Y in older adults.

1.4 Classification and Antigenicity

Meningococcal virulence is related to several factors, such as capsular polisaccharide, outer membrane proteins (OMPs) and endotoxin or lipo-

oligosaccharide (LPS). A schematic representation of bacterial compartment and of main components is reported in Figure 3.

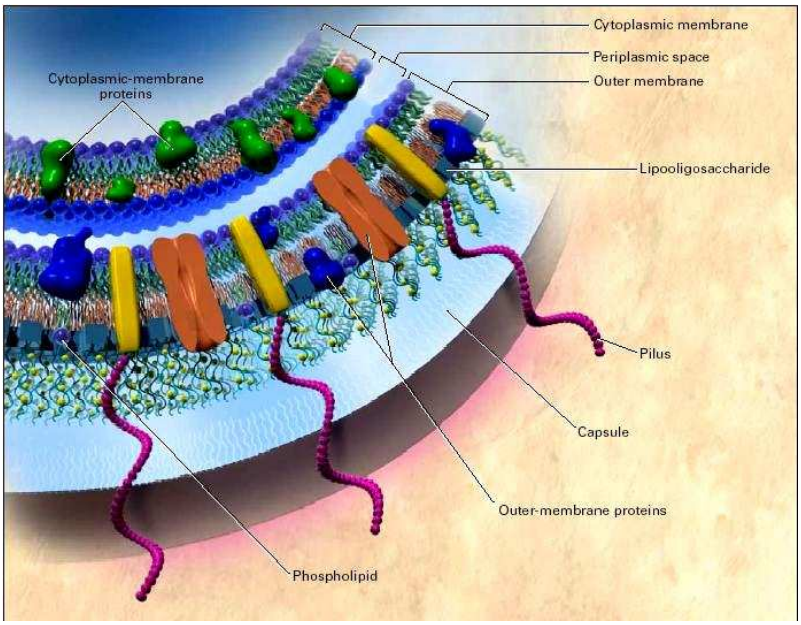


Figure 3. Section of the Meningococcal cell membrane.

The polysaccharide capsule, which covers the bacterium surface, is the major virulence factor. It is a homopolymer or a heteropolymer consisting of monosaccharide, disaccharide, or trisaccharide repeating units. Except for serogroup A, the capsules are composed of sialic acid (N- acetyl neuraminic acid, NANA) derivatives.

The peculiarity of the membrane is to allow the bacterium to survive in the blood stream and to confer resistance to phagocytosis and complement-mediated lysis. As a consequence, acapsulated bacteria are inactive *in vivo* since cannot persist in the blood.

The outer membrane is composed by many proteins that contribute to meningococcal virulence, such as Pili which are filamentous structures located on the meningococcus surface, essential in mediating the initial interaction of

meningococci and both epithelial and endothelial cells. Meningococcus also expresses variable proteins such as Opa and Opc, which are important in adhesion to and invasion into eukaryotic cells and outer membrane porins, PorA and PorB, which are involved in host-cells interaction and as target for bactericidal antibodies.

1.5 Mechanisms of infection

The human naso-oropharyngeal mucosa is the only natural reservoir of *N. meningitidis.*

Three crucial steps are involved in the development of the disease:

1. The crossing by the bacterium of the mucosal barrier;
2. The invasion in the blood stream;
3. The crossing of the blood-brain-barrier and final reaching of the meninges.

Meningococci are transferred from one person to another by direct contact or via droplets for a distance up to 1 m. During periods of endemic infection, 8 to 20% of adults are asymptomatic nasopharyngeal carriers of strains of *N. meningitidis,* most of which are not pathogenic (20, 21).

The duration of meningococcal carriage can vary from days to many months. Only under particular conditions and in susceptible individuals, *N. meningitidis* penetrates the mucosa and gains access to the bloodstream, causing systemic disease (22). In most persons, nasopharyngeal carriage is an immunizing process, resulting in a systemic protective antibody response (23).

A general scheme of the steps of invasion of meningococcus is shown on figure 4.

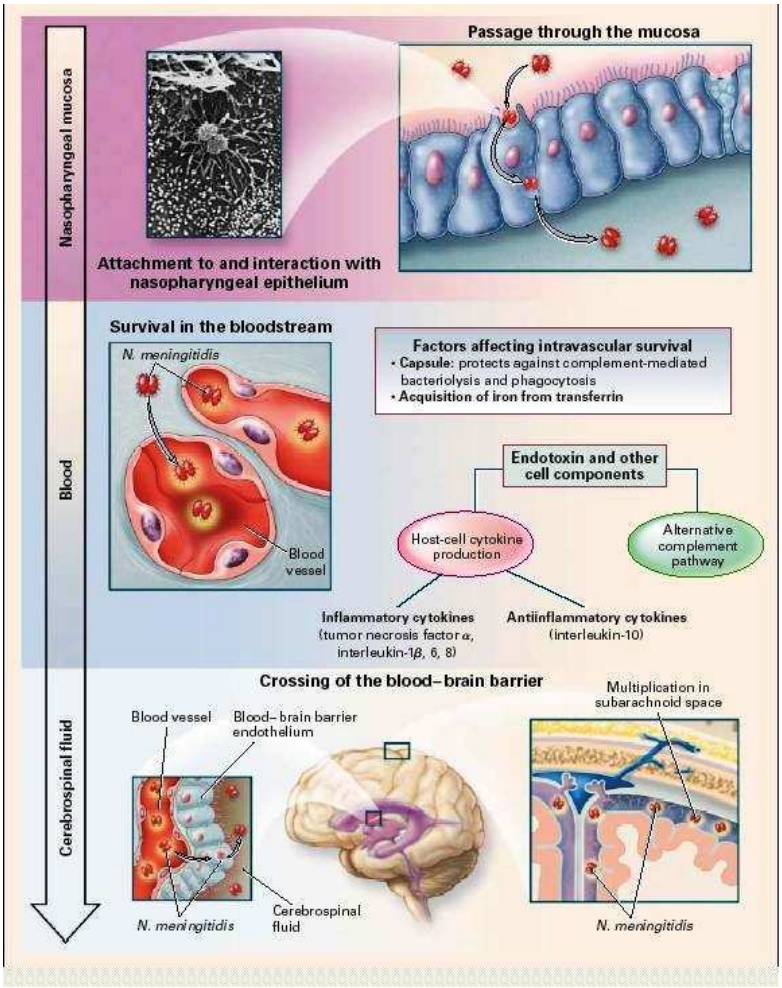


Figure 4.

Colonization of nasopharinx by *N. meningitidis* and entry into the bloodstream and cerebrospinal fluid (Rosenstein, 2001).

From the bloodstream, the bacterium can cross the blood-brain barrier and spreads to the celebrospinal fluid, causing meningitis. The mechanisms behind the propensity of meningococci to invade the meninges and their passage across the blood-brain barrier are poorly understood. Neurological damage in meningitis may be caused by a mixture of direct bacterial toxicity, indirect inflammatory processes and cerebral oedema.

1.6 Clinical manifestations, treatment and prevention

Initial general symptoms like fatigue, fever and headache, can rapidly progress to septicemia, coma and death, which arises predominantly by celebral oedema.

Currently an antibiotic therapy represent the elected treatment for meningitis, since effective antibiotics immediately stop the proliferation of bacteria. Antibiotic therapy should be started as early as possible, in order to avoid the bacteria proliferation and the progression of internal lesions. Benzylpenicillin, cefotaxime, ceftriaxone and chloramphenicol are effective antibiotics. In particular Chloramphenicol is bactericidal for meningococcus and penetrates the blood brain barrier more effectively.

1.7 Vaccines

Although antimicrobial therapy has contributed to reduce the mortality rate of 65-80%, cases of fulminant meningococcemia are unfortunately generally hopeless.

Prevention of disease can effectively be accomplished by vaccination. Immunization was made possible in 1969 when it was discovered that protection from disease correlates with the presence in the serum of antibodies able to induce complement-mediated killing of bacteria, and that purified capsular polysaccharide was able to induce them.

Tetravalent vaccines against serogroups A, C, W-135, and Y have been available since 1984 (24, 25).

Although efficacious in adolescents and adults, this vaccine is poorly immunogenic in infants and do not induce immunological memory.

Unfortunately, the conjugate approach cannot be easily applied to serogroup B (26) because the B capsular polysaccharide is a polymer of *N*-acetyl neuraminic acid or polysialic acid, a molecule which is identical to a widely distributed human carbohydrate. Being the serogroup B polysaccharide identical to a self-antigen, man is immunologically tolerant to it and therefore it

fails to induce an effective immune response. On the other hand, an immune response against it would risk to induce autoimmunity (27, 28).

To develop a vaccine against serogroup B, surface-exposed proteins contained in outer membrane vesicles (OMVs) have been used (29, 30). These vaccines elicit serum bactericidal antibody responses and protect against meningococcal disease (31, 32). However, while they induce complement-mediated bactericidal antibodies against the homologous strain, they fail to induce bactericidal antibodies against heterologous strains (33).

Today, the development of vaccines against serogroup B (which is responsible for 40% of all cases in USA and 45-80% in Europe) remains a challenge, since conventional biochemical and microbiological approaches have so far failed.

1.7.1 The Reverse vaccinology

The availability of the whole genome has permitted the development of a new strategy called “reverse vaccinology” (34): the genome represents a list of virtually of all the protein antigens that a pathogen can express at any time. It becomes possible to choose potentially surface exposed proteins in a reverse manner, starting from the genome rather than from the microorganism, with the great advantage that cultivation of dangerous microorganisms is not needed anymore. Consequently the process can be equally applied to cultivable and non-cultivable microorganisms.

This approach, starting from the genomic sequence by computer analysis, predicts those antigens that are most likely to be vaccine candidates (34). The feasibility of the approach relies heavily on the avaibility of a high-throughput system to screen protective immunity.

On figure 5 a general scheme of the steps describing the reverse vaccinology is represented.

Fragments of DNA are screened by computer analysis and novel genes are predicted to code for, for example, surface-exposed proteins.

These are cloned and express in *Escherichia Coli*, purified and used to immunized mice. The sera obtained are used to confirm the surface exposure of the proteins by ELISA and FACS analysis and tested for the ability to induce complement mediated in vitro killing of bacteria.
This approach permitted to discover novel surface exposed proteins which are able to induce bactericidal antibodies.

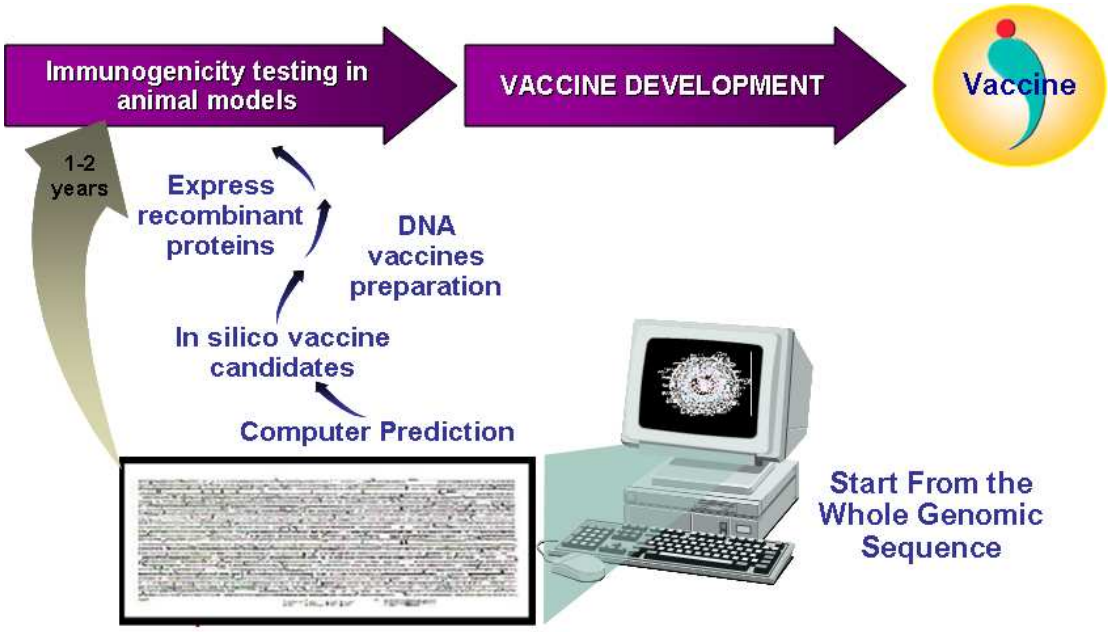


Figure 5. Schematic representation of Reverse Vaccinology approach for the identification of new vaccines candidates (Rappuoli, 2001).

1.7.2 Reverse vaccinology and the structural vaccinology approaches for Meningococcus B (MenB)

Group B meningococcus (MenB) represents the first example of a successful application of reverse vaccinology.

The antigens selected by reverse vaccinology were choosen based on their ability to induce broad protection.

Novartis Vaccines combined five antigens into a multicomponent vaccine, named 5CVMB (5 Component Vaccine against MenB), that entered Phase III clinical trials in 2008 (35, 36), trusting that this would increase the breadth of vaccine coverage and avoid selection of escape mutants. This vaccine, formulated with an adjuvant suitable for human use as alum, covers 78% of the strains and coverage can be increased to major 95% of the strains by using other adjuvants.

The major antigen molecule of this universal vaccine against meningococcus B is the factor H binding protein (fHbp), a 27-kDa surface exposed lipoprotein of *Neisseria meningitidis* discovered while screening the bacterial genome for vaccine candidates.

1.7.3 Factor H binding protein (fHbp), a new antigen of *N.Meningitidis*

Factor H-binding protein (fHbp) is essential for pathogenesis because it allows the bacterium to survive and grow in human blood by binding the human complement factor H (hfH).

hfH is one of the most important complement component in the frame of the immunity system and allows humans to distinguish between invasive agents and self-tissues by coating the latter and avoiding the complement attack.

One of the most popular strategies among pathogens to avoid complement system consists in covering themselves with factor H in order to mimick the host surface. fHbp of *N. meningitidis* is a typical example of this class of molecules.

The protein is expressed by all the pathogenic strains of *Neisseria meningitidis* and can be classified in three distinct sequence variants (37) with levels of sequence similarity ranging from 62.8% between variants 1 and 3 to 74.1% between variants 1 and 2 and 84.7% between variants 2 and 3 (figure 6).

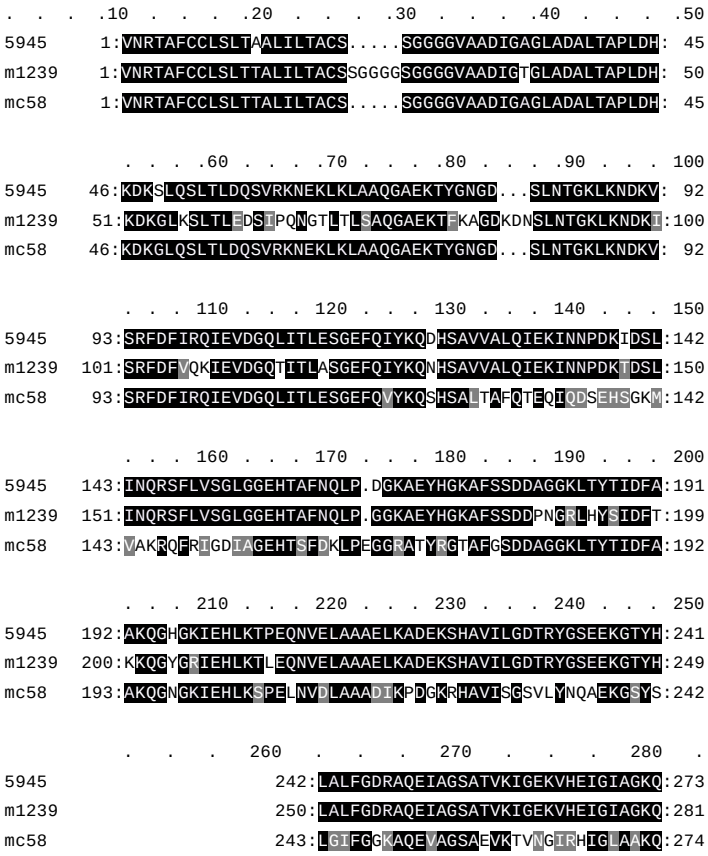


Figure 6. Aminoacidic sequence alignment of variant 1, 2, 3 of fHbp. In white, grey and black are indicated not conserved, conserved and identical residues respectively.

As already mentioned before the structural approach as been applied to MenB and through NMR spectroscopy in 2006 the solution structure of the carboxyl terminal domain of fHbp ($fHbp_c$) was determined (38).

The main project of my PhD consisted in the structural study of fHbp full length, being the latter the major component of the universal vaccine; the structural approach applied to *N.Meningitidis* allowed to obtain important information about this crucial antigen and to subsequentely provide the basis to

understand the molecular mechanisms behind the function of this important molecule.

1.8 Steps of the research

My research was first devote to the study of the interaction between $fHbp_C$ and MAb502, a bactericidal monoclonal antibody elicited in mice by fHbp, using a combination of serological data, NMR measurements and docking calculations. The data showed that MAb502 recognizes, within a well-defined area of the protein, a conformational epitope, whose identification provides valuable information for the rational engineering of fHbp, aimed to create a chimeric protein presenting a wider collection of neutralizing epitopes to the host immune system (chapter 3).

After that, the following determination of the solution structure of the full-length fHbp variant 1 by NMR spectroscopy, as anticipate before, improved the knowledge about the distribution of protective epitopes on the protein surface and provided useful indications on the possible localization of the binding site of the factor H protein. The protein is composed of two independent domains connected by a short link. The structure therefore provides the basis for designing improved vaccine molecules (chapter 4).

The subsequent investigation through NMR measurements of the interaction between fHbp and a construct of human factor H comprising domains 5, 6 and 7 (hfH567) allowed to predict residues involved in the binding with hfH (chapter 5).

My study on fHbp proceeded with the investigation on the interaction between the protein and the monoclonal antibody JAR4, a mAb that was isolated from a mouse immunized with recombinant fHbp in the variant 1 group (39).

This mAb was not bactericidal with human complement when tested alone (39); however, it elicited cooperative, complement mediated bactericidal activity with other non-bactericidal mAbs specific for fHbp in the variant 1 or 2 group (40, 41) (chapter 6).

After that, in order to identify epitopes that are variant-specific, fHbp variant 1 mutants were performed and, at present, I am studying 10 mutants of the fHbp. Then, the determination of the solution structure of other variants of fHbp (in my particular case, variant 3) will allowed to proceed on the investigation of the interaction with other mAbs.

The aim of the global project is to obtain a vaccine which could be cross-reactive and immunogenic against the most of meningococcal strains; for this reason we are interesting in studying how fHbp interacts with several mAbs, which are residues involved in this interaction and how it would possible to create a chimeric molecule containing an epitope able to cover the most possible strains of Meningococcus; indeed structural vaccinology could represent a powerful tool for the rational design or modification of vaccine antigens to improve their immunogenicity and safety.

References

1. Serruto, D., Rappuoli, R., 2006, *FEBS letters,* 580, 2985-2992.

2. Lundstrom, K., 2007, *J Cell Mol Med* , 11, 224–238.

3. Description of illness resembling meningococcal disease dates back to the 16th century. Meningococcal disease was described by Vieusseux in 1805 during an outbreak with 33 deaths in the vicinity of Geneva, Switzerland.

4. Marchiafava E, Celli A. Spra i micrococchi della meningite cerebrospinale epidemica. *Gazz degli Ospedali*, 1884, 5:59.

5. Weichselbaum A. Ueber die Aetiologie der akuten Meningitis cerebrospinalis. *Fortschr Med* , 1887, 5, 573-83.

6. Flexner S., 1913, *J Exp Med*; 17, 553-76.

7. Jochmann G. Versuche zur Serodiagnostik, 1906, *Dtsch Med Wochenschr*; 1:788.

8. Rosenstein NE, Perkins BA, 2000, *Pediatr Clin North Am*;47:337-52.

9. Kirsch EA, Barton RP, Kitchen L, Giroir BP, 1996, Pediatr Infect Dis J;15, 967-79.

10. Frasch CE, Zollinger WD, Poolman JT, 1985, *Rev Infect Dis*;7, 504-10.

11. Scholten RJ, Kuipers B, Valkenburg HA, Dankert J, Zollinger WD, Poolman JT., 1994, *J Med Microbiol.*, 41, 236-43.

12. Mandrell RE, Zollinger WD., 1977, *Infect Immun*., 16, 471-5.

13. Connolly M, Noah N. Is group C meningococcal disease increasing in Europe? A report of surveillance of meningococcal infection in Europe 1993-6. Epidemiol Infect, 1999;122:41-9.

14. Fischer M, Perkins BA. *Neisseria meningitidis* serogroup B:emergence of the ET-5 complex. *Semin Pediatr Infect Dis* ,1997;8, 50-6.

15. Caugant DA. Population, APMIS, 1998, 106, 505-25.

16. Achtman M. Global epidemiology of meningococcal disease. In: Meningococcal disease. Cartwright K, Editors. John Wiley & Sons Ltd: Chichester, United Kingdom; 1995. p. 159-75.

17. Schwartz B, Moore PS, Broome CV., 1989, *Clin Microbiol Rev*, 2, 118-24.

18. World Health Organization Working Group. Control of epidemic meningococcal diseases: WHO practical guidelines. Lyon, France: Edition Foundation Marcel Merieux; 1995.

19. Rosenstein NE, Perkins BA, Stephens DS, Lefkowitz L, Cartter ML, Danila R, *et al*., 1999, *J Infect Dis*, , 180, 1894-901.

20. van Deuren M, Brandtzaeg P, van der Meer JW., 2000, Clin Microbiol Rev, 13, 144-66.

21. Cartwright KAV. Bacterial meningitis, Chapter 17. In : Topley and Wilson's Microbiology and Microbial Infections, 9th edn. Collier L, Balows A, Sussman M, editors. Oxford University Press: New York; 1996, p. 299-318.

22. Aycock WL, Mueller JH., 1950, *Bacteriol Rev*, 14, 115-60.

23. Stephens DS., *Lancet,* 1999, 353, 941-2.

24. Gotschlich, E.C., T.Y. Liu, and M.S. Artenstein, 1969, J. Exp. Med. 129: 1349–1365.

25. Lepow, M.L., J. Beeler, M. Randolph, J.S. Samuelson, and W.A. Hankins, 1986, *J. Infect. Dis.* 154, 1033–1036.

26. Nakai, K., and M. Kanehisa, 1991, *Proteins,* 11, 95–110.

27. Hayrinen, J., H. Jennings, H.V. Raff, G. Rougon, N. Hanai, R. Gerardy-Schahn, and J. Finne, 1995, *J. Infect. Dis.,* 171, 1481–1490.

28. Finne, J., D. Bitter-Suermann, C. Goridis, and U. Finne. 1987, *J. Immunol.,* 138, 4402–4407.

29. Sierra, G.V., H.C. Campa, N.M. Varcacel, I.L., Garcia, P.L. Izquierdo, P.F. Sotolongo, G.V. Casanueva, C.O. Rico, C.R. Rodriguez, and M.H. Terry. 1991. Vaccine against group B *Neisseria meningitidis*: protection trial and mass vaccination results in Cuba. *NIPH Ann.* 14:195–207.

30. Bjune, G. E.A. Hoiby, J.K. Gronnesby, O. Arnesen, J.H. Fredriksen, A. Halstensen, E. Holten, A.K. Lindbak, H. Nokleby, E. Rosenqvist, et al., 1991, *Lancet.* 338:1093–1096.

31. Tappero, J.W., R. Lagos, A.M. Ballesteros, B. Plikaytis, D. Williams, J. Dykes, L.L. Gheesling, G.M. Carlone, E.A. Hoiby, J. Holst, et al., 1999, *J. Am. Med. Assoc.* 281, 1520–1527.

32. Perkins, B.A., K. Jonsdottir, H. Briem, E. Griffiths, B.D. Plikaytis, E.A. Hoiby, E. Rosenqvist, J. Holst, H. Nokleby, F. Sotolongo, et al., 1998, *J. Infect. Dis.,* 177, 683–691.

33. Jódar, L., I.M. Feavers, D. Salisbury, and D.M. Granoff, 2002, *Lancet*., 359, 1499–1508.

34. Rappuoli R., 2001, *Vaccine*,17-19, 2688-91.

35. Giuliani MM, Adu-Bobie J, Comanducci M, Arico B, Savino S, et al., 2006, *Proc Natl Acad Sci*, 103, 10834–10839.

36. Rappuoli R., 2008, The application of reverse vaccinology, Novartis MenB vaccine developed by design. 16th International Pathogenic Neisseria Conference, Rotterdam, The Netherlands: http://www.IPNC2008.org. Abstr. 81 p.

37. Masignani, V., Comanducci, M., Giuliani, M. M., Bambini, S., Adu-Bobie, J., Arico, B., Brunelli, B., Pieri, A., Santini, L., Savino, S., Serruto, D., Litt, D., Kroll, S., Welsch, J. A., Granoff, D. M., Rappuoli, R. & Pizza, M., 2003, *J. Exp. Med.,* 197, 789-799.

38. Cantini, F., Savino, S., Scarselli, M., Masignani, V., Pizza, M., Romagnoli, G., Swennen, E., Veggi, D., Banci, L. & Rappuoli, R., 2006, *J. Biol. Chem.,* 281, 7220-722.

39. Welsch, J. A., Rossi, R., Comanducci, M. & Granoff, D. M., 2004, *J Immunol,* 172, 5606-5615.

40. Giuliani, M. M., Santini, L., Brunelli, B., Biolchi, A., Arico, B., Di Marcello, F., Cartocci, E., Comanducci, M., Masignani, V., Lozzi, L., Savino, S., Scarselli, M., Rappuoli, R. & Pizza, M., 2005, *Infect. Immun.*, 73, 1151-1160.

41. Beernink PT, Welsch JA, Bar-Lev M, Koeberling O, Comanducci M, Granoff DM, 2008*, Infect Immun*., 76(9), 4232-40.

Chapter 2

Methodological Aspects

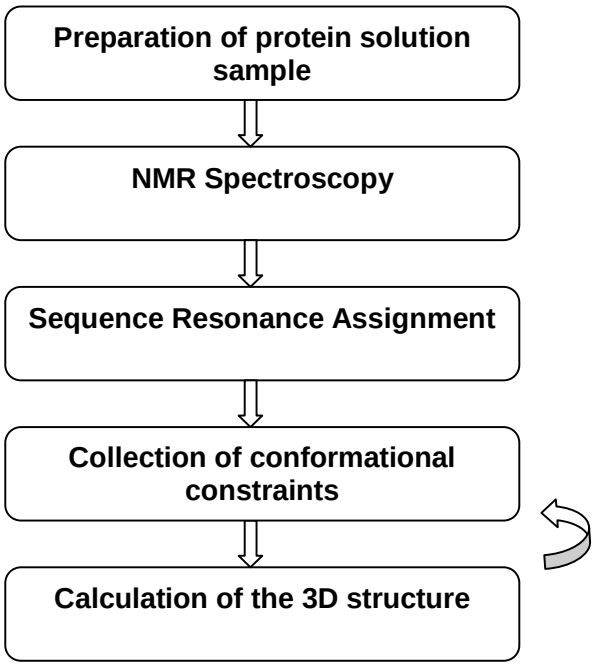
2.1 NMR and high-resolution structure determination

X-ray crystallography and NMR spectroscopy are the two main techniques that can provide structures of macromolecules at atomic resolution. Both techniques play a key role in structural biology for understanding of molecular functions and mechanisms which are involved in several physiological processes.

Whereas X-ray crystallography requires single crystals, NMR measurements can be carried out in solution with possibility to adjust conditions, such as temperature, pH and salt concentration, to be as close as possible to the physiological fluids. Moreover NMR measurements not only provide structural data but also can supply information on the internal mobility of proteins on various time scales, on protein folding and on intra-, as well as, intermolecular interactions.

In my thesis the latter studies are also supported by molecular docking programs, such as HADDOCK (*H*igh *A*mbiguity *D*riven protein-protein *DOCK*ing).

The principles of NMR structure determination can be summarized with the following steps:



The standard protocol includes the preparation of a homogeneous sample of the protein solution, the recording and handling of the NMR datasets, and the structural interpretation of the NMR data.

2.1.1 Sample preparation

The first step to solve the three dimensional structure of biological macromolecules is the preparation of the sample, since a highly purified protein preparation is required. In fact, an inhomogeneous preparation and/or aggregation of the protein may severely compromise the structure determination. The first step in every protein NMR study therefore involves optimization of the measurement conditions as pH, ionic strength, and temperature that can often be adjusted to mimic physiological conditions. The macromolecule under study should be stable in the chosen conditions for many weeks.

Proteins with a molecular weight larger than 10 kDa must be isotope enriched in ^{15}N and ^{13}C for an efficient structure determination. Triple labelled sample, ^{2}H, ^{13}C and ^{15}N, is necessary for protein with larger size (more than 30 kDa).

2.1.2 NMR spectroscopy

It is not possible to perform the data interpretation on 1D NMR, since biological macromolecules contain thousands of resonance lines.

Moreover, the interpretation of NMR data requires correlations between different nuclei, which are implicitly contained in 1D spectra but often difficult to extract. Therefore multidimensional NMR spectra are required for proteins structural determination. The crucial step in increasing the dimensionality of NMR experiments lies in the extension from one to two dimensions. A higher dimensional NMR experiment consists of a combination of two-dimensional (2D) experiments and extension from a 2D to a n-dimensional (nD) NMR experiment consists in the combination of (n-1) two-dimensional experiments which contains only one excitation and one detection period but repeats the

evolution and mixing times (n-1) times. A typical nD NMR experiment thus follows the scheme:

excitation - (evolution - mixing)$_{n-1}$ – detection

where the bracket repeats (n–1) times. Only during the detection period the signal is physically measured and this period is often referred to as the direct dimension in contrast to the evolution periods which are referred to as indirect dimensions.

The NMR multidimensional measurements almost always use protons (^{1}H) and, depending on the isotope labelling, ^{13}C and/or ^{15}N nuclei. A 3D spectrum can for example be obtained by correlating the amide groups with the α-carbon nuclei attached to ^{15}N. The chemical shifts of these carbon nuclei are used to spread the resonances from the 2D plane into a third dimension. These 3D triple resonance NMR experiments are used to correlate backbone resonances to corresponding residues in the primary sequence of a protein.

The proton offers the best sensitivity and for this reason constitutes the preferred nucleus for detection of the NMR signal. The other nuclei are usually measured during evolution periods of multidimensional NMR experiments and their information is transferred to protons for detection.

Beside ^{1}H NMR, ^{13}C direct detection technique provides a valuable alternative to overcome fast relaxation (1). Due to the smaller magnetic moment of the ^{13}C nucleus, the transverse relaxation rates on ^{13}C spins are much slower than ^{1}H. By exciting ^{13}C spins directly, the signal loss due to the fast relaxation of ^{1}H during the magnetization transfer can be reduced. 2D and 3D pulse sequence for backbone assignment of ^{13}C direct detection and ^{13}C -^{13}C NOESY type experiments are available (2).

For small proteins (less than 10k Da), ^{13}C or ^{15}N labelling is not required. In this case the assignment strategy makes use of a combination of 2D homonuclear ^{1}H NMR experiments such as TOCSY, and NOESY spectra.
TOCSY-type experiments, where TOCSY stands for *To*tal *C*orrelation *S*pectrosco*py*, correlate different nuclei *via J* coupling (3, 4). In proteins which are isotope labelled with ^{15}N and ^{13}C *J* couplings between ^{1}H, ^{15}N and ^{13}C allow *through-bond* correlations across the peptide bond.

Differently, geometric information, required to determine the structure of a protein, are obtained by *through-space* correlation experiments via nuclear Overhauser effect (NOE), the strength of which relies mainly on dipolar coupling (5). NOEs connect pairs of hydrogen atoms separated by less than about 5 Å.
Moreover, in the NOESY-experiments, where NOESY stands for *NOE S*pectroscop*Y* (6), the nuclei involved in the NOE correlation can belong to amino acid residues that may be far apart along the protein sequence but close in space. For molecules with a molecular weight of more than 5 kDa the intensity of an NOE is approximately proportional to r^{-6} and to the molecular weight, where *r* is the distance between the two interacting spins.

2.1.3 Backbone and side chains Resonance Assignment

For a good analysis of NMR spectra and to solve the 3D structure, nearly complete assignments of signals in the spectra to individual atoms in the molecule are required.

The application of multidimensional NMR spectroscopy allowed the development of general strategies for the assignment of signals in proteins. All procedures use the known protein sequence to connect nuclei of amino acid residues which are neighbours in the sequence. As mentioned before, for unlabelled proteins smaller than 10 kDa the combination of the [^{1}H,^{1}H]-TOCSY, used for the sequential assignment, with the [^{1}H, ^{1}H]-NOESY spectrum allows the assignment of most proton NMR signals to individual protons (7).

For larger proteins extensive signal overlap prevents complete assignments of all ^{1}H signals in proton spectra. This problem can be overcome with 3D NMR technique and uniformly ^{13}C and ^{15}N labelled proteins. The resonance assignment of single (^{15}N or ^{13}C) labelled proteins using 3D experiments is basically an extension of Wüthrich's strategy which exclusively relies on homonuclear ^{1}H NMR experiments. With these methods, systems with molecular weights up to approximately 30 kDa can be studied.

In ^{13}C, ^{15}N-labelled proteins a sequential assignment strategy is based on through-bond correlations across the peptide-bond between sequential amino acids. This procedure circumvents the use of NOESY spectra already in the assignment step. Most of these correlation experiments use the three types of nuclei ^{1}H, ^{15}N, ^{13}C and are referred to as triple resonance experiments.

The 3D triple resonance experiments exclusively correlate the resonances of the peptide backbone (HN(i), N(i), Cα(i), Hα(i), Cα(i-1), CO(i) and CO(i-1)). Figure 1 shows the spin system of the peptide backbone and indicates the size of the coupling constants used for magnetization transfer in double ^{13}C-, ^{15}N-labelled proteins.

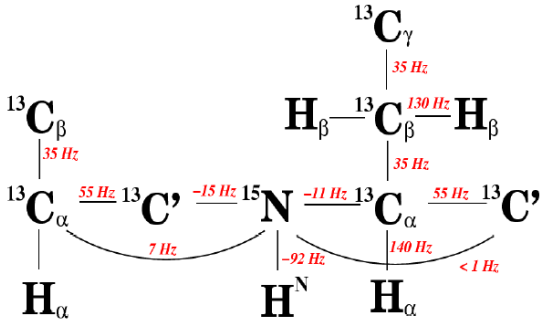


Figure 1. Spin system of the peptide backbone and the size of the ^{1}J and ^{2}J coupling constants that are used for magnetization transfer in ^{13}C-, ^{15}N-labelled proteins.

The 3D experiments used to identify the backbone resonances are, usually, HNCA or HNCACB, HN(CO)CA or HN(CO)CACB, HNCO, HN(CA)CO and HNHA (8). The HNCACB for example, correlates each H-^{15}N group with both the intra- and the neighbouring inter-residue Cα and and Cβ. These four types of connectivities are discriminated using the HN(CO)CACB experiment, in which only the inter-residue HN-Cα and Cβ couplings are observed.

Similar strategy can be used to assign the other resonances in the other triple resonance spectra.

In the case of proteins with a molecular weight larger than 30 kDa, the use of TROSY-type experiments is necessary (9). TROSY experiments can reduce the signal loss, which is the direct consequence of the slower correlation tumbling of large molecules that results in faster relaxation and consequently broader lines in the NMR spectrum.

TROSY uses constructive interference between different relaxation mechanisms and works best at the highest available magnetic field strengths in the range of 700 to 900 MHz proton resonance frequencies. With TROSY the molecular size of proteins accessible for detailed NMR investigations has been extended several fold. The TROSY technique benefits a variety of triple resonance NMR experiments as the 3D HNCA and HNCOCA (10) and the TROSY-based NOESY experiments (11) for the collection of structural constraints are also available.

CRINEPT-TROSY is another recently developed technique for even larger proteins. CRINEPT-TROSY can yield a further significant gain in sensitivity for molecular sizes above 200kDa (12, 13). The best magnetic field strengths for CRINEPT-TROSY are in the range 900-1000MHz proton resonance frequencies.

Since the Hα and Cα/β chemical shifts have been assigned, 3D H(C)CH-TOCSY and (H)CCH-TOCSY (14) experiments are then used to link the side chain spin systems to the backbone assignments. These two experiments

provide information for the assignment of the side chain protons and of the side chain carbons, respectively.

A complete set of backbone chemical shifts for all Hα, Cα, Cβ and CO resonances can be used to predict the secondary structure of the protein (15). One technique in particular, the Chemical Shift Index (CSI) (16), has been widely used for the quantitative identification and location of secondary structure in proteins.

The method relies on the fact that the chemical shifts of the different nuclei in the protein backbone are related both to the type of aminoacid and to the nature of the secondary structure they are located in. By comparing the actual chemical shift for a nucleus in a specific amino acid with a reference value, it is possible to predict in what secondary structure element the nucleus resides.

Another system for the secondary structure prediction is TALOS (Torsion Angle Likelihood Obtained from Shift and sequence similarity) which allowed to predict phi and psi backbone torsion angles using a combination of five kinds (HA, CA, CB, CO, N) of chemical shift assignments for a given protein sequence.

The method relies on the fact that the phi and psi of the different nuclei in the protein backbone are related both to the type of amino acid and to the nature of the secondary structure they are located in.

2.1.4 Collection of conformational constraints

The NMR data analysis permits to collect geometric conformational information in the form of distances and/or torsion angles, which are subsequently used for structure calculation. The latter can be performed combining information about the covalent structure of the protein, such as the amino acid sequence, bond lengths, bond angles, chiralities, and planar groups, as well as by steric repulsion between non-bonded atom pairs together with the experimental data.

Although a variety of NMR parameters contain structural information, the crucial information comes from NOE measurements which provide distance information between pairs of protons.

Supplementary constraints can be derived from through bond correlations in the form of dihedral angles (17). Further, CSI and TALOS too data provides, as mentioned before, information on the type of secondary structure. Such information can be included in a structure calculation by restricting the local conformation of a residue to the α̃-helical or β-sheet region of the Ramachandran plot through torsion angle restraints.
Moreover, hydrogen bonds can be experimentally detected via through-bond interactions (18) and they can be useful during structure calculations of larger proteins when not enough NOE data are available yet.

Another class of conformational restraints that can be introduced is constituted by residual dipolar couplings (RDCs) in the presence of an external orienting medium; RDCs give information on angles between covalent bonds and globally defined axes in the molecule, namely those of the magnetic susceptibility tensor (19, 20).

The intensity of a NOE, i.e. the volume *V* of the corresponding cross peak in a NOESY spectrum (21, 22), is related to the distance *r* between the two interacting spins by

$$V=\langle r^{-6}\rangle f(\tau_c) \qquad (1)$$

where averaging ***r*** indicates that in molecules with inherent flexibility the distance r may vary and thus has to be averaged appropriately. This volume is also dependent on the function ***f***(τ_c) that includes effects of global and internal motions of the molecule.

NOEs are usually treated as upper interatomic distance rather than as precise distance restraints. Adding the information on the distance between pairs of

protons and their position in the polypeptide sequence allows construction of possible arrangements in which all distance constraints are fulfilled.
Vicinal scalar coupling constants, $^3\boldsymbol{J}$, can also be used for torsion angle prediction, according to Karplus relations (23):

$$^3J(\theta) = A\cos 2\theta + B\cos\theta + C \qquad (2)$$

where θ is the torsion angle; ***A***, ***B*** and ***C*** are Karplus constants depending on the type of the torsion angle. $^3\boldsymbol{J}$ HNHa coupling constants are obtained from the ratio between the intensity of the diagonal peak and that of the cross-peak of the HNHA map.

Through the analysis of the HNHB spectrum the $^3J_{HNH\beta}$ coupling constants (χ torsion angles) can be derived. The process of defining useful conformational constraints on the basis of J coupling constants usually requires the translation of these quantities into molecular-geometry related dihedral-angle values.

Using NMR constrains, calculation programs fold a random generated 3D structure, in order to maximize the agreement between the structure and the structural constrains. A NMR structure is represented by a family of conformers which are in good agreement with the structural constrains imposed. The precision of the structure is measured by the *root-mean-square-deviation* (RMSD) of the coordinates of the protein atoms for each conformer of the family from the mean structure and the accuracy of the structure is measured by an average *target function* of the family; the latter measures the agreement between the structure and the given set of constrains.

2.1.5 Structure calculation, refinement and validation

Starting from chemical shifts and NOESY spectra all the calculation works have been done with the programs ATNOS-CANDID coupled with CYANA-2.0

by using torsion angle dynamics algorithm. In this algorithm, the molecular dynamics simulation uses torsion angles as degree of freedom, while bond lengths, angles, and backbone peptide plane angles are fixed.
In particular, ATNOS/CANDID program, based on the backbone and side-chains assignment, allow to obtain the automated NOESY peak picking (ATNOS) and the automated NOE assignment (CANDID).

During the integration time steps, the best conformation was searched by minimized the deviation between the constrains and obtained conformations. Compared with other algorithms, torsion angle dynamics provides at present the most efficient way to calculate NMR structures.

At the end of the structure calculation, an important question is if the structure calculation was “successful”. There are two important features, i) whether the calculated structures fulfill the given restraints, ii) whether the calculations are converged. The first feature could be described by statistical report of the target function. In each calculation, 20 or 30 conformers, among 200 or 350 conformers, with lowest final target functions were chosen to represent the solution structures. The average values of target function of the represent structures should be smaller than 1$Å^2$ and each single violation should be smaller than 0.3 $Å^2$. The second feature was judged by RMSD between these models. Only the family structure with backbone RMSD values close or less than 1 Å will be considered as “good”.
The “good” structures obtained from torsion angle dynamics finally will be refined in AMBER program (24). In the final refinement, force field parameters are also considered.
The refinement in explicit water box after the refinement *in vacuum* make the structure qualities further improved.

The solved structures could be flawed or present a wrong fold, even if RMSD and target function values are accepted. The defect could be induced by only a few wrong assigned NOEs or other local constraints. The validation package provides a tool to understand the quality of the structures. In this thesis, programs PROCHEK_NMR, WHATIF and QUEEN are used for judging the structure quality. The solution structures will be accepted if more than 90% residues fall into the allowed region of Ramachandran plot and less than 1% residues in the disallowed region.

PROCHEK_NMR provides a detailed check on the stereochemistry of a protein structure, while WHATIF (25) try to assess the quality of a structure primarily by checking whether a number of different parameters are in agreement with their values in databases derived from high-resolution X-ray structures.

QUEEN (*QU*antitative *E*valuation of *E*xperimental *N*mr restraints) is a method which describes the quantitative evaluation of experimental NMR restraints. This method successfully identifies the crucial restraints in a structure determination: those restraints that are both important and unique.

2.2 Mapping Protein-Protein Interactions in Solution by NMR Spectroscopy

A large majority of the functional processes in the cell occur through weak transient interactions among proteins (interactome). Unfortunately, the characterization at the atomic level of these weak interactions is difficult, due to their dynamic nature. NMR is a very well suited technique as it allows us to characterize also dynamical and kinetic properties. The available NMR methods to this end are reviewed and illustrated with applications from the recent biochemical literature: intermolecular NOEs, cross-saturation, chemical shift perturbation, dynamics and exchange perturbation, paramagnetic methods and dipolar orientation.

Most of these methods are now routinely applied for complexes with total molecular mass of 60 kDa and can likely be applied to systems up to 1000 kDa.

The study of protein-protein interactions by NMR appears to be one of the most important application; this evolution became possible because of better NMR equipment, development of more efficient isotopic labelling schemes, especially perdeuteration (26) and the advent of TROSY (9) and CRINEPT-TROSY (12, 13)

2.2.1 Chemical Shift Perturbation Mapping

Chemical shift perturbation is the most widely used NMR method to map protein-protein interfaces. The ^{15}N-^{1}H HSQC spectrum of protein “A” is monitored when the unlabelled interaction partner (protein “B”) is titrated in, and the perturbations of the chemical shifts are recorded. The approach requires that sequence-specific assignments have been determined for protein A at least in its free state.

The interaction causes environmental changes on the protein interfaces and, hence, affect the chemical shifts of the nuclei in this area. In some cases the entire protein may change conformation, and all chemical shifts may be affected; then the chemical shift perturbation fails as a mapping device but is an excellent indicator of allosteric processes (27, 28).

Shift perturbation measurements just yield the locations of the interfaces on the individual binding partners, but not how the partners interact at atomic level. One approach may be to combine the NMR shift mapping experiments with computational docking programs (29).

2.2.2 Titration with NMR

NMR titration allows, in addition to the mapping of the interface, a good estimation of the affinity, stoichiometry, and specificity of binding as well as the

kinetics of binding. How the chemical shifts of the labelled protein change during the titration is determined by the kinetics of the interaction.

If the complex dissociation is very fast, there is, even during the titration, only a single set of resonances whose chemical shifts are the fractionally weighted average of the free and bound chemical shifts. Here the resonances of nuclei at the interface move in a continuous way during the titration. This regime is referred to as fast chemical exchange and is often observed for weaker interactions.

If the complex dissociation is very slow, one observes one set of resonances for the free protein and one set for the bound protein. During the titration, the “free set” will disappear and will be replaced by the bound set. This regime is referred to as slow chemical exchange. In slow exchange one does not automatically know to which new location the resonance has moved, unless one carries out an

independent assignment procedure for the bound state.

In the intermediate chemical exchange case, the frequencies of the changing resonances become poorly defined, and extensive kinetic broadening sets in (30). A rule of thumb is that interactions with Kd <10 μM are slow exchange and intermediate/fast exchange otherwise.

2.3 Internal mobility in proteins

The study of the protein mobility is particular important to understand biological and chemical phenomena for the recognized importance of protein internal motions in biological processes. Macromolecular functions are often associated to energetic transition, which are intimately connected with structural changes and molecular flexibility.

The measurement of ^{15}N relaxation rates in isotopically enriched proteins is particularly useful for obtaining dynamics information since the relaxation of this nucleus is governed predominantly by the dipolar interaction

with directly bound proton and by the chemical shift anisotropy (CSA) mechanism.

In an ^{15}N relaxation experiment, one creates non-equilibrium spin order and records how this relaxes back to equilibrium. At equilibrium, the ^{15}N magnetization is aligned along the external field, and this alignment can be changed by radio frequency pulses. The magnetization will relax back to equilibrium along the direction of the magnetic filed with a constant time called longitudinal relaxation time T_1. When outside equilibrium the magnetization can have also a component perpendicular to the external magnetic field. The time constant for this spin component to return to equilibrium is called transverse relaxation time, T_2. A third source of relaxation parameter is the heteronuclear NOE. This is measured by saturating the proton (^1H) signal and observing changes in the ^{15}N signal intensities.

The relaxation parameters are related to the spectral density function, of the ^1H-^{15}N bond vector by the following equations (31, 32):

$$T_1^{-1} = R_1 = (d^2/4)\left[J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H + \omega_N)\right] + c^2 J(\omega_N) \quad (3)$$

$$T_2^{-1} = R_2 = (d^2/8)\left[4J(0) + J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H) + 6J(\omega_H + \omega_N)\right] + (c^2/6)\left[4J(0) + 3J(\omega_N)\right] + R_{ex} \quad (4)$$

$$NOE = 1 + (d^2/4R_1)(\gamma_H/\gamma_N)\left[6J(\omega_H + \omega_N) - J(\omega_H - \omega_N)\right] \quad (5)$$

in which $d=(\mu_0 h\gamma_N\gamma_H/8\pi r^3_{NH})$ and $c=\omega_N\Delta_N/\sqrt{3}$. R_{ex} is a term introduced to account for microsecond to millisecond conformational exchange contributions to R_2.

The dynamic information contained in the relaxation rates are represented by the values of the spectral density function, $J(\omega)$, at several frequencies. The descriptions of the dynamics requires therefore the identification of a suitable model for the spectral density function, which must

be consistent with the experimental relaxation rates. The spectral density is often expressed in terms of global tumbling parameters, and of local motional parameters.

2.3.1 Data Analysis

2.3.1a Spectra Density Mapping

The spectral density mapping approach was developed by Peng and Wagner in 1992 (33). It makes use of six different relaxation parameters, which are used to map the spectral density function. These parameters are: 1) the longitudinal ^{15}N relaxation rate, R_1, 2) the transverse ^{15}N relaxation rate, R_2, 3) the $^{1}H \rightarrow ^{15}N$ heteronuclear NOE, 4) the relaxation rate of longitudinal two-spin order $R_{NH}(2H_zN_z)$, 5) the relaxation rate of anti-phase ^{15}N coherence $R_{NH}(2H_zN_{x,y})$ and 6) longitudinal relaxation of the amide protons.

The method of spectral density mapping does not require any assumption regarding the form of the spectral density functions. Values of the spectral density functions of ^{1}H-^{15}N vectors are directly sampled at several relevant frequencies (e.g. 0, ω_H-ω_N, ω_H+ω_N, ω_H, ω_N).

Drawbacks in this method reside in the fact that the three measurable relaxation parameters, R_1, R_2 and heteronuclear NOE, are insufficient to determine uniquely the values of the spectral density function at the five frequencies in equations (1) to (3) and that anomalous behaviour can be expected for the spectral densities at the three highest frequencies (33). These problems can be overcome by a more recent approach called reduced spectral density mapping, in which the values of the spectral density function at ω_H-ω_N, ω_H+ω_N and ω_H frequencies can be combined in a average spectral density $<J(\omega_H)>$ (24, 25). A comparison of $J(\omega)$ measured at high and low frequencies provides a quantitative measure of the breadth of the frequency distribution accessed through the spatial fluctuations of the bond and the overall tumbling of the molecule.

2.2.1b Model Free Approach

A different method, called Model Free approach, has been introduced by Lipari and Szabo (34). This model assumes that the overall rotation of the molecule can be described by a single correlation time (isotropic motion) and that this overall motion and the internal motions are independent. Then the total correlation function can be factored as:

$$C(t)= C_O(t)\times C_I(t) \qquad (6)$$

Where C(t) is the total correlation function, $C_O(t)$ is the correlation function characterizing the overall rotation and $C_I(t)$ is the correlation function characterizing the internal motions.

As a consequence, the global spectral density function can be expressed as a weighted sum of Lorentzian functions. This is correct rigorously for isotropic rotational diffusion and approximately for anisotropic rotational diffusion.

The Lipari and Szabo formalism employs a minimum number of parameters to describe the overall isotropic tumbling motion of a macromolecule and the internal motions of the ^{15}N-^{1}H bond vector. The central equation in the Model Free Approach is:

$$J(\omega) = \left[\frac{S^2 \tau_m}{1+(\omega\tau_m)^2} + \frac{(1-S^2)\tau'}{1+(\omega\tau')^2} \right] \qquad (7)$$

where τ_m is the correlation time as a result of the isotropic tumbling motion of the entire molecule. The effective correlation time resulting from internal motions is described by τ_e, where $\tau^{-1} = \tau_m^{-1} + \tau_e^{-1}$. The order parameter S^2 describes the degree of spatial restriction of the internal motion of the ^{1}H-^{15}N bond vector. It satisfies the inequality $0 \leq S \geq 1$ and lower values indicate larger amplitudes of internal motions. As a consequence, for a nucleus rotating

as a whole with the molecule, all contributions to the spectral density function derive from the overall tumbling; alternatively, extra contributions will be described by other motions with correlation times faster than the overall tumbling.

An extended form of the model-free spectral density function has been developed by Clore and coworkers (35) to describe internal motions that take place on two distinct time scales, differing by at least an order of magnitude.

2.2.3 The contribution to relaxation of exchange processes

The presence of exchange processes occurring in the micro-millisecond time scale produces dephasing of magnetization and contributes to make the transverse relaxation time shorter.

A method to obtain a detailed analysis of exchange contribution is the measurement of R_2 rates as a function of the refocusing times $\tau_{CPMG}(=1/(2\nu_{CPMG})$, where ν_{CPMG} is the frequency of repetition of 180° pulses during the Carr-Purcell-Meiboom-Gill (CPMG) sequence (36).

The contribution of the exchange processes (R_{ex}) to the transverse relaxation rate can be expressed as follow (37):

$$R_{ex} = \frac{k_{ex}}{2} - 2\nu_{eff}\sinh^{-1}\left(\frac{k_{ex}}{\xi}\sinh\frac{\xi}{4\nu_{eff}}\right) \qquad (6)$$

where $\xi = (k_{ex}^2 - 4p_A p_B \delta\omega^2)^{1/2}$, $k_{ex} = 1/\tau_{ex}$ and $\nu_{eff}(s^{-1}) \doteq \frac{1}{2(T_\pi + \tau_{CPMG})}$

p_A and p_B are the populations of the sites A and B in a two-site exchange process, $\delta\omega$ is the difference of Larmor frequencies between the sites and τ_{ex} is the time constant for the exchange process.

2.4 Molecular Docking for the study of protein-protein interaction

Molecular docking have been investigated through the program HADDOCK 2.0 (38).

In the latter the docking process is driven by ambiguous interaction restraints (AIRs), which are ambiguous distances between all solvent exposed residues involved in the interaction. HADDOCK protocol defines active and passive residues.

The active residues are all residues showing a significant chemical shift perturbation after the formation of the complex, with a solvent accessibility usually > 50%; the passive residues are the residues which are surface neighbors of the active and have a solvent accessibility usually > 50%.

Solvent accessibility has been calculated with the program NACCESS (39).

In this work of thesis I considered a solvent accessibility > 40%.

2.4.1 Docking Protocol

The docking protocol consists of three consecutive steps:

a) rigid body minimization driven by interaction restraints (it0), in which 1000 structures of the complex have been generated;

b) the best 200 structures in terms of total intermolecular energy were further submitted to the semiflexible simulated annealing in which side-chains and backbone atoms of the interface residues are allowed to move (it1);

b) final refinement in Cartesian space in a explicit solvent (water).

Finally the 200 structures obtained are clustered using a threshold of 7.5 Å of RMSD among the structure of the cluster.

Clusters containing 5 structures were considered and I performed a detailed analysis on the first two clusters which showed a lower HADDOCK score value (HADDOCK score: 1Evdw+0.2Eelc+0.1Edist+1Esolv, where the four terms are respectively: Van der Walls energy, electrostatic energy, distance restrains energy and desolvation energy).

References

1. Bermel, W.; Bertini, I.; Felli, I.C.; Kümmerle, R.; Pierattelli, R., 2003, *J. Am. Chem. Soc.* 125, 16423-9.

2.Bermel, W.; Bertini, I.; Duma, L.; Felli, I.C.; Emsley, L.; Pierattelli, R., 2005, *Angew Chem Int Ed Engl.,* 44, 3089-92.

3. Aue, W.P., Bartholdi, E. & Ernst, R.R., 1976, *J.Chem.Phys.,* 64, 2229-2235.

4. Wider, G., Macura, S., Kumar, A., Ernst, R.R. & Wüthrich, K., 1984, *J.Magn.Reson.,* 56, 207-234.

5. Wider G., 1998, *Progress NMR Spectrosc,* 32, 193-275.

6. Kumar, A., Ernst, R.R. & Wüthrich, K., 1980, *Biochem.Biophys.Res.Commun.* 95, 1104.

7. Wüthrich, K., 1986, *NMR of Proteins and Nucleic Acids*, Wiley: New York .

8. Kay, L.E., Ikura, M., Tschudin, R. & Bax, A., 1990, *J.Magn.Reson.*, 89, 496-514.

9. Pervushin, K., 2000, *Q.Rev.Biophys.*, 33, 161-197.

10. Salzmann, M.; Pervushin, K.; Wider, G.; Senn, H.; Wüthrich, K., 1999, *J Biomol NMR.* 14, 85-8.

11. Pervushin, K.V.; Wider, G.; Riek, R.; Wüthrich, K.,1999, *Proc. Natl. Acad. Sci.*, 96, 9607-12.

12. Riek, R.; Wider, G.; Pervushin, K.; Wüthrich, K., 1999, *Proc. Natl. Acad. Sci.,* 96, 4918-23.

13. Rudiger, S.; Freund, SM.; Veprintsev, D.B.; Fersht, A.R., 2002, *Proc. Natl. Acad. Sci.,* 99 11085-90.

14. Kay, L.E., Xu, G.Y., Singer, A.U., Muhandiram, D.R. & Forman-Kay, J.D., 1993, *J.Magn.Reson.Ser.B,* 101, 333-337.

15. Wishart, D.S., Sykes, B.D. & Richards, F.M., 1991, *J.Mol.Biol.*, 222, 311-333.

16. Wishart, D.S., Sykes, B.D. & Richards, F.M., 1992, *Biochemistry*, 31, 1647-1651.

17. Wüthrich, K.,1986, *NMR of Proteins and Nucleic Acids*, Wiley: New York.

18. Cordier, F. & Grzesiek, S., 1999, *J.Am.Chem.Soc.,* 121, 1601-1602.

19. Tolman, J.R., Flanagan, J.M., Kennedy, M.A. & Prestegard, J.H., 1995, *Proc.Natl.Acad.Sci., USA* 92, 9279-9283.

20. Tjandra, N., Grzesiek, S. & Bax, A., 1996, *J.Am.Chem.Soc.,* 118, 6264-6272.

21. Jeener, J., Meier, B.H., Bachmann, P. & Ernst, R.R., 1979, *J.Chem.Phys.*, 71, 4546-4553.

22. Macura, S. & Ernst, R.R., 1980, *Mol.Phys.,* 41, 95-95.

23. Karplus, M. (1963) J.Am.Chem.Soc. 85, 2870-2871

24. Case, D.A.; Darden, T.A.; Cheatham, T.E.; Simmerling, C.L.; Wang, J.; Duke, R.E.; Luo, R.; Merz, K.M.; Wang, B.; Pearlman, D.A.; et al., 2004, *AMBER 8 (computer program)*. University of California, San Francisco.

25. Vriend G., 1990, *J Mol Graph*., 8(1):52-6, 29.

26. Venters, R. A., Farmer, B. T., II, Fierke, C. A., and Spicer, L. D., 1996, J. Mol. Biol. 264, 1101-1116

27. van Nuland, N. A., Kroon, G. J., Dijkstra, K., Wolters, G. K., Scheek, R. M., and Robillard, G. T., 1993, *FEBS Lett*., 315, 11-15.

28. Chen, Y., Reizer, J., Saier, M. H., Jr., Fairbrother, W. J., and Wright, P. E., 1993, *Biochemistry*, 32, 32-37.

29. Morelli, X., Dolla, A., Czjzek, M., Palma, P. N., Blasco, F., Krippahl, L., Moura, J. J., and Guerlesquin, F., 2000, *Biochemistry*, 39, 2530-2537.

30. Zuiderweg, E. R. P., Hamers, L. F., Rollema, H. S., De Bruin, S. H., and Hilbers, C. W., 1981, *Eur. J. Biochem.,* 118, 95- 104.

31. Abragam, A., 1961, *The Principles of Nuclear Magnetism*, Oxford University.

32. Wagner, G., 1993, *Curr.Opin.Struct.Biol.,* 3, 748-754.

33. Peng, J.W. & Wagner, G. ,1992, *J.Magn.Reson.*, 98, 308-332.

34. Lipari, G. & Szabo, A., 1982, *J.Am.Chem.Soc.,* 104, 4546-4559.

35. Clore, G.M., Szabo, A., Bax, A., Kay, L.E., Driscoll, P.C. & Gronenborn, A.M., 1990, *J.Am.Chem.Soc.,* 112, 4989-4991.

36. Mulder, F.A., Van Tilborg, P.J., Kaptein, R. & Boelens, R., 1999, *J.Biomol.NMR,* 13, 275-288.

37. Palmer, A.G., III, Williams, J. & McDermott, A., 1996, *J.Phys.Chem.*, 100, 13293-13310.

38. Dominguez, C., Boelens, R. & Bonvin, A. M., 2003, *J. Am. Chem. Soc*., 125, 1731–1733.

39. Hubbard, S. J. & Thornton, J. M., 1993, NACCESS. Department of Biochemistry and Molecular Biology, University College London, London, UK.

Chapter 3

Epitope Mapping of a Bactericidal Monoclonal Antibody against the Factor H Binding Protein of *Neisseria meningitidis*

Maria Scarselli[1], Francesca Cantini[2], Laura Santini[1], Daniele Veggi[1], Sara Dragonetti[2], Claudio Donati[1], Silvana Savino[1], Marzia M. Giuliani[1], Maurizio Comanducci[1], Federica Di Marcello[1], Giacomo Romagnoli[1], Mariagrazia Pizza[1], Lucia Banci[2] and Rino Rappuoli[1]

[1] Novartis Vaccines and Diagnostics, Via Fiorentina 1, 53100 Siena, Italy.

[2] Magnetic Resonance Center (CERM), University of Florence, Via L. Sacconi 6. 50019 Sesto Fiorentino, Italy

J Mol Biol., (2009)

doi:10.1016/j.jmb.2008.12.005 J. Mol. Biol. (2009) 386, 97–108

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Epitope Mapping of a Bactericidal Monoclonal Antibody against the Factor H Binding Protein of *Neisseria meningitidis*

Maria Scarselli[1], Francesca Cantini[2], Laura Santini[1], Daniele Veggi[1], Sara Dragonetti[2], Claudio Donati[1], Silvana Savino[1], Marzia M. Giuliani[1], Maurizio Comanducci[1], Federica Di Marcello[1], Giacomo Romagnoli[1], Mariagrazia Pizza[1], Lucia Banci[2]* and Rino Rappuoli[1]*

[1]*Novartis Vaccines and Diagnostics, Via Fiorentina 1, 53100 Siena, Italy*

[2]*Magnetic Resonance Center (CERM), University of Florence, Via L. Sacconi 6, 50019 Sesto Fiorentino, Italy*

Received 27 August 2008; received in revised form 11 November 2008; accepted 2 December 2008
Available online
11 December 2008

Edited by I. Wilson

The factor H binding protein (fHbp) is a 27-kDa membrane-anchored lipoprotein of *Neisseria meningitidis* that allows the survival of the bacterium in human plasma; it is also a major component of a universal vaccine against meningococcus B.

In this study, we used nuclear magnetic resonance spectroscopy, mutagenesis, and *in silico* modeling to map the epitope recognized by MAb502, a bactericidal monoclonal antibody elicited by fHbp. The data show that the antibody recognizes a conformational epitope within a well-defined area of the immunodominant C-terminal domain of the protein that is formed by two loops connecting different β-strands of a β-barrel and a short α-helix brought in spatial proximity by the protein folding. The identification of the protective epitopes of fHbp is an important factor for understanding the mechanism(s) of an effective immune response and provides valuable guidelines for designing variants of the protein able to induce broadly protective immunity.

Keywords: Neisseria; fHbp; antibody; epitope mapping; NMR

Introduction

Neisseria meningitidis causes severe, often fatal septicemia and meningitis. Polysaccharide-based vaccines are available in preventing diseases caused by pathogenic serogroups A, C, Y, and W-135 in older children and adults. The development of a safe and effective vaccine against serogroup B has been, on the contrary, hampered for a long time because of the poor immunogenicity and cross-reactivity of type B capsular polysaccharide with sialylated proteins of human tissues.

So far, outer membrane-based vaccines such as the New Zealand MeNZB[1] and the Cuban VA-MENGOC-BC® have been licensed against serogroup B meningococcus for local use.[2] Recently, a universal protein-based vaccine[3] containing five novel antigens discovered by reverse vaccinology, a process that involves whole bacterial genome screening, has been proposed.[4] One of these antigens is a membrane-anchored lipoprotein of 274 amino acids known as factor H binding protein (fHbp), also termed GNA1870 and lipoprotein 2086.[5,6] In addition to inducing bactericidal antibodies in mice, this protein is indeed able to bind human factor H, a negative regulator of the alternative complement pathway. It has been proposed that, in addition to triggering the classical complement pathway, which in turn results in bacterial killing, antibody binding to fHbp could enhance the protective power of the immune response by inhibiting the bacterium ability to down-regulate the alternative cascade.[7] All these observations increase the relevance of fHbp in

**Corresponding authors.* E-mail addresses: banci@cerm.unifi.it; rino.rappuoli@novartis.com.

Abbreviations used: fHbp, factor H binding protein; WB, Western blot; FACS, fluorescence-assisted cell sorting; HSQC, heteronuclear single-quantum coherence; CRINEPT, cross-correlated relaxation-enhanced polarization transfer; MAb, monoclonal antibody; FAb, antigen binding fragment; Fv, fragment variable; Fv_L, Fv light chain; Fv_H, Fv heavy chain; PBS, phosphate-buffered saline; BSA, bovine serum albumin; MD, molecular dynamics; AIR, ambiguous interaction restraint.

0022-2836/$ - see front matter © 2008 Published by Elsevier Ltd.

vaccine efficacy and stimulate interest in a detailed knowledge of its antigenic profile.

Sequence analysis on a panel of 71 clinical isolates of *N. meningitidis* showed that, while the 120 N-terminal amino acids of the protein are well conserved, the last 154 C-terminal residues show remarkable diversity. The sequence variability correlated with antigenic variation as polyclonal sera obtained immunizing with recombinant forms of the protein were protective *versus* strains carrying closely related variants but ineffective against those having more distant alleles.[5]

The nature and distribution of protective fHbp epitopes have been investigated in more detail for the protein encoded by the pathogenic strain MC58. It has been demonstrated that the 154 C-terminal residues of the protein contain most of the protective epitopes, with this domain being almost as good as the whole protein in inducing a protective immune response.[8] Moreover, the C-terminal portion of the protein contains residues targeted by bactericidal monoclonal antibodies, such as Arg204,[8] Glu146, Gly147, Gly148, and Arg149.[9]

The solution structure of such C-terminal immunodominant domain of fHbp ($fHbp_C$) has been determined by nuclear magnetic resonance (NMR) spectroscopy. This part of the protein forms an eight-stranded β-barrel whose strands are connected by loops of variable lengths. The barrel is preceded by a short α-helix and by a flexible N-terminal tail.[10]

Sequence alignments of fHbp alleles representative of the meningococcal population diversity showed that variable residues were not exclusively located in correspondence of loops, as expected for a domain embedded into the outer membrane, but were spread on the entire surface of the β-barrel. This profile of variability, together with the solvent exposure of hydrophilic residues and the observation that aromatic and hydrophobic residues are buried into the protein interior, suggested that the β-barrel is fully exposed on the bacterium surface and accessible to the immune system. The only exception could be represented by a hydrophobic patch present on one side of the β-barrel, which might represent a contact area to the outer membrane or to the N-terminal domain of the protein, in any case shielded to the immune system. The NMR structure also showed that amino acids previously identified as targets of protective monoclonal antibodies[8,9] clustered within a restricted area opposite to the hydrophobic patch.

In this study, we investigated the interaction between $fHbp_C$ and MAb502, a bactericidal monoclonal antibody elicited in mice by fHbp, using a combination of serological data, NMR measurements, and docking calculation. Residues forming the bactericidal epitope of $fHbp_C$ have been mapped by NMR titrations with the antigen binding fragment (FAb) of MAb502 (FAb502). The putative complex between $fHbp_C$ and the hypervariable fragment of FAb502 (Fv502) has been generated by molecular docking.

The data show that MAb502 recognizes a conformational epitope within a well-defined area of the immunodominant C-terminal domain of fHbp. The identification of the MAb502 epitope provides valuable information for the rational engineering of fHbp, aimed to create a chimeric protein presenting a wider collection of neutralizing epitopes to the host immune system.

Results

Western blot analysis

To test the binding of MAb502 to fHbp produced by natural isolates representative of the diverse meningococcal population, we selected 10 strains, including MC58, each of them carrying nonredundant fHbp sequences. The anti-MC58 fHbp polyclonal serum recognized the protein in Western blot (WB) on the cell extracts of all the strains tested, and fluorescence-assisted cell sorting (FACS) experiments confirmed the presence of the protein on the bacterial cell surface. As expected, the polyclonal serum elicited by the recombinant protein of MC58 was also able to induce complement-mediated killing of all strains. On the other hand, MAb502 was able to recognize fHbp in 6 of 10 strains in WB and FACS experiments and had a strong bactericidal activity only against MC58 (Table 1).

Hints on residues involved in the formation of the MAb502 epitope were searched by analyzing the sequence differences between the fHbp protein produced by MC58 strain and that of the other clinical isolates. Sequence alignment allowed classifying in two groups the amino acid positions variable between the $fHbp_C$ of MC58 and the other alleles. The first one contained amino acids varying between MC58 and WB-positive as well as WB-negative strains (shown in blue in Fig. 1). These amino acids in MC58 corresponded to Phe109, Ile114, His119, Glu146, Arg149, Gly163, Ala174,

Table 1. Summary of WB, FACS reactivity, and bactericidal titers (SBA) of MAb502 and anti-$fHbp_{MC58}$ antibodies against strains carrying different variant 1 fHbp alleles

	Polyclonal serum			MAb502		
Strain	WB	FACS	SBA	WB	FACS	SBA
MC58	+	+++	16,384	+	+++	256
M4030	+	++	2048	+	+	<4
M01-240149	+	+++	2048	+	+	<16
F6124	+	++	1024	+	+	<4
LPN17592	+	+	512	+	+	<4
M01-240345	+	++	512	+	+	<4
NZ98/254	+	++	64	−	−	<4
M2197	+	+++	512	−	−	<4
M6190	+	++	128	−	−	<4
M2937	+	++	1024	−	−	<16

SBA values are expressed as serum dilutions resulting in 50% decrease of bacterial CFUs per milliliter after a 60-min incubation of bacteria with the reaction mixture compared with the control CFUs per milliliter at time zero.

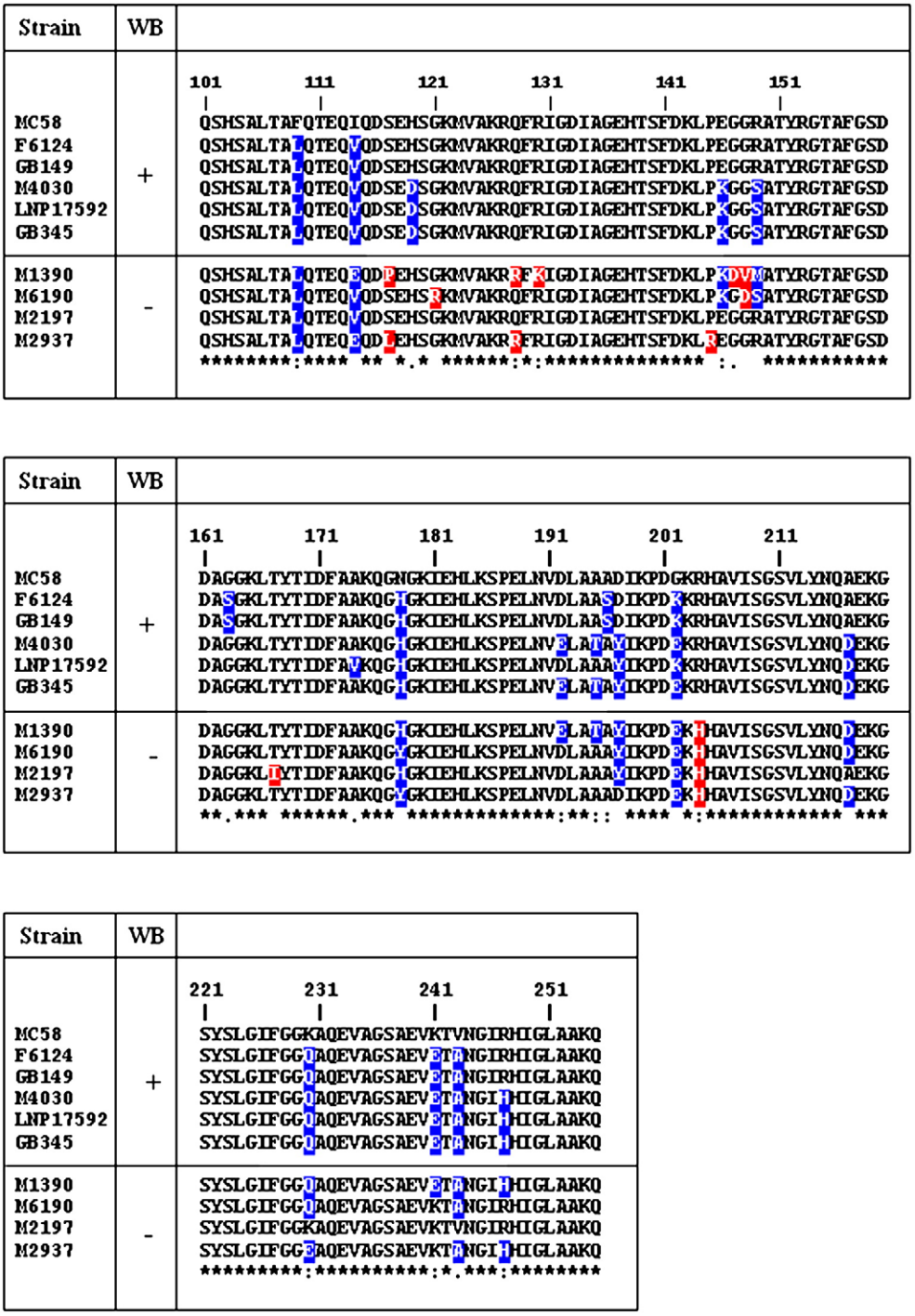


Fig. 1. Multiple sequence alignment of $fHbp_C$. Amino acids predicted relevant for MAb502 recognition are shown in red, while predicted dispensable residues are shown in blue. The second column (WB) subdivides strains recognized (+) or not (−) by MAb502 in WB.

Asn178, Asp192, Ala195, Ala196, Asp197, Gly202, Ala217, Lys230, Lys241, Val243, and Arg247. As the variability of these residues did not impact the WB recognition by MAb502, we classified their relative positions as dispensable.

The second group (red positions in Fig. 1) contained residues varying with respect to MC58 only in WB-negative strains. These amino acids are Ser117, Gly121, Gln128, Arg130, Pro145, Gly147, Gly148, Thr167, and Arg204. Among these, Ser117, Gly121, and Gln128 are located in the N-terminal flexible tail of $fHbp_C$, making it difficult to speculate on their role in epitope formation. With the exception of Thr167, all the other residues of this second group cluster on the same region on the β-barrel (Fig. 2a).

As already pointed out,[6] the Arg204→His substitution was the sole difference observed consistently in all the WB-negative strains. The relevance of this position for MAb502 binding was confirmed

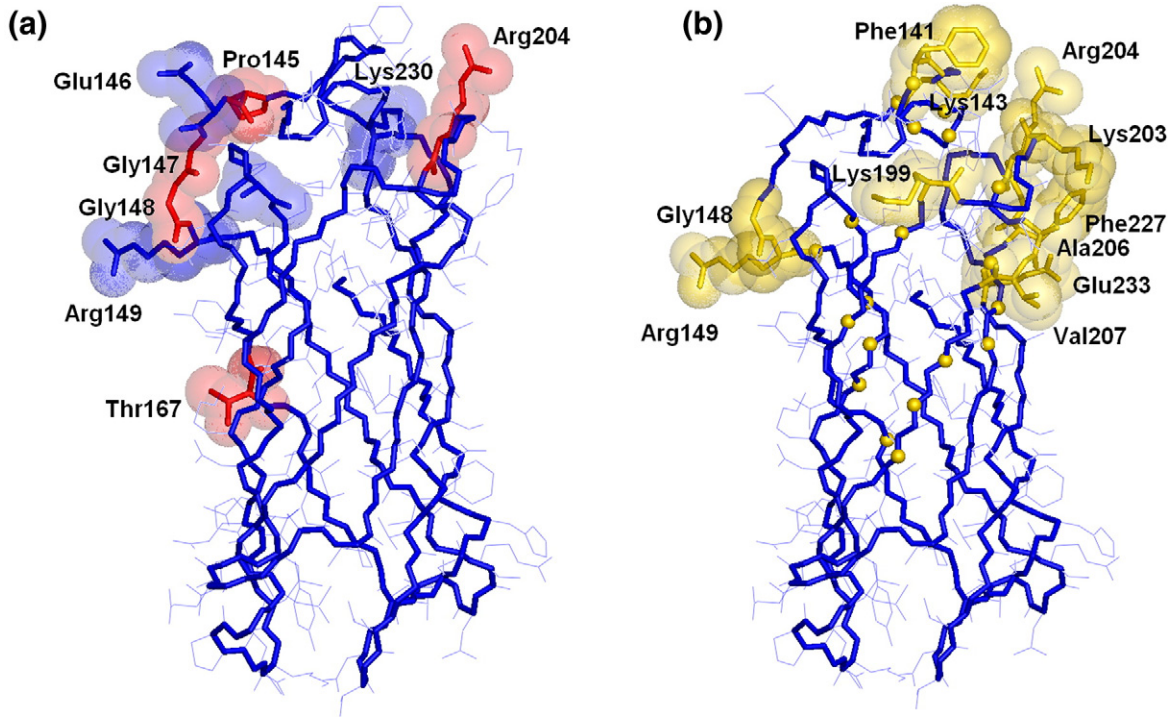


Fig. 2. Residues of fHbp that form the MAb502 epitope. (a) Spatial distribution of amino acids predicted as relevant (red surface) or accessory (blue surface) for MAb502 binding onto the β-barrel domain of fHbp. (b) Amino acids whose NMR signals were perturbed upon the addition of FAb502 to the NMR sample. van der Waals volumes are reported for surface-exposed amino acids; buried residues are marked by solid spheres in correspondence to their backbone nitrogen atoms.

by site-directed mutagenesis on the recombinant fHbp from MC58. When the Arg204 was changed to histidine, MAb502 was unable to recognize the protein. In close proximity to Arg204 were located three amino acids variable only in WB-negative strains, namely, Pro145, Gly147, and Gly148. On the basis of this observation, we predicted that the key residues of epitope recognized by MAb502 were Arg204 together with Pro145, Gly147, and Gly148.

In order to verify the assumption that residues classified as dispensable in Fig. 1 have a minor or no effect on MAb502 binding, we tested by WB the effects of Gly202→Lys and Asn178→His substitutions. Both mutations, as expected, did not impair the protein recognition by MAb502 in WB (Fig. 3). However, because of the absence of bactericidal activity of MAb502 against all the strains except MC58, we could not exclude *a priori* the contribution by some of these dispensable residues in determining antibody binding affinity. For this reason, we speculated that an additional contribution to the fHbp affinity for MAb502 could be given by the side chains of Glu146, Arg149, Ala174, and Lys230, located within the proximity of the residues predicted as crucial for epitope formation.

NMR mapping of the FAb502–fHbp$_C$ interaction

The predictions on fHbp residues interacting with MAb502 were tested through NMR spectroscopy, by analyzing the perturbations caused in the ^{1}H–^{15}N heteronuclear single-quantum coherence (HSQC) spectrum of fHbp$_C$ upon addition of the FAb (FAb502). Addition of FAb502 to ^{15}N-labeled fHbp$_C$ caused the progressive reduction in intensity of NH cross-peaks in the ^{1}H–^{15}N HSQC spectra, which completely disappeared at the 1:1 ratio (Fig. 4a). The only exception was represented by residues 100–136

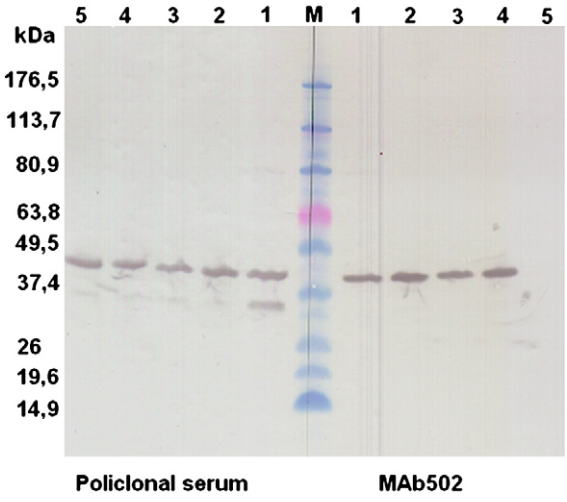


Fig. 3. WB of purified recombinant fHbp from MC58 and related mutants carried out by using the polyclonal antisera (left) or MAb502 (right). Lane 1, fHbp; lane 2, fHbp Asn178→His; lane 3, fHbp Gly202→Lys; lane 4, fHbp Asn178→His–Gly202→Lys; lane 5, fHbp Arg204→His. "M" indicates molecular weight marker.

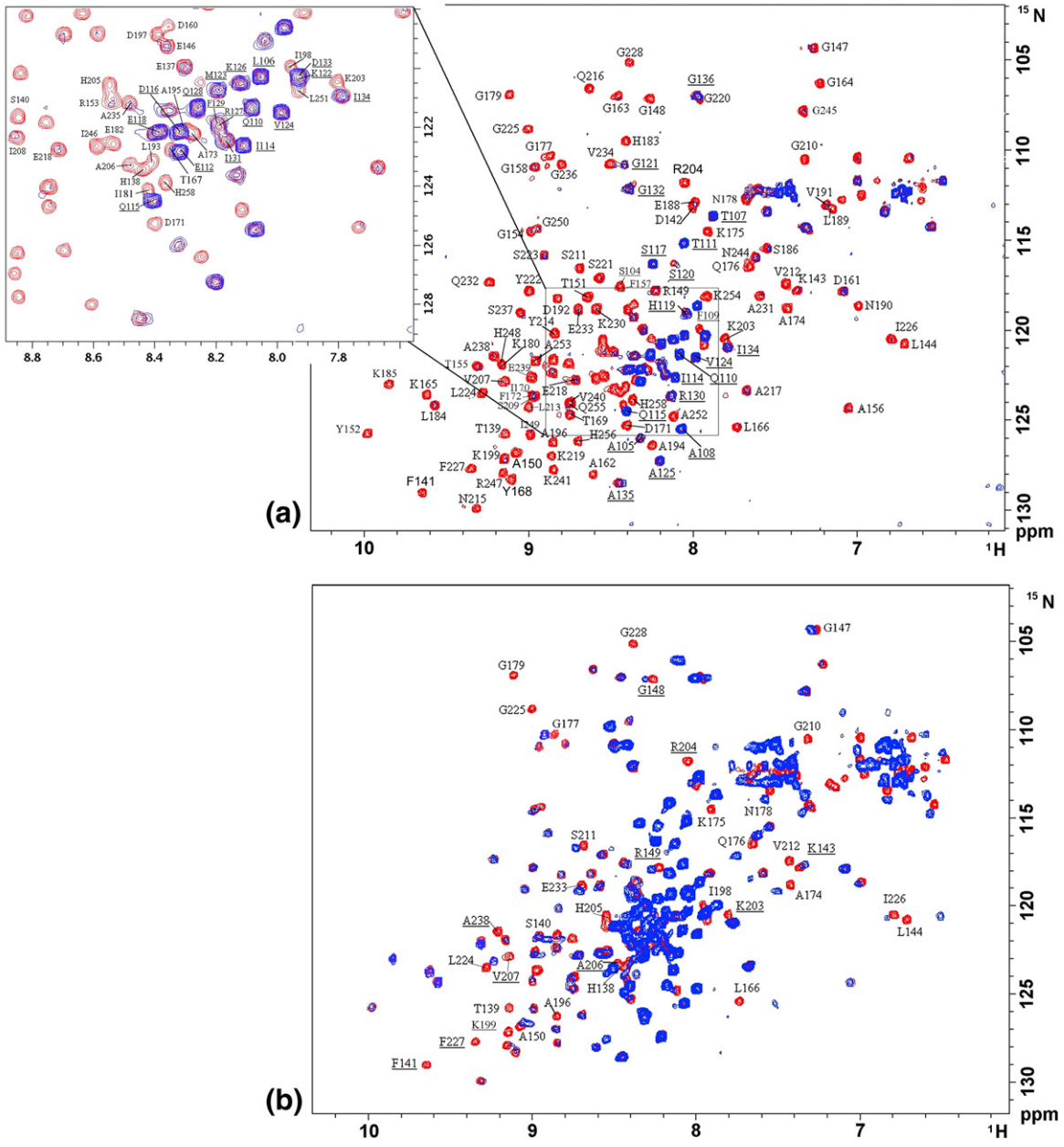


Fig. 4. Interaction of fHbp$_C$ with FAb502. (a) ^{1}H–^{15}N HSQC spectrum showing fHbp$_C$ (red) superimposed with fHbp$_C$ in the presence of one equivalent of FAb502 (blue). Labels of residues 100–136 belonging to the N-terminal tail are underlined. (b) ^{1}H–^{15}N HSQC spectrum of fHbp$_C$ (red) superimposed with the CRINEPT spectrum acquired on the fHbp$_C$–FAb502 1:1 mixture (blue). Only amino acids of fHbp$_C$ whose amide signals experience chemical shift changes upon the addition of FAb502 are labeled. Among the latter, only the labels of the ones that are solvent exposed are underlined. For the unstructured N-terminal tail, both multiplet components are observed in the CRINEPT spectrum.

belonging to the N-terminal tail, which maintain their full intensity. Disappearance of the protein backbone NHs in the HSQC spectra is a consequence of the formation of the antibody–protein complex, characterized by a high molecular mass (~70 kDa). The progressive disappearance of fHbp$_C$ signals upon addition of FAb502 and the absence of any signal shifting suggested that the complex exchanges with the free proteins at rates slower than the chemical shift differences between the two forms (i.e., in the range of milliseconds). The high mobile N-terminal tail, which reorients faster than the overall tumbling rate of fHbp$_C$, maintained this property in the complex. This behavior indicated that it was not involved in the interaction with FAb2502. Similar chemical shift values for the N-terminal region in fHbp$_C$ and the fHbp$_C$–FAb502 complex also indicated that no significant conformational change occurred in this region. When cross-correlated relaxation-enhanced polarization transfer (CRINEPT) NMR experiments, which are suitable for high-molecular-weight systems, were performed, the majority of the NH cross-peaks reappeared (Fig. 4b). The residues of fHbp$_C$, which experienced different chemical shifts upon

Table 2. Amino acids of $fHbp_C$ whose amide signals experience chemical shift changes upon the addition of FAb502

Exposure[a]	Residue
Buried	His138, Thr139, Ser140, Leu144 Ala150, Leu166, Ala174, Lys175, Gln176, Gly177, Asn178, Gly179, Ala196, Ile198, His205, Gly210, Ser211, Val212, Leu224, Gly225, Ile226, Gly228, Glu233
Exposed	Phe141, Lys143, Gly148, Arg149 Lys199, Lys203, Arg204, Ala206, Val207, Phe227, Ala238

[a] As it results from the solvent accessibility calculated with the program NACCESS.[11]

addition of FAb502, are shown in Fig. 2b and listed in Table 2. They cluster in one region of $fHbp_C$ and include Arg204, the residue essential for binding to the antibody,[8] as well as Gly148, Arg149, and Ala174. In addition, spectral perturbations were observed also for other NH cross-peaks of the surface-exposed residues, namely, Phe141, Lys143, Lys199, Lys203, Ala206, Val207, Phe227, Gly228, and Glu233 (Fig. 2b).

Computer modeling of the $fHbp_C$–Fv502 complex

Molecular docking is a computational methodology that uses the known structures of the components of a complex to generate a model for the complex. Docking approaches are becoming increasingly popular for investigating intermolecular complexes whose structures cannot be studied by classical structural methods.[12,13] They have been applied to a variety of protein–protein complexes,[14,15] including antibody–antigen complexes.[16]

To model the structure of the complex between $fHbp_C$ and the hypervariable FAb of MAb502 (Fv502), we performed docking calculations using as input the NMR structure of $fHbp_C$ and a computer homology model of Fv502.

NMR data were used to derive the interaction restraints on $fHbp_C$. In the absence of any experimental indication on Fv502 residues actually contacting the antigen, all the surface-exposed amino acids within the six hypervariable loops of the Fv fragment were used as restraints (Table S1).

As a consequence of the wide interaction surface defined on Fv502, more than one model of the complex was obtained, and the same portion of the $fHbp_C$ β-barrel was involved in the interaction with Fv502 in all of them. From the analysis of the energetic and scoring functions, it appeared that, essentially, only two clusters are meaningful. These two lowest-energy clusters experienced far much better scoring functions with respect to the others, larger interaction surfaces, and a much lower average number of ambiguous restraints, which indicate better agreement with the experimental data (Table S2). Therefore, we performed a detailed analysis only on these two clusters (Table 3). The lowest-energy complex for each of the two clusters is shown in Fig. 5.

The list of fHbp residues contacting Fv502 in each structural model is given in Table 4, and a summary of the intermolecular interactions calculated over the best five structures of each cluster is reported in Table S4.

Even if the reciprocal orientations of the molecules are different, the two clusters showed some common features, such as the involvement of Arg204 of $fHbp_C$. In particular, in the first model, hydrogen bonds were present between the side chain of Arg204 and the carbonyl oxygen of residues Thr158 and Ser159 of Fv502. In the second model, the side chain of Arg204 interacted with the same region of Fv502, forming indeed a salt bridge not only with Asp70 but also with Asp124, both of which are located in the Fv heavy chain (Fv_H) of Fv502 (Fig. 5). Another region of interaction present in both clusters involves residues Phe227–Lys230, located in the loop between β-strands 6 and 7 of $fHbp_C$ and residues Tyr71–Asp76 of Fv_H502.

The two structural models of the complex differed, on the other hand, in the mode of involvement of the stretch Glu146–Arg149. Hydrogen bonds are present in the structural models of the first cluster between the backbone of Gly146 and Arg149 of $fHbp_C$ and the side chains of Tyr51 and His49 of Fv light chain (Fv_L) 502. These interactions are absent in the second cluster, being the stretch Gly147–Arg149 completely exposed to the solvent (Fig. 4).

MAb502 does not inhibit the fHbp binding to fH

In order to investigate if MAb505 could interfere with the fH–fHbp interaction, we evaluated by FACS the ability of human factor H to bind meningococcal cell surface after preincubation with two monoclonal antibodies, namely, MAb502 and JAR5. The latter antibody was previously shown to cause complete inhibition of fH binding.[17] Results

Table 3. Statistics on the clusters of the structural models of the $fHbp_C$–Fv502 complex obtained through docking calculations

	Cluster 1[a](best five)	Cluster 2 (best five)
HADDOCK score[b]	−134 (4)	−133 (11)
RMSD (Å)[c]	1.87 (1.23)	0.76 (0.38)
BSA (Å^2)[d]	2062 (123)	2097 (42)
AIR viol[e]	11.0 (0.9)	12.00 (1.3)
	Cluster 1 (all)	Cluster 2 (all)
HADDOCK score	−104 (23)	−87 (25)
RMSD (Å)	3.1 (1.0)	3.4 (2.1)
BSA (Å^2)	1915 (135)	1851 (146)
AIR viol	12.4 (1.6)	14.8 (2.2)

The two clusters of solutions contained 35 and 69 structures, respectively. Averages (standard deviations are reported in parentheses) were calculated over the best five model structures and over all the structural models for each of the two clusters of solutions.

[a] Cluster rank according to the HADDOCK score.

[b] HADDOCK score defined as the weighted sum of different energetic terms: score = (1 E_{vdw} + 0.2 E_{elc} + 0.1 E_{dist} + 1 E_{solv}), where the four terms are van der Waals energy, electrostatic energy, distance restraint energy, and desolvation energy, respectively.

[c] Backbone RMSD from the lowest HADDOCK score structure in each cluster.

[d] Buried surface area.

[e] Average number of AIR violations for each cluster.

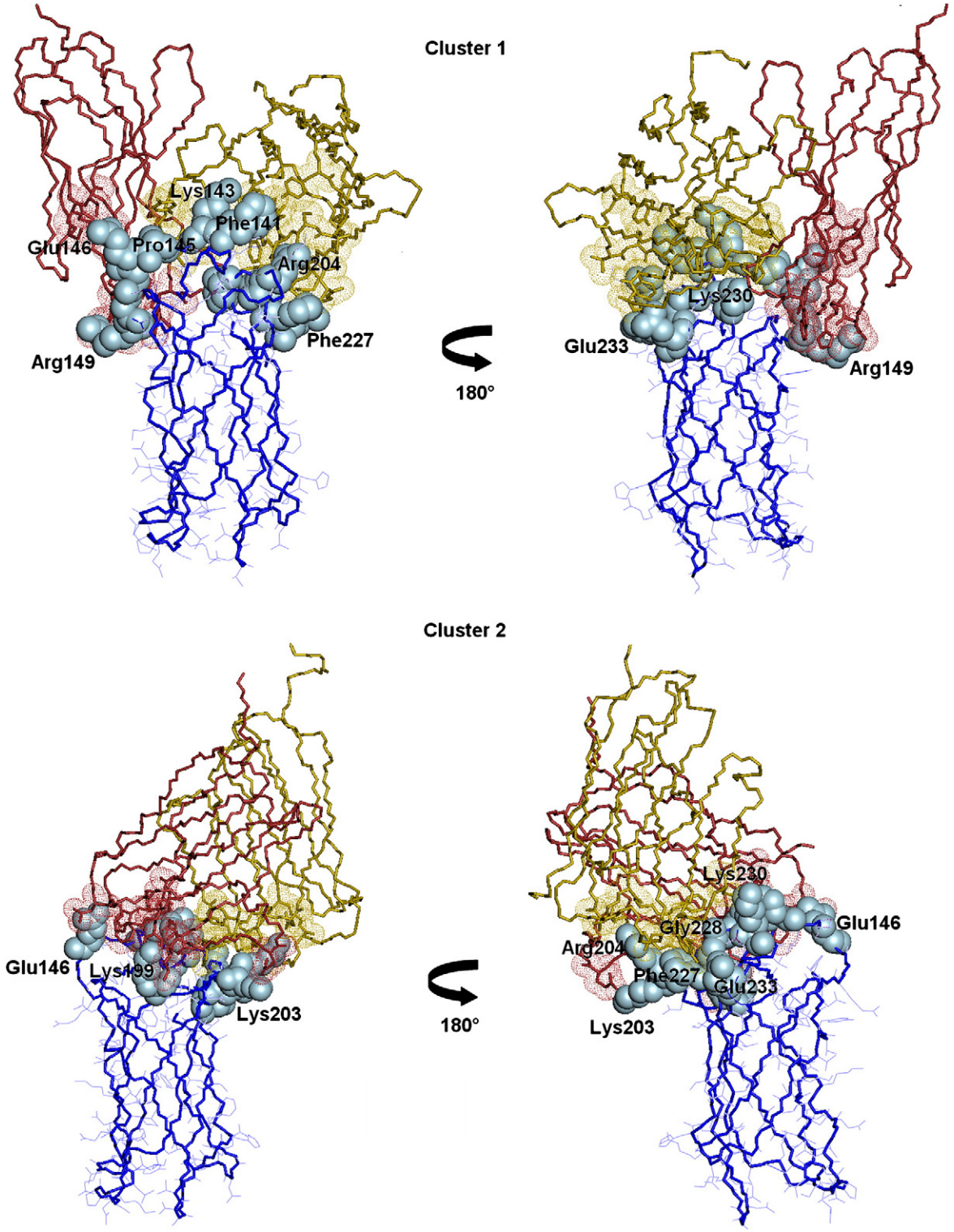


Fig. 5. Structural models of $fHbp_C$–Fv502 complex generated by molecular docking. Lowest HADDOCK score structure of cluster 1 (top) and that of cluster 2 (bottom) are represented. fHbp is shown in blue, while Fv502 chains are shown in gold (Fv_H) and red (Fv_L). Key residues involved in the intermolecular interactions are shown as solid ($fHbp_C$) and dotted (Fv502) surfaces.

are summarized in Fig. 6. While the presence of JAR5 led to the disappearance of the fluorescence peak diagnostic of fH binding (Fig. 6a), cells pre-incubated with MAb502 were still able to retain factor H on their surface (Fig. 6b). When the same experiment was repeated with the MC58 isogenic mutant lacking the fHbp gene, no difference in the FACS profile was observed (data not shown). These data suggested that the fH binding site does not overlap with the MAb502 epitope.

Discussion

Detailed knowledge of epitopes recognized by protective antibodies is necessary for understanding the molecular mechanisms responsible for effective immune response. In addition, epitope mapping of bactericidal monoclonal antibodies provides a powerful tool to identify *a priori* bacterial strains carrying susceptible antigens. For these reasons, we undertook the characterization of the epitope recognized by MAb502 raised against fHbp, a powerful antigen of *N. meningitidis*.

The present results, obtained from the analysis of immunological data, NMR mapping, and molecular docking carried out on fHbp protein in the presence of MAb502, unequivocally demonstrate that the fHbp contact area with the antibody is formed by three independent loops plus the N-terminal α-helix of the C-terminal domain $fHbp_C$. In particular, they evidenced the conformational nature of the MAb502 epitope and confirmed the key role of Arg204, as well the minor role of residues Glu146–Arg149.

The reactivity of MAb502 against a panel of meningococcal strains carrying different alleles of fHbp provided a first indication of which residues are more likely involved in epitope formation. FACS analysis and bactericidal activity of the polyclonal serum anti-fHbp showed that all the strains analyzed expressed and exported fHbp on their surface. We therefore assumed that differences in amino acid sequences could be interpreted as a major factor influencing the recognition. Consequently, we presumed that the MAb502 epitope could potentially contain all the amino acids that varied in WB-negative alleles. However, it should be considered that variability not necessarily implies involvement in the epitope and that not all the differences observed in WB-negative alleles could equally contribute to the MAb502 recognition. In previous work, we already indicated Arg204 as a necessary residue for MAb502 recognition.[8] The three-dimensional structure of $fHbp_C$ permitted us to assess that the surface-exposed side chains of Pro145, Gly147, and Gly148 are localized within the proximity of Arg204. This spatial distribution led us to the initial conclusion that Arg204, Pro145, Gly147, and Gly148 could likely form the binding region to MAb502. Remaining positions variable in WB-negative strains were located in the flexible tail of the NMR structure of fHbp, making it difficult to speculate *a priori* about their relevance for MAb502 binding. Among residues varying either in WB-negative or WB-positive alleles, we considered that Glu146, Arg149, Ala174, and Lys230, due to their exposure and closeness to Arg204, could form an accessory part of the MAb502 epitope.

NMR data confirmed in large extent the prediction based on the immunological data, as resonances of Gly148, Arg149, and Arg204 were clearly perturbed upon FAb502 interaction. The ^{15}N resonance of Glu146 is overlapped to other signals in the NMR spectra, preventing us from definitively confirming its inclusion into the epitope as predicted by sequence alignment. In addition, NMR data suggested the involvement of other amino acids, such as Phe227, Gly228, and Glu233, otherwise unpredictable from sequence alignment because of their conservation among all the alleles analyzed. NMR data also showed that the N-terminal flexible tail is not involved in the interaction with FAb502. A number of signals belonging to residues buried within the fHbp interior were also perturbed upon FAb502 addition. These spectral changes might be a result of structural rearrangements induced by fHbp binding

Table 4. fHbp residues forming the MAb502 epitope

Predicted	NMR	Cluster 1	Cluster 2
	Phe141	Phe141	Phe141
	Lys143	Lys143	Lys143
Pro145	[a]	Pro145	
Glu146	[b]	Glu146	Glu146
Gly147	[b]	Gly147	
Gly148	Gly148	Gly148	
Arg149	Arg149	Arg149	
Ala174	Ala174		
	Lys199		Lys199
	Lys203		Lys203
Arg204	Arg204	Arg204	Arg204
	Ala206		
	Val207		
	Phe227	Phe227	Phe227
	Gly228	Gly228	Gly228
Lys230	[b]	Lys230	Lys230
	Glu233	Glu233	Glu233

Predicted = predicted by sequence alignment. NMR = deduced by NMR experiments. Cluster 1 and Cluster 2 = identified in the docking complexes.

[a] Proline residue has no amide proton.

[b] These residues have not been reported due to spectral overlap of their backbone amides or as they have a chemical shift perturbation lower than 0.03 ppm.

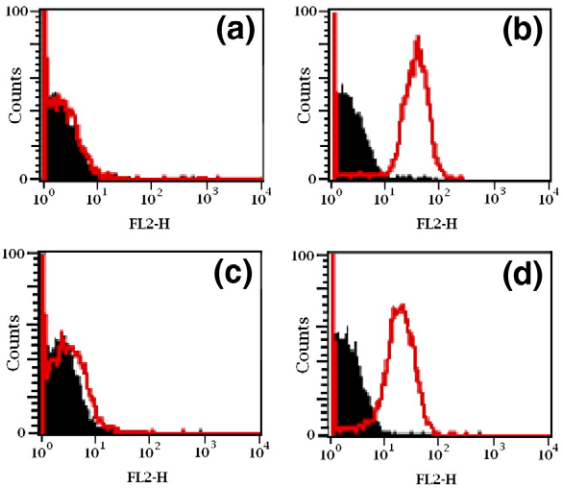


Fig. 6. Effects of MAb502 on fH binding. FACS profile of meningococcal cells in the absence (black) and in the presence (red) of human factor H preincubated with (a) JAR5, (b) MAb502, (c) mice polyclonal serum elicited against recombinant fHbp, and (d) PBS buffer.

more than of direct contact with the antibody. This behavior has been indeed observed in a number of protein–protein complexes characterized by NMR.[18]

In order to derive from these NMR and immunological data a structural model of the complex and to evaluate if all the 12 accessible amino acids identified by NMR mapping could constitute the MAb502 epitope, we modeled the fHbp–Fv502 complex by molecular docking, from which two possible conformations were obtained. The epitopes identified as complexes 1 and 2 are largely overlapping, with 8 residues involved in the antigen–antibody interaction for both models (Table 4). In both cases, Ala206 and Val207 are not directly involved in the interaction with the antibody, suggesting that alterations observed for their NH cross-peaks were likely due to induced conformational changes.

The main difference between the two clusters lies in the level of involvement of loop Glu146–Arg149. These residues are all located in the contact area of the first model, while only Glu146 experiences some interaction with Fv502 in the second.

Overall, our results demonstrated the direct interaction of Arg204 with the antibody binding site and evidenced the involvement of Glu146–Arg149 in the formation of the antigen–antibody contact area, albeit to a minor extent.

This epitope mapping study also provides indirect information on the localization of the factor H binding site. Here, we showed that MAb502 was unable to prevent fH binding onto the fHbp surface in contrast to JAR5. Although we cannot exclude that such absence of inhibition was due to the lower affinity of MAb502 to fHbp, these data seem to suggest that the area targeted by this antibody is distinct and sufficiently far away from the factor H binding site to not compromise the formation of the fH–fHbp complex. Our results fit the conclusions of the epitope mapping study recently reported by Beernink *et al.*,[17] who observed fH binding inhibition by monoclonals recognizing Gly121 in N-terminal domain of the protein and partially by the ionic pair Lys180–Glu192 localized on the opposite side of $fHbp_C$ with respect to Arg204.

The approach used here, which combined immunological, FACS, WB, and NMR characterization, was particularly powerful in defining the MAb502 epitope. These results represent the first step of an experimental strategy in which vaccine candidates can be designed to contain broad repertoires of natural protective epitopes identified by molecular mapping.

Materials and Methods

Bacterial strains

Escherichia coli DH5 [F^- 80*lacZ*M15 (*lacZYA-argF*) U169 *deoR recA1 endA1 hsdR17*(r_k^- m_k^+) *phoA supE44 thi-1 gyrA96 relA1 tonA*] and BL21 Star (DE3) [F^- *ompT hsdS*$_B$ (r_B^- m_B^-) *gal dcm rne131* (*DE3*)] (Invitrogen) were used as the cloning strain and expression host, respectively. *N. meningitidis* strains MC58, F6124, M2197, NZ98/254, and M4030 were previously described.[8] Strains M01-240149 and M01-240345 were isolated in the UK in 2001 (a kind gift of Steve Gray, HPA, Manchester, UK). LNP17592 was a W-135 strain, isolated in 2000 (a kind gift of Muhamed-Kheir Taha, Institut Pasteur, Paris, France). Strains M6190 and M2937 were isolated in the United States in 1999 and 1996, respectively (kind gifts of Tanija Popovic, CDC, Atlanta, GA).

SDS-PAGE and WB analysis

Purified fHbp from MC58, the mutants Asn178→His, Gly202→Lys, and Arg204→His, and the double mutant Asn178→His Gly204→Lys (1 μg/lane) were analyzed by WB. Samples were loaded onto 4%–12% NuPAGE polyacrylamide gel (Invitrogen) in Mes SDS running buffer (Invitrogen) and transferred to a nitrocellulose membrane (Bio-Rad).[19]

WB analysis was performed according to standard procedures[20] using anti-MC58 fHbp mouse polyclonal antisera at a 1:20,000 dilution in phosphate-buffered saline (PBS) containing 3% (w/v) skimmed milk and 0.01% Triton and anti-fHbp MAb502 at a 1:5000 dilution (0.34 μg/ml), followed by a 1:2000 dilution of horseradish peroxidase-labeled antimouse immunoglobulins (DAKO). The signal was developed with a Bio-Rad Opti-4CN substrate kit. Scanning was performed using LabScan (Pharmacia) and Imagemaster software (Pharmacia).

FACS analysis

The analysis was determined using a FACScan flow cytometer as described previously.[8] Polyclonal sera were tested at a 1:200 dilution. Antibody binding was detected using a secondary antimouse antibody (whole molecule) that was FITC conjugated (Sigma).

Serum bactericidal assay

Serum bactericidal activity (BCA) against *N. meningitidis* strains was evaluated as reported in Ref. 8. Pooled baby rabbit sera have been used as a complement source (Cedarlane or Pel-Freez).

Sequencing of Fv502

Approximately 5×10^6 MAb502-secreting hybridoma cells were collected and homogenized via passage through a 20-guage needle using an RNase-free 5-ml syringe. Poly (A)+RNA was isolated using an Oligotex Direct mRNA Mini Kit according to the manufacturer's instructions (QIAGEN®). cDNA was produced via reverse transcription using ~200 ng of the poly(A)+RNA template, an oligo-$(dT)_{12–18}$ primer, and SuperScript™ II RT (Invitrogen). Reactions were incubated at 42 °C for 50 min, followed by inactivation of the reverse transcriptase at 70 °C for 15 min. RNA was removed from the reactions by incubation with RNase H (Invitrogen) at 37 °C for 20 min. The resulting cDNA was used as a template for PCR amplification using Platinum® *Taq* DNA Polymerase High Fidelity (Invitrogen) and various 5′ degenerate primers and 3′ primers specific for Fv_H and Fv_L gene fragments (Mouse Ig-Primer Set, Novagen).

PCRs for Fv_H were incubated at 94 °C for 2 min, followed by 30 cycles of 94 °C for 1 min, 58 °C for 1 min 30 s, and

68 °C for 1 min. A final extension at 68 °C for 10 min completed the reactions. For the Fv_L, PCR was performed using an additional primer complementary to the aberrant light chain CDR3 (5′-ACC TAT TAC TGT CAG CAC ATT A-3′) in molar excess to force truncation of the aberrant cDNA product. PCRs for Fv_L were incubated at 94 °C for 2 min, followed by 30 cycles of 94 °C for 1 min, 62 °C for 1 min, and 72 °C for 1 min 30 s. A final extension at 72 °C for 10 min completed the reactions. Ten microliters of the PCR products was analyzed on 1% agarose gel.

The PCR product of Fv_H was purified using a QIAprep Spin Miniprep Kit (QIAGEN®). The PCR product of the Fv_L gene fragment of the correct size (350–450 base pairs) was excised from agarose gel and purified with a QIAquick Gel Extraction kit (QIAGEN®).

The Fv_H and Fv_L purified fragments were converted to blunt phosphorylated forms using an End Conversion Mix (Novagen) and were inserted into the EcoRV site of the pSTBlue-1 Perfectly Blunt™ vector (Novagen) and transformed into NovaBlue Singles™ competent cells (Novagen) according to the manufacturer's instructions. Positive colonies were selected by white/blue screening after overnight growth, at 37 °C, on Luria–Bertani agar plates containing ampicillin (50 μg/ml), X-gal (70 μg/ml), and IPTG (80 μg/ml). White (positive) colonies, containing the plasmid with insert, were incubated in 4 ml of Luria–Bertani broth containing ampicillin (100 μg/ml) for 12–16 h with vigorous shaking at 37 °C, and plasmids were isolated using a High Pure Plasmid Isolation Kit (ROCHE), according to the manufacturer's instructions. BamHI and HindIII digestion of ~1 μg of the resulting plasmids and resolution of the products via agarose gel electrophoresis confirmed the presence of the appropriately sized inserts. Positive clones were sequenced with T7 PR (TAA TAC GAC TCA CTA TAG GG) and pSTB rev (CCA CAC GTG TGG TCT AGA GC) primers.

Nucleotide sequences of Fv_L and Fv_H are reported in Fig. S1. The GenBank accession numbers are EU835941 for Fv_L and EU835942 for Fv_H.

Preparation of FAb502

MAb502 was produced and purified as previously described.[8] MAb502 was dialyzed in 20 mM sodium phosphate and 10 mM ethylenediaminetetraacetic acid, pH 7.0, buffer and then incubated with immobilized papain cross-linked on agarose support with respect to the condition indicated in the kit specifics (Pierce code 20341) in order to generate the corresponding FAb (FAb502). FAb502 was subsequently purified by affinity chromatography using HiTrap Protein A Fast Flow (GE Healthcare code 17-5079-01) in 20 mM sodium phosphate, pH 7.0, buffer and eluted in 0.1 M sodium citrate, pH 5.0, buffer. The fractions eluted were checked on SDS-PAGE gel in nonreducing conditions, confirming the correct molecular size of 50 kDa. The purest fractions were pooled and dialyzed 1:100 overnight at 4 °C against 10 mM sodium phosphate, pH 7.0, buffer. The purified FAb502 was then concentrated using a Centricon YM-30 (molecular weight cutoff = 30,000) concentrator reaching a final concentration of 0.1 mM. FAb502 was then tested in WB against fHbp to assess the persistence of the activity.

^{15}N-labeled fHbp cloning, expression, and purification

^{15}N-labeled $fHbp_C$ protein fragment was expressed and purified in *E. coli* as previously described.[10] Purified ^{15}N-labeled $fHbp_C$ was concentrated using Centricon YM-10 (molecular weight cutoff = 10,000) to a final concentration of 0.8 mM suitable for the NMR analyses.

NMR mapping of FAb502–$fHbp_C$ interaction

The interaction of FAb502 with ^{15}N-labeled $fHbp_C$ was investigated by NMR spectroscopy with ^{1}H–^{15}N HSQC and CRINEPT experiments. The latter was recorded for 12 h, with a data size of 1024 × 232 points, a CRINEPT transfer of 3.3 ms, and a recycle time of 1.1s. All the NMR experiments were performed at 298 K on a Bruker Avance 900 spectrometer, working at 900.13-MHz frequency and equipped with a cryogenically cooled probe.

Titrations were performed on 0.8 mM ^{15}N-labeled $fHbp_C$ protein samples with the unlabeled FAb502 up to an $fHbp_C$–FAb502 molar ratio of 1:1. ^{1}H and ^{15}N resonance assignments for the $fHbp_C$ protein were already available.[10] All these samples were in 20 mM phosphate buffer (90% H_2O and 10% D_2O) at pH 7.0.

Modeling of Fv502

The Fv fragment is the smallest FAb of the antibody made of the variable domains of Fv_L and Fv_H. We determined the nucleotide sequences of Fv_L and Fv_H of MAb502. The corresponding amino acid sequences (Fig. S1) were then used to scan the Brookhaven Protein Data Bank (PDB) in order to retrieve suitable templates for modeling. Eighteen and 29 hits were retrieved for Fv_L and Fv_H, respectively, with percentages of identity ranging from 96% to 91% for Fv_L and from 82% to 74% for Fv_H. Crystal structures of an antibody variable heavy chain (PDB entry 32C2) and an antibody variable light chain (PDB entry 1WEJ), showing 96% and 82% of identity to Fv_H and Fv_L, respectively, of MAb502 were selected as templates. Models of Fv_L and Fv_H of MAb502 were obtained with Swiss-PDB viewer[21] software by alignment and superimposition to the corresponding template chains and used to assemble the whole Fv502. The resulting model was refined by energy minimization and molecular dynamics (MD). The molecule was solvated in a water box of size 53.7 × 60.7 × 58.77, and ions were then added to neutralize the resulting configuration. The system was then equilibrated using the following schedule: (i) the system was minimized for 300 steps, fixing the position of the protein; (ii) the system was minimized for a further 300 steps followed by 1000 steps of MD, fixing the position of water molecules and charged ions; (iii) the system was minimized for a further 300 steps followed by 10,000 MD steps, fixing the position of the protein; and (iv) the system was minimized for a further 300 steps followed by a final MD simulation of 5×10^5 MD steps. All simulations were performed using NAMD 2.6[22] and the Charmm force field in the NPT ensemble at a pressure of 1 atm and a temperature of 300 K†. In all simulations, an integration time step of 2 fs was used. Quality of models was evaluated by using the program PROCHECK[23] and by visual inspection with the computer graphics program VMD.[24]

†NAMD was developed by the Theoretical and Computational Biophysics Group in the Beckman Institute for Advanced Science and Technology at the University of Illinois at Urbana–Champaign.

Docking calculations

Docking calculations were performed with HADDOCK2.0.[25] The docking calculations in HADDOCK are driven by ambiguous interaction restraints (AIRs) between all residues involved in the intermolecular interactions.[26] Active and passive residues (Table S1) were defined according to the HADDOCK protocol. Solvent accessibility was calculated with the program NACCESS Fv_L.

As structures of the isolated proteins, 20 NMR conformers of the $fHbp_C$ structure were used. For FAb502, structural models were generated as described above and the lowest-energy structural model among those generated by MD was chosen for docking calculations. Flexible regions of the proteins were defined based on the active and passive residues plus two preceding and following residues.

In the initial rigid-body docking calculation phase, 1000 structures of the complex were generated, and the best 200 in terms of total intermolecular energy were further submitted to the semiflexible simulated annealing and final refinement in water. These 200 structures were then clustered using a cutoff of 7.5 Å of RMSD among any structure of a cluster. Only clusters containing at least 5 structures were taken and analyzed. Table S2 reports the statistics calculated over the best five models of each cluster as well as over all the components of the cluster. Table S3 reports average pairwise RMSDs among the clusters.

Antibody inhibition and fH binding reactions

Bacteria were isolated in a single colony and then inoculated in 7 ml of Mueller–Hinton broth plus 0.25% D-glucose until the optical density reached 0.1. Cultures were successively grown at 37 °C with 5% CO_2 until the optical density reached 0.6.

The culture was transferred to a sterile 50-ml conical tube and washed with 40 ml of PBS plus 1% bovine serum albumin (BSA) and then centrifuged at 3700 rpm (2908 RCF) for 10 min. The pellet was resuspended into 2 ml of PBS plus 1% BSA [~3 to 5×10^8 colony-forming units (CFUs)/ml].

Ten microliters of each MAb (final concentration = 50 μg/ml) or polyclonal serum (final dilution = 1:400) was added to an aliquot of 85-μl bacterial suspension, and the mixtures were incubated at 37 °C for 10 min with gentle shaking. Five microliters of human factor H (Calbiochem) solution was added to achieve a final concentration of 50 μg/ml in the final reaction mixture. The samples were incubated at 37 °C for 20 min with shaking. The mixture was then centrifuged at 13,400 RCF in a microcentrifuge for 3 min to pellet the bacteria. After the aspiration of the supernatant, the bacteria were resuspended in 100 μl of PBS plus 1% BSA containing 1:100 dilution (final) goat antihuman fH (Calbiochem). After incubation at room temperature for 15 min, bacteria were pelleted and resuspended in 100 μl of PBS plus 1% BSA containing FITC-conjugated donkey antigoat immunoglobulin G (Jackson Immunoresearch) (final working concentration = 1:100).

After 15 min at room temperature, bacteria were again pelleted and resuspended in 1 ml of filter-sterilized, freshly made 0.5% formaldehyde in 1× PBS for FACS analysis, carried out as already described.[5]

Accession numbers

The GenBank accession numbers are EU835941 for Fv_L and EU835942 for Fv_H.

Acknowledgements

This work was supported by MIUR-FIRB (Fondo per gli Investimenti della Ricerca di Base) through MIUR no. RBLA032ZM7 (12/09/05-12/09/10) ("Piattaforme NMR per lo studio dell'interazione proteine-leganti di interesse farmacologico").

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2008.12.005

References

1. Oster, P., O'Hallahan, J., Aaberge, I., Tilman, S., Ypma, E. & Martin, D. (2007). Immunogenicity and safety of a strain-specific MenB OMV vaccine delivered to under 5-year olds in New Zealand. Vaccines, Immunisation and Immunotherapy—Based on the Fifth World Congress on Vaccines, Immunisation and Immunotherapy, 25, 3075–3079.
2. de Moraes, J. C., Perkins, B. A., Camargo, M. C., Hidalgo, N. T., Barbosa, H. A., Sacchi, C. T. *et al.* (1992). Protective efficacy of a serogroup B meningococcal vaccine in Sao Paulo, Brazil. *Lancet*, **340**, 1074–1078.
3. Giuliani, M. M., Adu-Bobie, J., Comanducci, M., Arico, B., Savino, S., Santini, L. *et al.* (2006). Inaugural article: A universal vaccine for serogroup B meningococcus. *Proc. Natl Acad. Sci. USA*, **103**, 10834–10839.
4. Pizza, M., Scarlato, V., Masignani, V., Giuliani, M. M., Aricò, B., Comanducci, M. *et al.* (2000). Identification of vaccine candidates against serogroup B meningococcus by whole-genome sequencing. *Science*, **287**, 1816–1820.
5. Masignani, V., Comanducci, M., Giuliani, M. M., Bambini, S., Adu-Bobie, J., Arico, B. *et al.* (2003). Vaccination against *Neisseria meningitidis* using three variants of the lipoprotein GNA1870. *J. Exp. Med.* **197**, 789–799.
6. Fletcher, L. D., Bernfield, L., Barniak, V., Farley, J. E., Howell, A., Knauf, M. *et al.* (2004). Vaccine potential of the *Neisseria meningitidis* 2086 lipoprotein. *Infect. Immun.* **72**, 2088–2100.
7. Madico, G., Welsch, J. A., Lewis, L. A., McNaughton, A., Perlman, D. H., Costello, C. E. *et al.* (2006). The meningococcal vaccine candidate GNA1870 binds the complement regulatory protein factor H and enhances serum resistance. *J. Immunol.* **177**, 501–510.
8. Giuliani, M. M., Santini, L., Brunelli, B., Biolchi, A., Arico, B., Di Marcello, F. *et al.* (2005). The region comprising amino acids 100 to 255 of *Neisseria meningitidis* lipoprotein GNA 1870 elicits bactericidal antibodies. *Infect. Immun.* **73**, 1151–1160.
9. Welsch, J. A., Rossi, R., Comanducci, M. & Granoff, D. M. (2004). Protective activity of monoclonal antibodies to genome-derived neisserial antigen 1870, a *Neisseria meningitidis* candidate vaccine. *J. Immunol.* **172**, 5606–5615.
10. Cantini, F., Savino, S., Scarselli, M., Masignani, V., Pizza, M., Romagnoli, G. *et al.* (2006). Solution structure of the immunodominant domain of protective antigen GNA1870 of *Neisseria meningitidis*. *J. Biol. Chem.* **281**, 7220–7227.

11. Hubbard, S. J. & hornton, J. M. (1993). *NACCESS.* Department of Biochemistry and Molecular Biology, University College London, London, UK.
12. Bonvin, A. M. (2006). Flexible protein–protein docking. *Curr. Opin. Struct. Biol.* **16**, 194–200.
13. Gray, J. J. (2006). High-resolution protein–protein docking. *Curr. Opin. Struct. Biol.* **16**, 183–193.
14. van Dijk, A. D., Ciofi-Baffoni, S., Banci, L., Bertini, I., Boelens, R. & Bonvin, A. M. J. (2007). Modeling protein–protein complexes involved in the cytochrome C oxidase copper-delivery pathway. *Proteome Res.* **6**, 1530–1539.
15. Bertini, I., Cavallaro, G. & Rosato, A. (2005). A structural model for the adduct between cytochrome *c* and cytochrome *c* oxidase. *J. Biol. Inorg. Chem.* **10**, 613–624.
16. Fontayne, A., De Maeyer, B., De Maeyer, M., Yamashita, M., Matsushita, T. & Deckmyn, H. (2007). Paratope and epitope mapping of the antithrombotic antibody 6B4 in complex with platelet glycoprotein Ibalpha. *J. Biol. Chem.* **282**, 23517–23524.
17. Beernink, P. T., Welsch, J. A., Bar-Lev, M., Koeberling, O., Comanducci, M. & Granoff, D. M. (2008). Fine antigenic specificity and cooperative bactericidal activity of monoclonal antibodies directed at the meningococcal vaccine candidate factor H-binding protein. *Infect. Immun.* **76**, 4232–4240.
18. Banci, L., Bertini, I., Cantini, F., Felli, I. C., Gonnelli, L., Hadjiliadis, N. *et al.* (2006). The Atx1-Ccc2 complex is a metal-mediated protein-protein interaction. *Nat. Chem. Biol.* **2**, 367–368.
19. Towbin, H., Staehelin, T. & Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Natl Acad. Sci. USA*, **76**, 4350–4354.
20. Laemmli, U. K., Virji, M., Kayhty, H., Ferguson, D. J., Alexandrescu, C. & Moxon, E. R. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227**, 680–685.
21. Guex, N. & Peitsch, M. C. (1997). SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*, **18**, 2714–2723.
22. James, C., Phillips, R. B., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E. *et al.* (2005). Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781–1802.
23. Laskowski, R. A., MacArthur, M., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291.
24. Humphrey, W. (1996). VMD: visual molecular dynamics. *J. Mol. Graphics*, **14**, 33–38.
25. de Vries, S. J., van Dijk, A. D., Krzeminski, M., van Dijk, M., Thureau, A., Hsu, V. *et al.* (2007). HADDOCK *versus* HADDOCK: new features and performance of HADDOCK2.0 on the CAPRI targets. *Proteins*, **69**, 726–733.
26. Dominguez, C., Boelens, R. & Bonvin, A. M. (2003). HADDOCK: a protein–protein docking approach based on biochemical or biophysical information. *J. Am. Chem. Soc.* **125**, 1731–1733.

Chapter 4

Solution Structure of the Factor H-binding Protein, a Survival Factor and Protective Antigen of *Neisseria meningitidis*

Francesca Cantini[1], Daniele Veggi[2], Sara Dragonetti[1], Silvana Savino[2], Maria Scarselli[2], Giacomo Romagnoli[2], Mariagrazia Pizza[2], Lucia Banci[1], and Rino Rappuoli[2]

[1] Magnetic Resonance Center (CERM), University of Florence, Via L. Sacconi 6, 50019 Sesto Fiorentino, Italy

[2] Novartis Vaccines and Diagnostics, Via Fiorentina 1, 53100 Siena, Italy.

J Biol Chem., (2009)

ACCELERATED PUBLICATION

This paper is available online at www.jbc.org

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 284, NO. 14, pp. 9022–9026, April 3, 2009

Solution Structure of the Factor H-binding Protein, a Survival Factor and Protective Antigen of *Neisseria meningitidis*[*S]

Received for publication, November 12, 2008, and in revised form, December 16, 2008
Published, JBC Papers in Press, February 4, 2009, DOI 10.1074/jbc.C800214200

Francesca Cantini[‡], Daniele Veggi[§], Sara Dragonetti[‡], Silvana Savino[§], Maria Scarselli[§], Giacomo Romagnoli[§], Mariagrazia Pizza[§], Lucia Banci[‡1], and Rino Rappuoli[§2]

From the [‡]Magnetic Resonance Center (CERM), F50019 Sesto Fiorentino and [§]Novartis Vaccines and Diagnostics, 53100 Siena, Italy

Factor H-binding protein is a 27-kDa lipoprotein of *Neisseria meningitidis* discovered while screening the bacterial genome for vaccine candidates. In addition to being an important component of a vaccine against meningococcus in late stage of development, the protein is essential for pathogenesis because it allows the bacterium to survive and grow in human blood by binding the human complement factor H. We recently reported the solution structure of the C-terminal domain of factor H-binding protein, which contains the immunodominant epitopes. In the present study, we report the structure of the full-length molecule, determined by nuclear magnetic resonance spectroscopy. The protein is composed of two independent barrels connected by a short link. Mapping the residues recognized by monoclonal antibodies with bactericidal or factor H binding inhibition properties allowed us to predict the sites involved in the function of the protein. The structure therefore provides the basis for designing improved vaccine molecules.

Neisseria meningitidis, a Gram-negative bacterium that colonizes the upper respiratory tract of 10% of healthy human population, is adapted to grow only in humans. With a frequency of 1 in 100,000 population, the bacterium invades the bloodstream and becomes a severe pathogen, causing sepsis and meningitis. Vaccination with capsular polysaccharides induces serogroup-specific protective antibodies. Meningococcal capsular polysaccharide vaccines are available against serogroups A, C, Y, and W135 (1–3). On the contrary, the development of a vaccine against serogroup B, still responsible for a significant percentage of invasive diseases, has been protracted due to the immunologic cross-reactivity of B polysaccharide with human tissues. Recently, new perspectives to meningococcal B (menB) prevention have been opened by the identification of suitable protein-based vaccine antigens identified by mining the bacterial genome (4). One of the most promising antigens is the factor H-binding protein (fHbp).[3] This is a membrane-anchored lipoprotein (3, 5, 6) that binds human factor H, a negative regulator of the alternative complement activation pathway.(7). Coating the bacterial surface with factor H allows the bacterium to mimic a human tissue and avoid complement-mediated lysis. fHbp is expressed by all the pathogenic strains of *N. meningitidis* and can be classified in three distinct sequence variants (5). This diversity has a remarkable impact on the immunological properties of fHbp given that members of each variant induce a strong protective immunity against meningococcal strains carrying homologous alleles but are ineffective against strains that express distantly related fHbp alleles.

A number of studies using monoclonal antibodies (8–11) have identified residues involved in protective epitopes and factor H binding. Initially, Arg-204 and the cluster Glu-146–Arg-149 were identified as targets of bactericidal monoclonal antibodies elicited by the recombinant fHbp of variant 1 (v.1) (8, 9). The same epitope was later shown to contain also Phe-277, Gly-228, Lys-230, and Glu-233 (10). Recently, Beernink *et al.* (11) used a panel of monoclonal antibodies obtained by immunizing mice with all three variants of fHbp and identified Gly-121 and Lys-122 as critical for binding by anti-v.1 antibodies. Ser-216 was shown to be important for variant 2 (v.2), whereas amino acid positions 174, 180, and 192 were shown to be key residues to discriminate between variant 2 and variant 3 (v.3).

Although the molecular distribution of residues from 141 to 255 was appreciable onto the structure of C terminus domain, whose solution structure was already solved (12), spatial arrangement of Gly-121 and Lys-122 remained so far undetermined. This incompleteness of information hampered a detailed evaluation of the molecular distribution of variant-specific epitopes, as well as the opportunity to rationalize the reported differences in the ability of monoclonals to induce complement-mediated killing and inhibit the protein binding to human factor H (11). In the present study, we determined the structure of the full-length fHbp by NMR, which improves the knowledge about the distribution of protective epitopes on the protein surface and provides useful indications on the possible localization of the factor H (fH) binding site.

EXPERIMENTAL PROCEDURES

Sample Production and NMR Measurements—Recombinant fHbp v.1 (residues Met-7–Gln-255) was expressed in *Escherichia coli* as already described. Analytical gel filtration analysis showed that the recombinant protein was eluted in fractions

* This work was supported by Ministero dell'Istruzione, Università e della Ricerca Scientifica-Fondo per gli Investimenti della Ricerca di Base (MIUR-FIRB) Grant RBLA032ZM7.

The atomic coordinates and structure factors (code 2KC0) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

[S] The on-line version of this article (available at http://www.jbc.org) contains seven supplemental figures three supplemental tables, and supplemental references.

[1] To whom correspondence may be addressed: Magnetic Resonance Center (CERM) – University of Florence, Via L. Sacconi 6, 50019 Sesto Fiorentino, Italy. E-mail: banci@cerm.unifi.it.

[2] To whom correspondence may be addressed: Novartis Vaccines and Diagnostics. Via Fiorentina, 1, 53100 Siena, Italy. Tel.: 39-0577-243414; Fax: 39-0577-243564; E-mail: rino.rappuoli@novartis.com.

[3] The abbreviations used are: fHbp, factor H-binding protein; RDC, residual dipolar coupling; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser effect spectroscopy; HSQC, heteronuclear single quantum correlation; v.1, variant 1; v.2, variant 2; v.3, variant 3.

corresponding to a monomeric state of the molecule. Electrospray mass ionization-mass spectrometry spectrum indicated a mass of 27281.43 Da, which corresponded to the cloned construct. NMR spectra were acquired at 298 K on Avance 900, 700, and 500 MHz Bruker spectrometers, all equipped with a triple resonance cryoprobe. The NMR experiments, used for the backbone and the aliphatic side chain resonances assignment recorded on $^{2}H/^{13}C/^{15}N$, $^{13}C/^{15}N$, and ^{15}N enriched and on unlabeled fHbp samples, are summarized in supplemental Table S1. The ^{1}H, ^{13}C, and ^{15}N resonance assignments of fHbp are reported in supplemental Table S2. All the amide protons of the fHbp protein were assigned, with the only exceptions of Val-8, His-26, Gly-202, and Gly-229. The assignment of the aromatic spin systems was performed with two-dimensional NOESY and total correlation spectroscopy maps acquired on the sample dialyzed against deuterated buffer. Backbone dihedral angle constraints were derived from ^{15}N,$^{13}C'$,$^{13}C\alpha$,$^{13}C\beta$, and Ha chemical shifts, using TALOS (13). Standard errors of values predicted by TALOS were used as allowed ranges of variations in the dihedral angle constraints.

Distance constraints for structure determination were obtained from ^{15}N-edited and ^{13}C-edited three-dimensional NOESY-HSQC. Residual dipolar couplings have been measured in the presence of an external orienting medium constituted by a binary mixture of C12E5 (penta-ethylene glycol dodecyl ether, > 98% purity, Fluka) and neat *n*-hexanol (puriss., Fluka), which form a stable liquid crystalline phase made of neutral aggregates (called bicelles) in the temperature range from 298 to 312 K (14). The molar ratio of C12E5 to *n*-hexanol was 0.96, and the C12E5/water ratio was 7% weight. One-bond ^{1}H-^{15}N coupling constants were measured at 298 K and 800 MHz by using the inphase anaphase (IPAP) method (15). A total of 100 RDC values have been measured from amide signals not overlapped in the HSQC spectrum. Out of them, 73 RDC values, derived from residues not experiencing R_1 and/or R_2 values outside the average values, were used for structure calculations.

2987 meaningful proton-proton distance restraints (supplemental Fig. S4), together with 158 ϕ and 158 ψ backbone dihedral angles restraints, were included in structure calculations. The exchangeability of the backbone amide hydrogens with solvent protons was investigated through an ^{1}H-^{15}N HSQC experiment performed on the protein previously extensively dialyzed against deuterated buffer. Hydrogen bond constraints for the slowly deuterium-exchanging amide protons of the β-strands were introduced at later steps of structure calculations.

Structure calculations were performed using the program CYANA-2.1 (16). 900 random conformers were annealed in 13,000 steps. The χ tensor parameters were obtained with FANTAORIENT, and they were optimized through iterative cycles of PSEUDOCYANA until convergence (17).

The family of the best 30 structures in terms of target function was then subjected to restrained energy minimization with the AMBER 10.0 package in explicit water solvent (18). The distance and torsion angle constraints and the RDCs were applied with force constants of 50 kcal $mol^{-1}Å^{-2}$ and 32 kcal mol^{-1} $radians^{-2}$, respectively. The conformational and energetic analysis together with selected quality parameters from PROCHECK-NMR, WHATIF (19, 20) analysis, and QUEEN program (21) of the family of the best 30 structures are reported in supplemental Table S3. The program MOLMOL was subsequently used for structure analysis (22). The root mean square deviation values per residue of the restrained energy minimization (REM) family of 30 conformers for the segment 14–255 are reported in supplemental Fig. S5.

^{15}N R_1, R_2, and steady-state heteronuclear NOEs were measured with pulse sequences as described by Farrow *et al.* (23) R_2 values were measured using a refocusing time of 450 μs. In all experiments, the water signal was suppressed with the "water flip-back" scheme. Average R_1, R_2, and ^{1}H-^{15}N NOE values of 0.91 ± 0.04 s^{-1}, 24.2 ± 0.2 s^{-1}, and 0.70 ± 0.03 are found, respectively, at 500 MHz. The experimental relaxation rates were used to map the spectral density function values, $J(\omega_H)$, $J(\omega_N)$, J(0) following a procedure available in literature (24, 25).

RESULTS

The far-UV CD spectrum of fHbp had features characteristic of a folded protein with a high content of β-strands, whereas the negative band at 198 nm, characteristic of random coil conformation, was not present (data not shown). Consistently, the ^{1}H-^{15}N HSQC spectra show well dispersed resonances indicative of an overall well folded protein (supplemental Fig. S1).

Heteronuclear relaxation rates, whose values are affected by the dynamic properties of the molecule, were essentially homogeneous along the entire polypeptide sequence, with the exception of residues located at the C and N termini and some loops. The correlation time for the molecule tumbling (τ_c), as estimated from the R_2/R_1 ratio, is 20.3 ± 1.5 ns, consistent with the molecular weight of the protein being in the monomeric state.

The solution structure of the full-length fHbp consisted of two domains, fHbpN (residues 8–136) and fHbpC (residues 141–255), composed of 10 and 8 antiparallel β-strands, respectively. fHbpN and fHbpC were connected by a five-residue linker and had a topology of secondary structure elements as shown in Fig. 1*A*, with an overall three-dimensional structure as shown in Fig, 1B. Both domains are structurally well defined, with only few loops showing some conformational disorder (supplemental Fig. S2). From heteronuclear relaxation measurements (supplemental Fig. S3), it appears that most of these loop regions, and in particular, those including residues 85–90 and 118–123 located in the N-terminal domain, have significantly lower heteronuclear NOE values as a consequence of local motions in the ns–ps time scale. Consistently, they were characterized by a very low number of long range ^{1}H-^{1}H NOEs (supplemental Fig. S4), and consequently, their conformation was less defined (supplemental Fig. S2).

The fHbpN domain has an elongated barrel-like structure with a length of about 49 Å and a width of 22 Å. It is organized in six antiparallel β-strands of different lengths (β_2(45–50), β_3(53–58), β_5(72–84), β_6(87–100), β_7(104–111), β_8(128–136)), forming an extended β-sheet, and two short β-strands (β_1(33–35) and β_4(62–64)), which face the β-sheet and a short N-terminal α-helix (16–21). In 50% of the family conformers, a short β-hairpin, including residues 114–116 (β_7') and 123–125

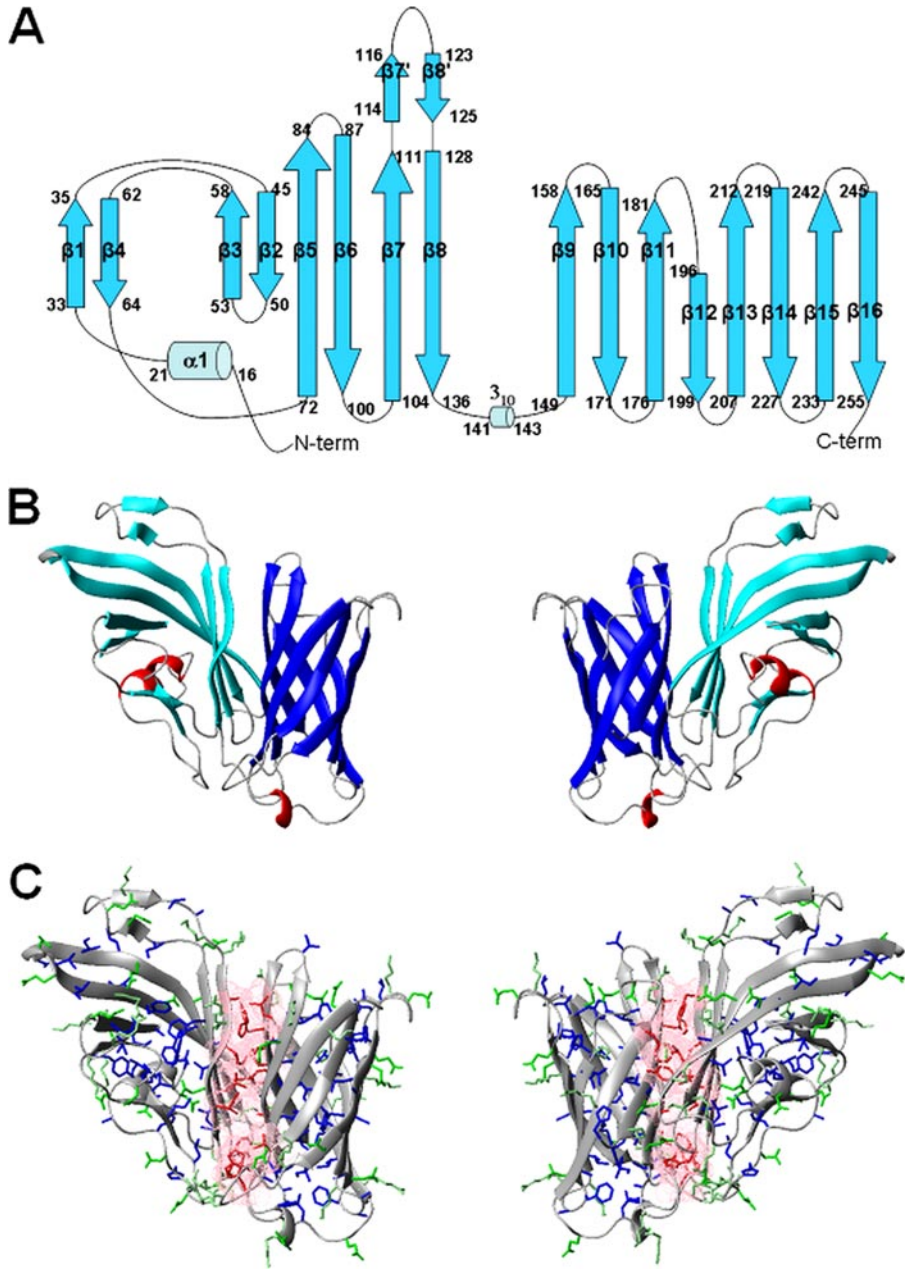


FIGURE 1. *A*, topology diagram of the fHbp protein. The α-helices are represented by *sky blue cylinders*, and the β-strands are *cyan arrows*. *N-term*, N terminus; *C-term*, C terminus. *B*, ribbon diagram of fHbp. Secondary structure elements are shown. β-strands of the N-terminal domain are shown in *cyan* and helices are shown in *red*, whereas β-strands of fHbpC are shown in *blue*. *C*, the side chains of the hydrophobic residues involved in interdomains contacts are shown as *sticks* in *red*, and *pink*. Contact surfaces are also reported. The other hydrophobic residues are shown in *blue*. The charged residues are shown in *green*. The backbone is shown as a *ribbon*.

(β_8'), is observed between β7 and β8 (Fig. 1*A*), whose presence is supported by the chemical shift index analysis.

The core of fHbpN is characterized by several hydrophobic interactions that involve aromatic and aliphatic residues. In particular, hydrophobic contacts are present between the residues clustered on one side of the β-sheet (Leu-46, Leu-48 (β_2), Ala-53 and Tyr-57 (β_3), Phe-76, Phe-78 (β_5), Phe-96 and Val-98 (β_6), Ala-105 and Phe-109 (β_7), Ala-135(β_8) and those on the two other β-strands β_1 (Leu-34 and Leu-36) and β_4 (Leu-63), determining the spatial proximity of the two regions and the overall shape of the N-terminal domain. Further long range interactions between residues Leu-34, Thr-35, and Asp-133 and Ala-135 maintain β_1 close to β_8, stabilizing the overall structure. The three-dimensional arrangement of the β-strands defines a hydrophobic core, which is closed on one side by helix α_1. Contacts between some residues located in β_3, β_4, and β_5 and those in helix α_1 are present that anchor such a helix to the rest of the protein. The latter interactions define the conformation of a long loop (segment 22–32), which is completely solvent-exposed and rich in charged residues. The latter loop, although stabilized by the contacts of two hydrophobic residues (Leu-24 and Leu-31) with the rest of the protein, experiences high conformational disorder. Also, the region between β_5 and β_6, opposite to the N-terminal helix, is solvent-exposed despite the fact that it is rich in hydrophobic residues (Fig. 1*C*).

The C-terminal domain fHbpC is constituted by an eight-stranded antiparallel β-barrel (141–255), whose strands are connected by loops of variable lengths. The same arrangement was previously found in the solution structure of the same domain in the isolated fHbpC domain (12). The 3_{10}-helix present in the solution structure of isolated fHbpC domainis was still present in the majority of 30 of the full-length conformers, although shorter (141–143) with respect to that observed the isolated fHbpC (138–141). Sizable variations for the backbone NH chemical shifts of fHbpC domain in the two constructs (>0.25 ppm as combined chemical shifts) were observed on β_7, β_8, β_9, β_{14}, β_{15}, and β_{16} (supplemental Fig. S6). The largest differences were observed for residues located in β_7, β_8, which formed a long disordered tail in the isolated fHbpC domain. The other four β-strands of the C-terminal domain are located on the one side of the β-barrel and define the domain-domain interface. Interestingly, residues located in β_9, β_{14}, β_{15}, and β_{16} formed a hydrophobic patch in fHbpC molecule. The present structure confirms our previous insights on the role of this region of the β-barrel (12).

The domain-domain interface, formed by about 850 Å^2 of buried surface, is stabilized by intermolecular contacts involving exclusively hydrophobic amino acids (Leu-106, Tyr-99, Phe-129, Ile-134, N-terminal domain; Thr-155, Phe-157, Phe-

227, Leu-251, C-terminal domain, Fig. 1*C*). About 70 long range interdomain NOEs were found that defined, together with the RDCs, the reciprocal orientation of the domains.

DISCUSSION

Human complement is the first barrier of innate immunity and kills infectious agents when they try to invade the bloodstream (26). To do this, it attacks anything that is not self. Factor H and C4b are complement components that allow the system to distinguish between invasive agents and self-tissues by coating the latter and avoiding the complement attack. During the evolutionary battle for survival, virtually all pathogens have found multiple and redundant ways to escape the complement-mediated killing (27). One of the most popular strategies among pathogens is to avoid complement by mimicking host surfaces, covering themselves with factor H. fHbp of *N. meningitidis* is a typical example of this class of molecules. By binding human factor H, this protein allows the bacterium to survive and grow in human blood and cause a terrible disease. This property, together with the ability to induce in humans a strong antibody response, makes this protein an attractive vaccine antigen able to induce antibodies with a double function: killing the bacterium by direct activation of the classical complement cascade and preventing the formation of fH-fHbp complex on the surface of the bacterium.

Based on the knowledge of the complete structure of fHbp, we can identify the location of the variable amino acids (Fig. 2*A*), the residues involved in binding protective monoclonal antibodies (Fig. 2*B*), the region bound by antibodies against variant 1 (Fig. 2*C*), the region bound by antibodies against variant 2 (Fig. 2*D*), and the residues recognized by the antibodies that inhibit factor H binding (Fig. 2*E*). The variable residues (Fig. 2*A*) cluster in the upper part of the molecule. This is consistent with a model where the protein is anchored to the bacterial cell wall through the palmitic acid extension of Cys-1 (5) and exposes the upper part to the outside, where it is under the selective pressure of the immune system. Accordingly, the amino acids known to be part of epitopes (8, 9, 10, 11) are all localized in correspondence of the zone of higher variability (Fig. 2*B*). Arg-204, Gly-121, and the loop formed by Glu-146 until Arg-149 were identified as involved in the formation of bactericidal v.1 epitopes (8, 9, 11). Arg-204 is located in the loop between β_{12} and β_{13}, whereas the segment 146–149 corresponds to the region connecting the two fHbp domains and containing the second short helix α2. Gly-121, which was localized in the disordered region of the isolated fHbpC domain, now occupies the loop between $\beta7'$ and $\beta8'$ strands. NMR-driven epitope mapping carried out on the isolated fHbpC domain identified Phe-141, Lys-199, Arg-204, Glu-146–Arg-149, Phe-227, Lys-230, and Glu-233 as part of the same epitope recognized by the bactericidal monoclonal antibody Mab502 elicited by fFbp v.1 (10). The distribution of these residues on the full-length protein supports the hypothesis that the fHbpC domain contains the major part of the native epitope, which consists of all those residues still accessible in the whole fHbp with the exception of Phe-227 and Glu-233, which are now shielded by the presence of the N-terminal domain. Epitopes of v.2 and v.3 have recently been shown to include residues in

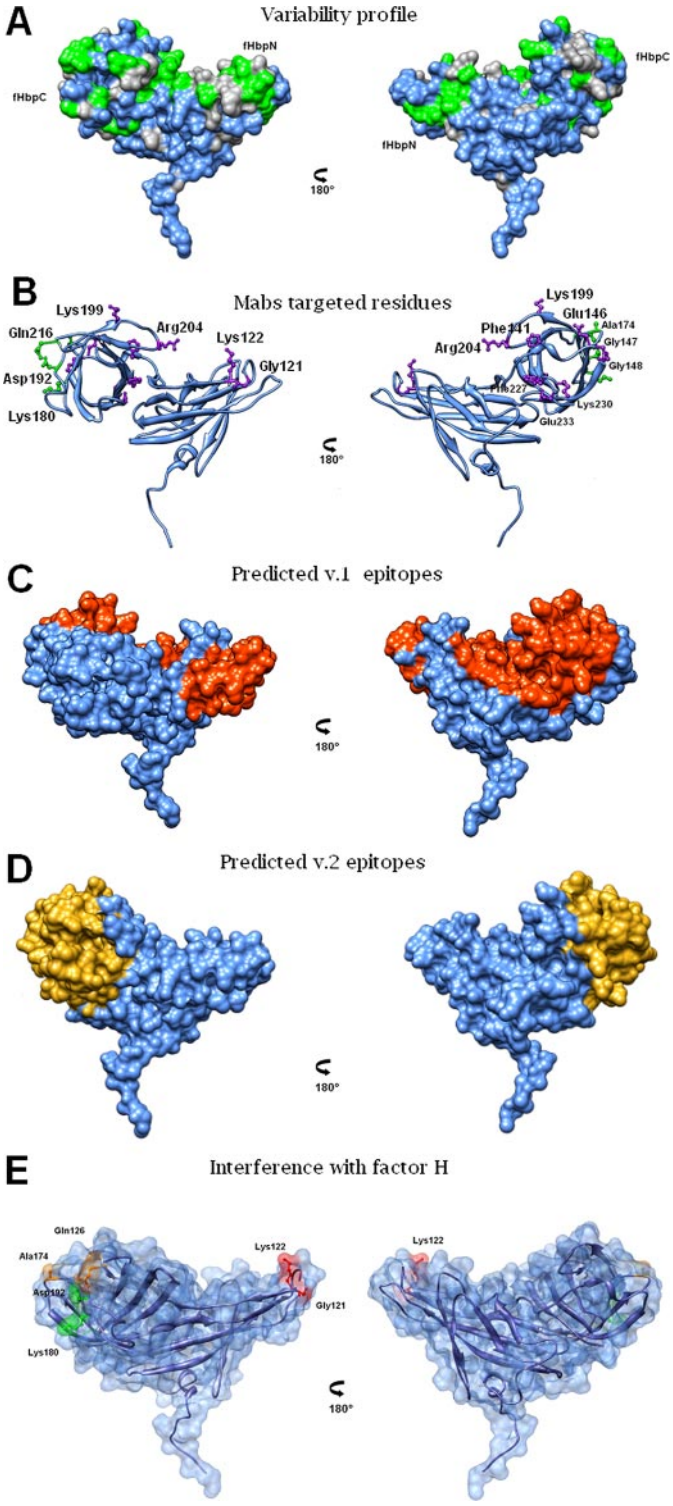


FIGURE 2. *A*, variability profile of the three fHbp variants. Conserved amino acids are colored in *blue*, conservative substitutions are colored in *gray*, and variable amino acids are colored in *green*. *B*, distribution of amino acids recognized by monoclonal antibodies (*Mabs*) raised against v.1 (*purple*) and v.2/v.3 (*green*). *C* and *D*, molecular areas recognized by the bactericidal pairs of monoclonal antibodies against v.1 (*orange*) and v.2 (*gold*) (11). *E*, molecular distribution of fHbp residues recognized by monoclonal antibodies inhibiting (*red*), partially inhibiting (*orange*), and not inhibiting (*green*) the fH binding (11).

positions 174, 216, 180, and 192 (11). Although none of the corresponding monoclonal antibodies were bactericidal individually, nevertheless some pairs of them were able to perform

meningococcal killing when combined. The same was true for a pair of monoclonal antibodies raised against v.1 (11). We tried to identify the fHbp regions targeted by these pairs of bactericidal monoclonal antibodies, placing residues recognized by each component at the center of a molecular area of 950 A^2, a value chosen as the representative value of the average size for a protein conformational epitope (28–31). The interesting observation that the bactericidal epitopes crucial for the killing of v.1 and v.2 are only partially overlapping suggests that it may be possible to engineer molecules containing epitopes of all variant proteins (Fig. 2, *C* and *D*). Finally, Fig. 2*E* shows that the residues recognized by the monoclonal antibodies that inhibit factor H binding (11) are located In the upper part of both in the N-terminal and in the C-terminal part of the molecule, suggesting that this region is involved in the interaction with factor H.

In conclusion, the complete structure of fHpb allowed us to map residues involved in important functions, such as binding to factor H and eliciting bactericidal antibodies, and provides the basis to understand the molecular mechanisms behind the function of this important molecule. The structure may be used to design improved vaccine antigens.

REFERENCES

1. Borrow, R., Andrews, N., Goldblatt, D., and Miller, E. (2001) *Infect. Immun.* **69,** 1568–1573
2. Girard, M. P., Preziosi, M. P., Aguado, M. T., and Kieny, M. P. (2006) *Vaccine* **24,** 4692–4700
3. Snape, M. D., Perrett, K. P., Ford, K. J., John, T. M., Pace, D., Yu, L.-M., Langley, J. M., McNeil, S., Dull, P. M., Ceddia, F., Anemona, A., Halperin, S. A., Dobson, S., and Pollard, A. J. (2008) *J. Am. Med. Assoc.* **299,** 173–184
4. Giuliani, M. M., Adu-Bobie, J., Comanducci, M., Aricò, B., Savino, S., Santini, L., Brunelli, B., Bambini, S., Biolchi, A., Capecchi, B., Cartocci, E., Ciucchi, L., Di Marcello, F., Ferlicca, F., Galli, B., Luzzi, E., Masignani, V., Serruto, D., Veggi, D., Contorni, M., Morandi, M., Bartalesi, A., Cinotti, V., Mannucci, D., Titta, F., Ovidi, E., Welsch, J. A., Granoff, D., Rappuoli, R., and Pizza, M. (2006) *Proc. Natl. Acad. Sci.* **103,** 10834–10839
5. Masignani, V., Comanducci, M., Giuliani, M. M., Bambini, S., Adu-Bobie, J., Aricò, B., Brunelli, B., Pieri, A., Santini, L., Savino, S., Serruto, D., Litt, D., Kroll, S., Welsch, J. A., Granoff, D. M., Rappuoli, R., and Pizza, M. (2003) *J. Exp. Med.* **197,** 789–799
6. Pizza, M., Scarlato, V., Masignani, V., Giuliani, M. M., Aricò, B., Comanducci, M., Jennings, G. T., Baldi, L., Bartolini, E., Capecchi, B., Galeotti, C. L., 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. e., Hood, D. W., Jeffries, A. C., Saunders, N. J., Granoff, D. M., Venter, J. C., Moxon, E. R., Grandi, G., and Rappuoli, R. (2000) *Science* **287,** 1816–1820
7. Madico, G., Welsch, J. A., Lewis, L. A., McNaughton, A., Perlman, D. H., Costello, C. E., Ngampasutadol, J., Vogel, U., Granoff, D. M., and Ram, S. (2006) *J. Immunol.* **177,** 501–510
8. Welsch, J. A., Rossi, R., Comanducci, M., and Granoff, D. M. (2004) *J. Immunol.* **172,** 5606–5615
9. Giuliani, M. M., Santini, L., Brunelli, B., Biolchi, A., Arico, B., Di Marcello, F., Cartocci, E., Comanducci, M., Masignani, V., Lozzi, L., Savino, S., Scarselli, M., Rappuoli, R., and Pizza, M. (2005) *Infect. Immun.* **73,** 1151–1160
10. Scarselli, M., Cantini, F., Santini, L., Veggi, D., Dragonetti, S., Donati, C., Savino, S., Giuliani, M. M., Comanducci, M., Di Marcello, F., Romagnoli, G., Pizza, M., Banci, L., and Rappuoli, R. (2008) *J. Mol. Biol.* **386,** 97–108
11. Beernink, P. T., Welsch, J. A., Bar-Lev, M., Koeberling, O., Comanducci, M., and Granoff, D. M. (2008) *Infect. Immun.* **76,** 4232–4240
12. Cantini, F., Savino, S., Scarselli, M., Masignani, V., Pizza, M., Romagnoli, G., Swennen, E., Veggi, D., Banci, L., and Rappuoli, R. (2006) *J. Biol. Chem.* **281,** 7220–7227
13. Cornilescu, G., Delaglio, F., and Bax, A. (1999) *J. Biomol. NMR* **13,** 289–302
14. Ruckert, M., and Otting, G. (2000) *J. Am. Chem. Soc.* **122,** 7793–7797
15. Ottiger, M., Delaglio, F., and Bax, A. (1998) *J. Magn. Reson.* **131,** 373–378
16. Guntert, P. (2000) *Methods Mol. Biol.* **278,** 353–378
17. Banci, L., Bertini, I., Huber, J. G., Luchinat, C., and Rosato, A. (1998) *J. Am. Chem. Soc.* **120,** 12903–12909
18. Perlman, D. A., Case, Caldwell, J. W., Ross, W. S., Cheatham, T. C., DeBolt, S. E., Ferguson, D. M., Seibel, G. L., and Kollman, P. A. (1995) *Comput. Phys. Commun.* **91,** 1–41
19. Laskowski, R., Rullmann, J., McArthur, M., Kaptein, R., and Thornton, J. (1996) *J. Biomol. NMR* **8,** 477–486
20. Vriend, G. (1990) *J. Mol. Graph.* **8,** 52–56
21. Nabuurs, S., Spronk, C., Krieger, E., Maassen, H., Vriend, G., and Vuister, G. W. (2003) *J. Am. Chem. Soc.* **125,** 12026–12034
22. Koradi, R., Billeter, M., and Wüthrich, K. (1996) **14,** 29–32
23. Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D., and Kay, L. E. (1994) *Biochemistry* **33,** 5984–6003
24. Grzesiek, S., and Bax, A. (1993) *J. Biomol. NMR* **3,** 185–204
25. Peng, J. (1992) *J. Magn. Reson.* **98,** 308–332
26. Rittirsch, D., Flierl, M. A., and Ward, P. A. (2008) *Nat. Rev. Immunol.* **8,** 776–787
27. Lambris, J. D., Ricklin, D., and Geisbrecht, B. V. (2008) *Nat. Rev. Microbiol.* **6,** 132–142
28. Davies, D. R., Padlan, E. A., and Sheriff, S. (1990) *Annu. Rev. Biochem.* **59,** 439–473
29. Farady, C. J., Egea, P. F., Schneider, E. L., Darragh, M. R., and Craik, C. S. (2008) *J. Mol. Biol.* **380,** 351–360
30. Lok, S. M., Kostyuchenko, V., Nybakken, G. E., Holdaway, H. A., Battisti, A. J., Sukupolvi-Petty, S., Sedlak, D., Fremont, D. H., Chipman, P. R., Roehrig, J. T., Diamond, M. S., Kuhn, R. J., and Rossmann, M. G. (2008) *Nat. Struct. Mol. Biol.* **15,** 312–317
31. Ménez, R., Michel, S., Muller, B. H., Bossus, M., Ducancel, F., Jolivet-Reynaud, C., and Stura, E. A. (2008) *J. Mol. Biol.* **376,** 1021–1033

Chapter 5

Investigation on the interaction between factor H binding protein and domains 5, 6 and 7 of human factor H

manuscript in preparation

5.1 Abstract

The factor H binding protein (fHbp), a 27-kDa membrane-anchored lipoprotein of *Neisseria meningitidis*, is the major component of a universal vaccine against meningococcus B. Moreover the protein is essential for pathogenesis because it allows the bacterium to survive and grow in human blood by binding the human complement factor H (hfH).

In this study, we used nuclear magnetic resonance spectroscopy to investigate the interaction between fHbp and a construct of the factor H comprising domains 5, 6 and 7 (hfH567).

The data show that the hfH567 binds an extensive region of fHbp which involved both N- and C-domains of the protein and that residues previously targeted as essential for the binding with MAb502 are not affected by the presence of hfH.

5.2 Introduction

Neisseria meningitidis, a Gram-negative bacterium that colonizes the upper respiratory tract of 10% of healthy human population, is adapted to grow only in humans. With a frequency of 1 in 100,000 population, the bacterium invades the bloodstream and becomes a severe pathogen, causing sepsis and meningitis. Vaccination with capsular polysaccharides induces serogroup-specific protective antibodies.

Meningococcal capsular polysaccharide vaccines are available against serogroups A, C, Y, and W135 (1, 2, 3). On the contrary, the development of a vaccine against serogroup B, still responsible for a significant percentage of invasive diseases, has been protracted due to the immunologic cross-reactivity of B polysaccharide with human tissues. Recently, new perspectives to meningococcal B (menB) prevention have been opened by the identification of suitable protein-based vaccine antigens identified by mining the bacterial genome (4).

One of the most promising antigens is the factor H-binding protein (fHbp), whose solution structure was recently determined by NMR spectroscopy (5).

This is a 27kDa membrane-anchored lipoprotein (6, 7) present on the surface of all strains of *N. meningitidis* and elicits protective bactericidal antibodies. fHbp binds human factor H (hfH), a negative regulator of the alternative complement activation pathway (8). The bacterium is able to cover its surface with factor H allowing mimicking a human tissue and avoiding complement-mediated lysis.

Factor H is a single polypeptide chain plasma glycoprotein (155 kDa), composed of 20 repetitive units of 60 amino acids (9), named short consensus repeats (SCR) or complement control protein modules (CCP). The 3D structure of several SCRs have been determined, showing a globular structure with six-stranded antiparallel β-sheets connected with loops and turns (10, 11, 12).

Recently Tang and co-workers solved the x-ray structure of the complex between fHbp and a construct of hfH including domains 6 and 7 (hfH67) (13); this study pointed out the crucial role of the sixth domain for the interaction. In particular they defined two residues on fHbp (Glu 218 and Glu 239) located on the sixth domain as absolutely required for the binding.

In this work we investigated through NMR measurement on the interaction between fHbp and the factor H construct comprising domains 5, 6 and 7 (hfH567). Residues involved in the interaction have been mapped by NMR titrations with the hfH567 on fHbp.

The data showed that hfH567 recognizes an extended area of fHbp, including both N- and C-domains.

5.3 Methods

The interaction of hfH567 with ^{15}N-labelled fHbp was investigated by NMR spectroscopy with ^{1}H–^{15}N TROSY-HSQC experiments. All the NMR experiments were performed at 298 K on a Bruker Avance 900 spectrometer, working at 900.13MHz frequency and equipped with a cryogenically cooled probe.

Titrations were performed on 0.2 mM ^{15}N-labelled fHbp protein samples with the unlabelled hfH567 up to an fHbp–hfH567 molar ratio of 1:1. ^{1}H and ^{15}N resonances assignment for the fHbp protein was already available (5). All the samples used were in 20 mM PBS (90% H_2O and 10% D_2O) at pH 7.0.

Solvent accessibility for fHbp residues was calculated with the program NACCESS (14); only residues with a global solvent accessibility larger than 40% were considered for the mapping.

5.4 NMR mapping of the hfH567-fHbp interaction

The interaction between fHbp and a construct of the human factor H comprising domains 5, 6 and 7 (hfH567) was investigated through NMR spectroscopy, by analyzing the perturbations caused in the ^{1}H–^{15}N TROSY-HSQC spectra of ^{15}N-labelled fHbp upon addition of unlabelled hfH567 up to a fHbp:hfH567 ratio of 1:1.

Chemical shift variation of NH cross peaks observed in the ^{1}H–^{15}N TROSY-HSQC spectra upon addition of hfH567 indicated that the two proteins are interacting.

The fHbp-hfH567 complex exchanged with the free proteins at rates slower than the resonance frequency differences between the two forms (i.e., in the range of milliseconds).

Residues of fHbp experiencing different chemical shift upon addition of hfH567 are shown in Figure 1A and listed in Table 1. They comprise an extensive region which involves both the N- and C-terminal domains of fHbp. Notably, surface-exposed residues (magenta in Figure 1), located in the linker connecting the N- and C-domains of fHbp (Thr139, Ser140, Phe141, Asp142 and Lys143) and several buried residues located at the domain-domain interface of fHbp (Gln97, Tyr99, Gln101, His103, Phe129, Gly132, Ala135, Ile226, Gly236, Ser237, His248, Ile249, Gly250 and Leu251 shown in red in Figure 1) result perturbed by the addition of hfH567, suggesting that a

molecular rearrangement of fHbp occurs during the formation of the fHbp-hfH567 complex.

Previous studies have shown that residues Gly121 and Lys122 are recognized by antibodies that inhibit hfH binding as JAR3 and JAR5 (15). These residues are solvent-exposed and they are located in the loop between β7 and β8 of the N-terminal domain of the molecule. In agreement with ELISA data, spectral perturbations were observed for their NH cross peaks upon addition of hfH567 to fHbp. We have instead recently demonstrated by FACS experiments (16), that the hfH-hfHbp interaction cannot be hindered by the monoclonal antibody MAb502 binding to fHbp. Residues involved in the interaction with MAb502 are Arg204, Glu146-Arg149, Ala174, Lys230 and Glu233. The latter residues are indeed not affected by hfH567 binding confirming that hfH567 interacts in a different region with respect to MAb502.

Recently, the crystal structure of the complex of fHbp with a construct of human factor H including domains 6 and 7 (hfH67) was determined (13). The hfH-fHbp complex is held together by extensive interactions between both the N- and C-terminal domains of fHbp and the domain 6 of hfH, while fewer contacts were found between fHbp and domain 7 of hfH. In particular two residues (Glu218 and Glu239, located on β14 and β15 of fHbp, respectively) were found to be essential for the interaction with the domain 6 of hfH67.

The total number of perturbed, surface-exposed residues in solution define a larger contact area with respect to that found in the crystal, but still contains all the residues located at the interface between hfH67 and fHbp, indicating that also in solution and in the presence of a longer construct of hfH, the same interacting surface of the fHbp is maintained (Figure1).

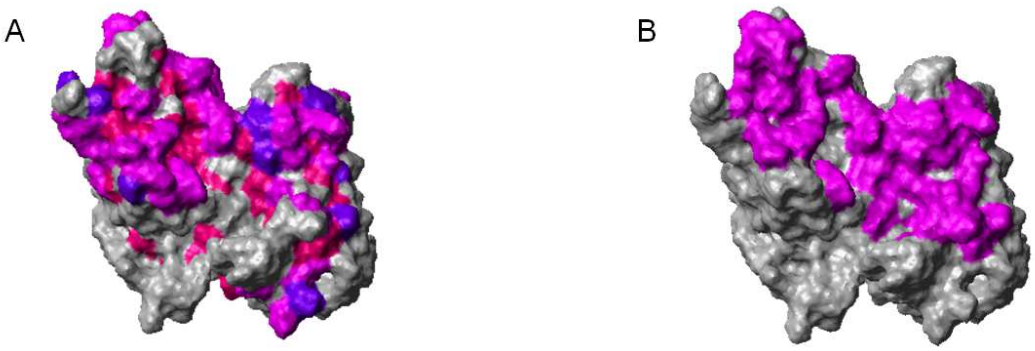
The higher number of residues perturbed by the presence of hfH567 with respect to the crystal complex with hfH67 can be justified by a model of interaction for fHbp-hfH567 complex in which the reciprocal orientation of fHbp N- and C-domains changed if compared with the structure of the free fHbp and

in which domain 5 of hfH, absent in the crystal, made contacts with residues Ile198, Lys 199 and Asp142.

Table 1. Amino acids of fHbp whose amide signals experience different chemical shift (> 0.1 ppm) upon the addition of hfH567. B=buried; E=exposed as it results from the solvent accessibility calculated with the program NACCESS (14).

Residues	
B	Leu31,Gln32,Leu36,Ser39,Val40,Glu44,Tyr57,Asn64,Ser74,Phe76,Phe78,Arg80 Gln81,Ile82,Ile89,Ser93,Phe96,Gln97,Val98,Tyr99,Gln101,His103,Thr107,Phe109, Gln110,Thr111,Ile114,Asp116,Phe129,Gly132,Ala135,Tyr152,Lys165,Gly177, Gly179,Ala196,Ile198,Ala206,Val207,Gly210,Val212,Gly220,Leu224,Gly225,Ile226, Gly236,Ser237,Ala238, Thr242,His248,Ile249,Gly250,Leu251
E	Asp37,Gln38,Arg41,Lys42,Asn43,Lys45,Thr56,Glu83,Gly86,Leu91,Glu95,Glu112, His119,Gly121,Lys122,Val124,Lys126,Arg127,Gln128,Arg130,Thr139,Ser140,Phe1 41,Asp142,Lys143,Asp160,Gly163,Glu188,Lys199,Ser211,Leu213,Lys219,Ser221, Ser223,Lys241

Figure 1. (A) Residues experiencing chemical shift differences mapped on the molecular surface of the free fHbp. Surface-exposed and buried residues, whose amide NMR signals were perturbated (> 0.1 ppm) upon addition of hfH567, are shown in magenta and red, respectively. Residues whose amide signals were not perturbated upon addition of hfH567, prolines, and residues that could not be accurately analyzed due to spectral overlap, are in gray. (B) Residues mapped on the molecular surface of the free fHbp (PDBID 2KC0) and located at the interface of the fHbp-hfH67 complex in the crystal structure are shown in magenta.



References

1. Borrow, R., Andrews, N., Goldblatt, D. & Miller, E., 2001, *Infect. Immun.* 69, 1568-1573.

2. Girard, M. P., Preziosi, M.-P., Aguado, M.-T. & Kieny, M. P., 2006, *Vaccine* 24, 4692-4700.

3. Snape, M. D., Perrett, K. P., Ford, K. J., John, T. M., Pace, D., Yu, L.-M., Langley, J. M., McNeil, S., Dull, P. M., Ceddia, F., Anemona, A., Halperin, S. A., Dobson, S. & Pollard, A. J., 2008, *JAMA*, 299, 173-184.

4. Giuliani, M. M., Adu-Bobie, J., Comanducci, M., AricÃ², B., Savino, S., Santini, L., Brunelli, B., Bambini, S., Biolchi, A., Capecchi, B., Cartocci, E., Ciucchi, L., Di Marcello, F., Ferlicca, F., Galli, B., Luzzi, E., Masignani, V., Serruto, D., Veggi, D., Contorni, M., Morandi, M., Bartalesi, A., Cinotti, V., Mannucci, D., Titta, F., Ovidi, E., Welsch, J. A., Granoff, D., Rappuoli, R. & Pizza, M., 2006, A universal vaccine for serogroup B meningococcus. *Proceedings of the National Academy of Sciences,* 103, 10834-10839.

5. Cantini, F., Veggi, D., Dragonetti, S., Savino, S., Scarselli, M., Romagnoli, G., Pizza, M., Banci, L., and Rappuoli, R., 2009, *J Biol Chem.,* 284, 9022-6.

6. Masignani, V., Comanducci, M., Giuliani, M. M., Bambini, S., Adu-Bobie, J., Arico`, B., Brunelli, B., Pieri, A., Santini, L., Savino, S., Serruto, D., Litt, D., Kroll, S., Welsch, J. A., Granoff, D. M., Rappuoli, R., and Pizza, M., 2003, *J. Exp. Med.,* 197, 789–799.

7. Pizza, M., Scarlato, V., Masignani, V., Giuliani, M. M., Arico`, B., Comanducci, M., Jennings, G. T., Baldi, L., Bartolini, E., Capecchi, B., Galeotti, C. L., 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. e., Hood, D. W., Jeffries, A. C., Saunders, N. J., Granoff, D. M., Venter, J. C., Moxon, E. R., Grandi, G., and Rappuoli, R., 2000, *Science,* 287, 1816–1820.

8. Madico, G., Welsch, J. A., Lewis, L. A., McNaughton, A., Perlman, D. H., Costello, C. E., Ngampasutadol, J., Vogel, U., Granoff, D. M., and Ram, S., 2006, *J. Immunol.,* 177, 501–510.

9. The complete amino acid sequence of human complement factor H. Ripoche, J., Day, A.J., Harris, T.J.R., Sim, R.B., 1988*, Biochem. J*., 249, 593–602.

10. Perkins, S.J., Haris, P.I., Sim, R.B., Chapman, D., 1988, Biochemistry, 27, 4004–4012.

11. Barlow, P.N., Norman, D.G., Steinkasserer, A., Horne, T.J., Pearce, J., Driscoll, P.C., Sim, R.B., Campbell, I.D., 1992, *Biochemistry*, 31, 3626–3634.

12. Barlow, P.N., Steinkasserer, A., Norman, D.G., Kieffer, B., Wiles, A.P., Sim, R.B., Campbell, I.D., 1993, *J. Mol. Biol*., 232, 268–284.

13. Schneider MC, Prosser BE, Caesar JJ, Kugelberg E, Li S, Zhang Q, Quoraishi S, Lovett JE, Deane JE, Sim RB, Roversi P, Johnson S, Tang CM, Lea SM., 2009, *Nature*, 458 (7240), 890-3.

14. Hubbard, S. J. & hornton, J. M. ,1993, NACCESS. Department of Biochemistry and Molecular Biology, University College London, London, UK.

15. Beernink PT, LoPasso C, Angiolillo A, Felici F, Granoff D., 2009, *Mol Immunol*., 46, 1647-53.

16. Scarselli M, Cantini F, Santini L, Veggi D, Dragonetti S, Donati C, Savino S, Giuliani MM, Comanducci M, Di Marcello F, Romagnoli G, Pizza M, Banci L, Rappuoli R., *J Mol Biol.*, 2009, 386(1), 97-108

Chapter 6

Interaction between fHbp and the monoclonal Antibody JAR4

in preparation

6.1 Abstract

Antibodies to factor H (fH) binding protein (fHbp), a meningococcal vaccine antigen, activate classical complement pathway serum bactericidal activity (SBA) and block binding of the complement inhibitor fH. To understand these 2 functions in protection, it results very interesting to investigate the interactions of human complement and anti-fHbp monoclonal antibodies (mAbs) with encapsulated *Neisseria meningitidis* (1). In this work we focused on the study, through NMR measurements, of the interaction between fHbp and mAb JAR4, which cross-reacted with both variant 1 and 2 fHbp in 95% of strains and elicited human complement-mediated, cooperative bactericidal activity with other non-bactericidal anti-fHbp mAbs.

Our data showed that JAR4 recognized a well defined area on the N-domain of fHbp, demonstrating that this region, previously indicated as involved on JAR4 binding (2), is actually implicated in the interaction with this mAb.

6.2 Introduction

Meningococcal factor H-binding protein (fHbp) is a promising vaccine candidate that is part of two vaccines in clinical development (3).

The protein binds human factor H (hfH) (4, 5) which is a down-regulatory molecule in the complement pathway (6).

By the binding of hfH on the surface of Meningococcus, the latter avoid the complement system by mimicking host surfaces and survive in the in human blood causing meningitis.

Antibodies against fHbp both activated the classical complement pathway and also block the binding of hfH to the surface of the bacteria (1, 7). Among different strains of *N. meningitidis*, fHbp exists in several antigenic variant groups based on antibody cross-reactivity and amino acid sequence identity. These fHbp groups have been designated variant 1, 2 or 3 by Masignani et al. (8).

A panel of anti-fHbp mAbs were previously characterized (9); several of the mAbs were bactericidal against strains with relatively high expression of fHbp

but were not bactericidal against other strains with lower expression of fHbp. Moreover a huge number of mAbs, showed a bactericidal activity only if combined with others mAbs. One example is JAR4, an antibody isolated from a mouse immunized with recombinant variant 1 fHbp. It indeed elicited a cooperative activity in combination with mAb JAR3 (1, 7), but the mechanism through which this occurs is not clear.

In fact, while mAb JAR3 was bactericidal only against strains with relatively high fHbp expression, mAb JAR4 was not bactericidal against any strains. The low-expressing fHbp strains required both mAbs together for complement-mediated bactericidal activity.

JAR3 and 4 bind to non-overlapping epitopes because, as shown by ELISA, neither MAb inhibited binding of the other mAb to fHbp. JAR3 but not JAR4 also inhibited binding of purified fH to *N. meningitidis* (5).

From filamentous bacteriophage libraries containing random peptides that were recognized by JAR4, a consensus tripeptide DHK that matched residues 25–27 in the N-terminal domain of fHbp was identified.

Since DHK was present in both JAR 4-reactive and non-reactive fHbps, the tripeptide was necessary but not sufficient for reactivity.

Collectively, the data indicated that the epitope for JAR4 was discontinuous and involved DHK stretch, together with a three residues stretch (YGN) beginning at position 57 (2).

Thus, the region of fHbp encompassing residues 25–59 in the N-terminal domain is important for eliciting antibodies that can cooperate with other anti-fHbp antibodies for cross-reactive bactericidal activity against strains expressing fHbp from different antigenic variant groups.

In our NMR study we demonstrated, through NMR titration, this evidence. Indeed the region encompassing the first 60 aa was perturbed upon addition of JAR4, while the C- terminal domain of fHbp is not affected by JAR4 binding.

6.3 Methods

The interaction of the antigen binding fragment (FAb) of JAR4 (FAb4) (50 kDa) with ^{15}N ^{2}H labelled fHbp was investigated through NMR spectroscopy with ^{1}H–^{15}N TROSY HSQC experiments. All the NMR experiments were performed at 298 K on a Bruker Avance 900 spectrometer, working at 900.13MHz frequency and equipped with a cryogenically cooled probe.

Titrations were performed on 0.15 mM ^{15}N ^{2}H labelled fHbp sample with the unlabelled FAb4 up to an fHbp–FAb4 molar ratio of 1:1.45.

All the samples used were in 20 mM PBS (90% H_2O and 10% D_2O) at pH 7.0.

6.4 NMR mapping of the JAR4-fHbp interaction

The interaction between fHbp and the JAR4 was investigated through NMR titration, by analyzing the chemical shift perturbations caused in the HSQC spectra upon addition of JAR4.

Residues whose NH cross-peaks were affected by presence of JAR4 all belonged to the N-domain of fHbp, forming a clustered region of interaction.

These residues include the first α-helix, the loop connecting α1 and β1 and the β strands β1, β2, β3, β4, β5.

The mapping of residues involved in the binding with JAR4 is shown on figure 1 and included also the tripeptide DHK (25-27) and YGN (57-59), whose resonances were clearly perturbed in the presence of FAb4.

This work is still in progress. Previous immunological data have been now demonstrated and the next step of this project will regards the investigation of the interaction between the fHbp and other mAbs.

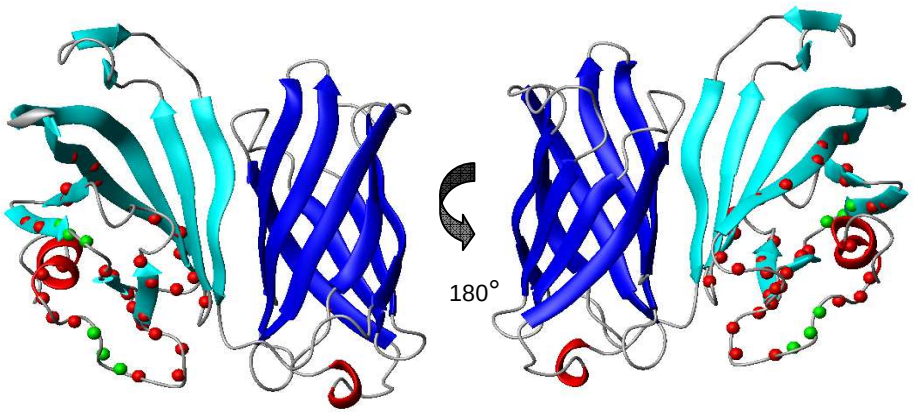


Figure 1. Aminoacids whose NMR signals were perturbed upon the addition of FAb4 to the NMR sample. They are indicated as solid spheres in correspondence to their backbone nitrogen atoms. In green are indicated DHK (25-27) and YGN (57-59) segments.

References

1. Beernink PT, Welsch JA, Bar-Lev M, Koeberling O, Comanducci M, Granoff DM., 2008, *Infect Immun*., 76(9), 4232-40.

2. Peter T. Beernink, Carla LoPasso, Antonella Angiolillo, Franco Felici, Dan Granoff, 2009, *Mol Immunol*., 46(8-9), 1647-53.

3. Rappuoli, R., 2008, The application of reverse vaccinology, Novartis MenB vaccine developed by design. In: van Alphen, L., van Ley, P., van den Dobbelsteen, G.(Eds.), Proceedings of the 16th International Pathogenic Neisseria Conference. Rotterdam, The Netherlands.

4. Granoff, D.M., Welsch, J.A., Ram, S., 2009, *Infect. Immun*., 77, 764–769.

5. Madico, G., Welsch, J.A., Lewis, L.A., McNaughton, A., Perlman, D.H., Costello, C.E., Ngampasutadol, J., Vogel, U., Granoff, D.M., Ram, S., 2006, *J. Immunol.*, 177, 501–510.

6. Schneider, M.C., Exley, R.M., Ram, S., Sim, R.B., Tang, C.M., 2007, *Trends Microbiol.,*15, 233–240.

7. Welsch, J.A., Ram, S., Koeberling, O., Granoff, D.M., 2008, *Infect. Dis*., 197, 1053–1061.

8. Masignani, V., Comanducci, M., Giuliani, M.M., Bambini, S., Adu-Bobie, J., Arico, B., Brunelli, B., Pieri, A., Santini, L., Savino, S., Serruto, D., Litt, D., Kroll, S., Welsch, J.A., Granoff, D.M., Rappuoli, R., Pizza, M., 2003*, J. Exp. Med*., 197, 789–799.

9. Welsch JA, Rossi R, Comanducci M, Granoff DM., 2004, *J Immunol*; 172 (9), 5606–15.

Chapter 7

General Discussion and Perspectives

General Discussion and Future Perspectives

Structural biology is a branch of the life science which provides a detailed analysis at atomic level of the structure of biological macromolecules, how they acquire the fold they have and how alterations in their structures affect their function, with the ultimate goal of investigating the relationship between these two latter aspects.

This subject is of great interest because macromolecules carry out most of the functions of cells and because it is only by arranging into specific three-dimensional shapes that they are able to perform their physiological roles.

The high resolution three-dimensional structure at the atomic level of a protein can be solved through X-ray crystallography or through NMR spectroscopy. The latter method has the advantage of studying proteins under conditions which can be as close as possible to their physiological state.

With NMR spectroscopy, a complete description of the molecule can be obtained, ranging from the folding process to the internal motions of the protein, all aspects which are essential for a deep understanding of the biological function. In addition, the recent development of methodology and algorithms for the quick determination of structures represent an enormous advance in structural biology in solution. The newly solved genomes from an increasing number of organisms bring an enormous number of new proteins to our knowledge, the majority of which have unknown function. This huge collection of new data needs to be analyzed and further investigated. In particular the structural and dynamic characterization of proteins is essential to shed light on their physiological functions.

Structure determination has already proven useful for the identification and evaluation of chemical leads (1). The number of high-resolution structures available in public databases today exceeds 30,000 and will definitely aid in structure-based drug design (2) and in improving drug potency and selectivity (3).

Using high resolution structures, antigens can be designed to cover more efficiently the majority of the strains of a pathogen.

By engineering antigens amenable for use in combination vaccines, immunization regimens can be simplified. By knowing which parts of an antigen must be retained to preserve basic characteristics and which can be altered, vaccine antigens can be modified more rapidly in response to changing epidemiology.

My work of thesis was developed in collaboration with Novartis-Vaccines and it was focused, in the frame of the extensive project concerning the development of a universal vaccines against Meningococcus B (MenB), on structural studies on new antigen candidates against *Neisseria Meningitidis.*

In the year 2000 Novartis laboratory published the use of the genomic sequence of a particular strain of MenB (MC58 strain) to discover novel antigens capable of inducing protection against serogroup B of *N. Meningitidis* (4).

Factor H binding protein (fHbp) is currently the more effective antigen discovered while screening the bacterial genome (5, 6) able to induce bactericidal antibodies, so the most important constituent of a universal vaccine.

The peculiarity of fHbp consists in its double function: killing the bacterium by direct activation of the classical complement cascade and preventing the formation of the complex human factor H(hfH)-fHbp on the surface of the bacterium (7, 8).

During my PhD I focused my attention on structural and dynamic characterization of fHbp. This study had a big scientific impact because not only allowed to solve a structure unknown so far, but also because it is the starting point for following important studies, like the investigation of the interaction with several monoclonal Antibodies or other important human proteins, as the hfH.

NMR studies turned out to be crucial for a deep investigation of the problem and provided important information also for future designing of new vaccine antigens.

Thanks to the improved NMR technology, the avaibility of high magnetic fields and new kind of NMR experiments like TROSY and CRINEPT, which allow the structural characterization of proteins with high molecular weight, it was possible to determine structures like fHbp, of more than 250 residues.

We defined two different domains of fHbp with a β barrel like conformation, connected by a five residue linker and from mobility data the value of the correlation time (τ_c) for the molecule tumbling (estimated of 20.53 ± 1.5 ns) confirms that the protein rotates in solution as a single entity, being in a monomeric state.

The discovery of this structure has given the bases for a series of analysis, focused on the study of the interaction of fHbp with other important macromolecules, crucial targets of the immunity system, as for example monoclonal Antibodies (mAbs) and the hfH.

The structural approach applied to vaccinology may help the rational engineering of future antigens; structural studies of antigen epitopes may help the understanding of the molecular nature of epitopes and their mechanisms of action. Moreover the application of the structural biology to the vaccinology field allowed to investigate on the interaction Antigen-Antiboby, process which is a crucial step for the immunity response.

The latter aspect was firstly investigated: the study of the interaction between the monoclonal Antibody MAb502 and the C-terminal domain of fHbp ($fHbp_c$) allowed to define critical residues for the binding.

Through NMR indeed, it is possible to monitor the NH chemical shift variation of residues which are directly involved in the interaction with MAb502.

The information obtained together with biological data have been fundamental to define a conformational epitope within a well define area on $fHbp_c$.

Besides NMR technique was used to study the interaction of fHbp full length and a construct of factor H comprising domains 5, 6 and 7.
The importance of this study consists on the fact that the binding with hfH is a form of resistance of MenB to the immunity system, because covering its surface with hfH, fHbp mimicks the host surface, avoiding the complement attack.
Especially in this case NMR spectroscopy resulted to be essential because it was possible, at high magnetic field and with the use of TROSY experiments, to study a complex of about 80 kDa.
Moreover this investigation confirmed the fact that residues involved in binding with hfH do not overlap with the epitope for MAb502, supporting the previous theory that MAb502 is a monoclonal Antidoby not inhibiting-hfH binding.

My project then concentrated on the study of the interaction with other variant-specific mAbs, which carry on their bactericidal function only in presence of other mAbs, as for example the anti-fHbp mAb JAR4. Also in this case this study allowed to map the region of interaction on fHbp.

With respect to expression of fHbp, *N. meningitidis* strains can broadly be subdivided into 2 groups: strains associated with high level of expression of fHbp and strains associated with low level of expression of fHbp (5).
Several data underscore the correlation between low expression of fHbp by a strain and the requirement for >1 anti-fHbp mAb to elicit bactericidal activity (5).
Based on this statement, we started to investigate through NMR measurements on the mapping of the interaction between fHbp and JAR4.
This kind of analysis will proceed with the investigation, by NMR experiments, of the interaction with other mAbs to define the bactericidal synergy.

The MenB vaccine, containing as principal component fHbp variant 1, can be improved in order to increase the immunological coverage using engineered fHbps (also called patches) that display epitopes of different variants.
Several mutants containing different combination of epitopes of variant 1 and 2 have been so far expressed by site-directed mutagenesis, introducing groups of point mutations into the fHbp v1 scaffold. These mutants have been named: P2S, P5A, P5E, P8B, P9C, P10A, P10B, P11B, P11I and P13B. Multiple sequence alignment of wild-type fHbps and various engineered fHbps (patches) is reported on figure 1.
The impact of respective fHbp sequence modifications on the overall structure is addressed using various analytical tools.
Results may help to improve interpretation of different protective properties found for various engineered fHbps when studied in vivo from a more structural point of view and potentially allow further optimization of the antigen.
Preliminary NMR studies, through ^{1}H-^{15}N HSQC experiments analysis, allowed to assemble fHbp mutants on the basis of the similairty of the their HSQC specta, to further understand if a correlation between groups of mutants with similar HSQC profile and comparable immunological activity exists.
From the comparison *fHbp wild-type - fHbp patch* HSQCs, it appears that most of the NH-cross peaks chemical shifts are maintained, however easily reassignable, especially at the N-domain.
Together with respective immunization studies (using various patches in different vaccine formulations), the data may help to better understand the potential impact of structural changes within engineered fHbps on their cross-protective immunity and potentially allow a further optimization of the antigens.

Future additional structural and dynamic studies will be expand to other variants of fHbp, to further investigate on the interaction between these variants and variant-specific mAbs.

During the last part of my PhD my attention will be devote on the study of variant 3 (v3). In difference with variant 1, variant 3 results less stable, in fact the chemical shifts of some NH groups in the HSQC spectra are different for different samples; this indicates that the structure of the protein slightly changes according to the protein preparation. The most significant changes in chemical shift can be assigned to the residues of the N and C-terminal regions. These regions are also the most affected by degradation in time.

Up to now 85% of the backbone atoms have been assigned. Most of residues which are not assigned yet are located in loop regions of the β barrel at C-terminal domain, as indicated on a model of fHbp v3 shown on figure 2. Comparison of the secondary structures elements of variant 1 and 3, the latter estimated by the Chemical Shift Index (CSI) analysis (9), showed that some β strands in variant 3 are shorter than β strand on variant 1.
Sequence similarity is quite high between v1 and v3 (around 63%).

The final goal of the global project is to improve the vaccine preparation, in order to increase the immunological coverage to the majority of MenB strains, using engineered fHbps that display epitopes of different variants, aimed to obtain a single chimeric molecule.

Figure 1. Multiple sequence alignment of wild-type fHbps and various engineered fHbps (patches). Points of mutation are indicated in circles.

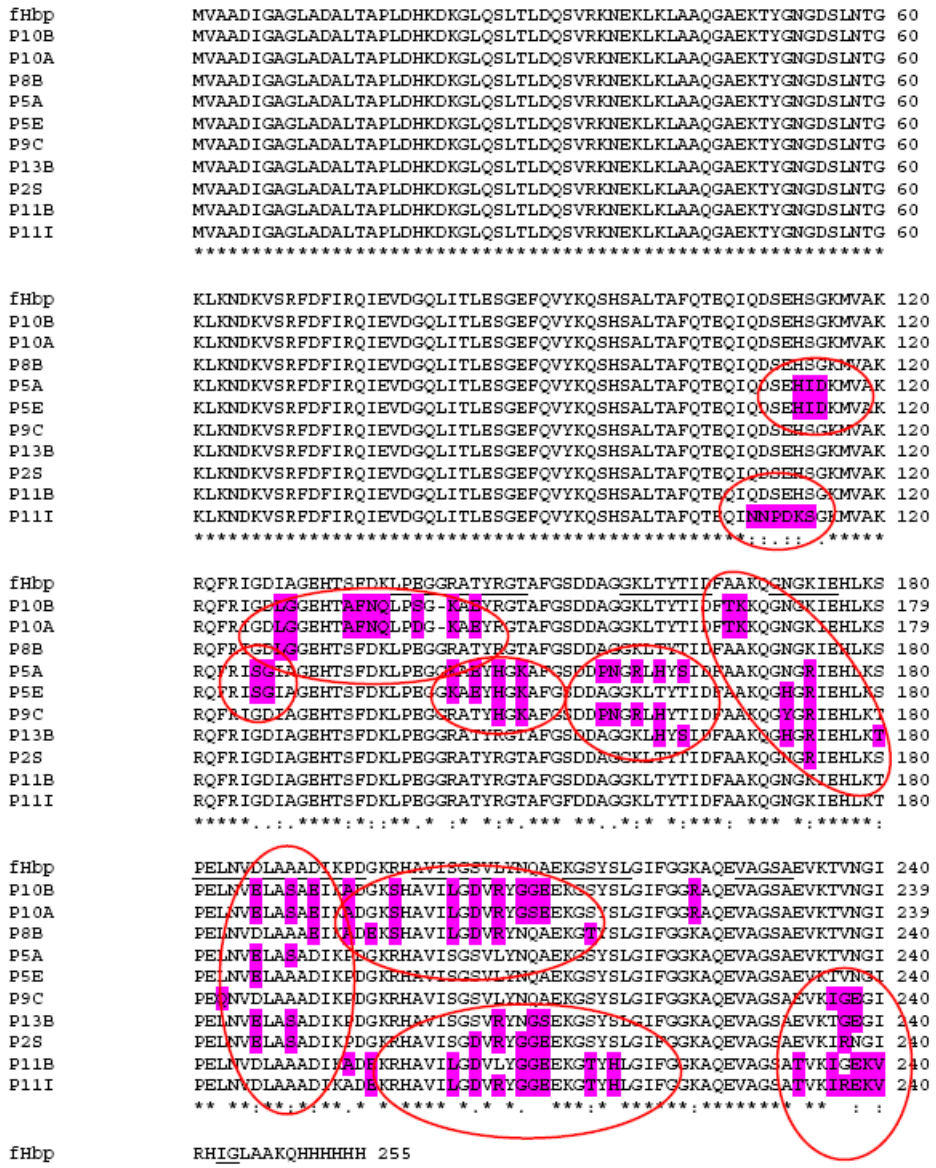
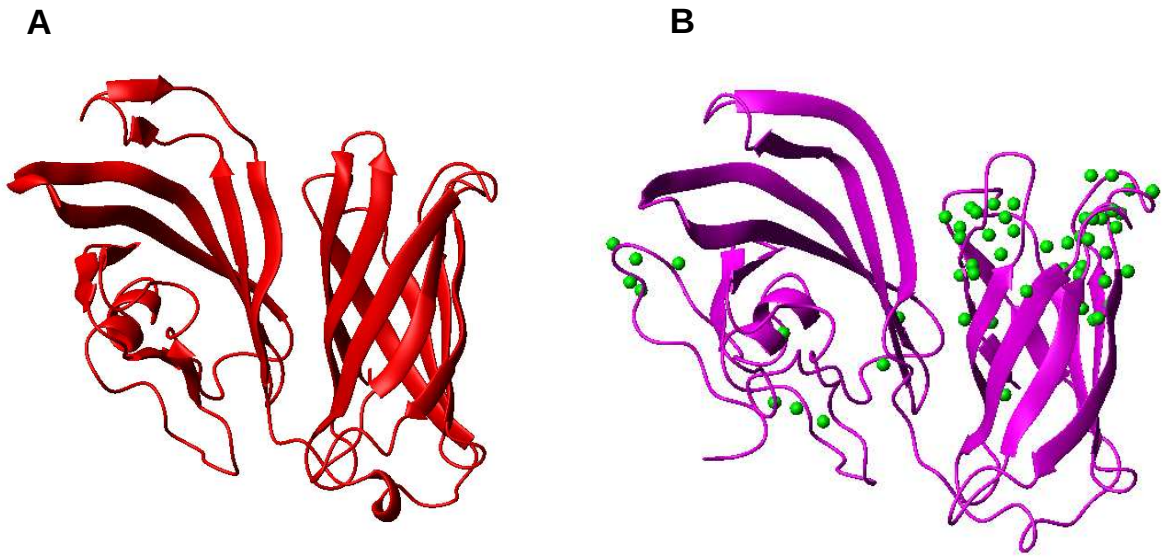


Figure 2. **A** Structure of fHbp variant 1. **B** Model of fHbp variant 3 obtained with MODELLER using the NMR structure of fHbp v1 as template. Residues not assigned yet are indicated as green sphere in correspondence to their nitrogen atom.



References

1. Moore JM., 1999, *Biopolymers.,* 51, 221–43.

2. Lundstrom K., 2007, *J. Cell. Mol. Med*., 11 (2), 224-238.

3. Campbell SF., 2000, *Clin Sci*., 99, 255–60.

4. Pizza, M., V. Scarlato, V. Masignani, M.M. Giuliani, B. Arico, M. Comanducci, G.T. Jennings, L. Baldi, E. Bartolini, B. Capecchi, et al,. 2000, *Science,* 287, 1816–1820.

5. Masignani V, Comanducci M, Giuliani MM, Bambini S, Adu-Bobie J, Arico B, Brunelli B, Pieri A, Santini L, Savino S, Serruto D, Litt D, Kroll S, Welsch JA, Granoff DM, Rappuoli R, Pizza M., 2003, *J Exp Med*., 197(6), 789-99.

6. Giuliani MM, Santini L, Brunelli B, Biolchi A, Aricò B, Di Marcello F, Cartocci E, Comanducci M, Masignani V, Lozzi L, Savino S, Scarselli M, Rappuoli R, Pizza M., 2005, *Infect Immun*., 73(2), 1151-60.

7. Beernink PT, Welsch JA, Bar-Lev M, Koeberling O, Comanducci M, Granoff DM., 2008, *Infect Immun*., 76(9), 4232-40.

8. Welsch JA, Ram S, Koeberling O, Granoff DM., 2008, *J Infect Dis*., 197(7), 1053-61.

9. Wishart, D.S., Sykes, B.D. & Richards, F.M., 1992, *Biochemistry*, 31, 1647-1651.