Structure of the C-terminal Domain of *Neisseria* Heparin Binding Antigen (NHBA), One of the Main Antigens of a Novel Vaccine against *Neisseria meningitidis**[S]

Veronica Esposito[‡1], **Valeria Musi**[‡], **Cesira de Chiara**[‡], **Daniele Veggi**[§], **Davide Serruto**[§], **Maria Scarselli**[§], **Geoff Kelly**[‡], **Mariagrazia Pizza**[§2], and **Annalisa Pastore**[‡2]

From the [‡]MRC National Institute for Medical Research, The Ridgeway, London NW71AA, United Kingdom and [§]Novartis Vaccines and Diagnostics, via Fiorentina 1, 53100 Siena, Italy

Background: NHBA, a surface-exposed lipoprotein from *Neisseria meningitidis*, is part of a multicomponent vaccine against serogroup B meningitis.
Results: We have solved the structure of a conserved C-terminal domain that adopts a β-barrel fold and seems to be the only independently folded region of the protein.
Conclusion: We observed a significant structural similarity with other *Nesseria* proteins.
Significance: Our data represent the first step toward understanding the structure/immunology relationship of NHBA.

Neisseria **heparin binding antigen (NHBA), also known as GNA2132 (genome-derived** *Neisseria* **antigen 2132), is a surface-exposed lipoprotein from** *Neisseria meningitidis* **that was originally identified by reverse vaccinology. It is one the three main antigens of a multicomponent vaccine against serogroup B meningitis (4CMenB), which has just completed phase III clinical trials in infants. In contrast to the other two main vaccine components, little is known about the origin of the immunogenicity of this antigen, and about its ability to induce a strong cross-bactericidal response in animals and humans. To characterize NHBA in terms of its structural/immunogenic properties, we have analyzed its sequence and identified a C-terminal region that is highly conserved in all strains. We demonstrate experimentally that this region is independently folded, and solved its three-dimensional structure by nuclear magnetic resonance. Notably, we need detergents to observe a single species in solution. The NHBA domain fold consists of an 8-strand β-barrel that closely resembles the C-terminal domains of** *N. meningitidis* **factor H-binding protein and transferrin-binding protein B. This common fold together with more subtle structural similarities suggest a common ancestor for these important antigens and a role of the β-barrel fold in inducing immunogenicity against** *N. meningitidis.* **Our data represent the first step toward understanding the relationship between structural, functional, and immunological properties of this important vaccine component.**

Neisseria heparin binding antigen (NHBA), also known as GNA2132 (genome-derived *Neisseria* antigen 2132), is a surface-exposed lipoprotein from *Neisseria meningitidis* that was originally identified by reverse vaccinology. It is one the three main antigens of a multicomponent vaccine against serogroup B meningitis (4CMenB), which has just completed phase III clinical trials in infants. In contrast to the other two main vaccine components, little is known about the origin of the immunogenicity of this antigen, and about its ability to induce a strong cross-bactericidal response in animals and humans. To characterize NHBA in terms of its structural/immunogenic properties, we have analyzed its sequence and identified a C-terminal region that is highly conserved in all strains. We demonstrate experimentally that this region is independently folded, and solved its three-dimensional structure by nuclear magnetic resonance. Notably, we need detergents to observe a single species in solution. The NHBA domain fold consists of an 8-strand β-barrel that closely resembles the C-terminal domains of *N. meningitidis* factor H-binding protein and transferrin-binding protein B. This common fold together with more subtle structural similarities suggest a common ancestor for these important antigens and a role of the β-barrel fold in inducing immunogenicity against *N. meningitidis.* Our data represent the first step toward understanding the relationship between structural, functional, and immunological properties of this important vaccine component.

Neisseria meningitidis, a Gram-negative capsulated bacterium, is a strictly human pathogen that induces meningitis and sepsis in children and young adults (1). It colonizes a substantial proportion of the population (3–30%), where it resides as a commensal having the nasopharynx as the only known reservoir (2). In susceptible individuals, *N. meningitidis* can cause septicemia by crossing the mucosal barrier and entering the bloodstream or meningitis by crossing the blood-brain barrier and multiplying into the cerebrospinal fluid (3). The bacterium is surrounded by capsular polysaccharides, the exact nature of which defines 13 different serogroups. Five of them (A, B, C, Y, and W135) are the main cause of disease.

Prevention from meningococcal meningitis can be effectively accomplished by vaccination as conjugate vaccines against serogroups A, C, W135, and Y based on polysaccharides are available or in the late phase of development (4). The capsular polysaccharide of meningococcus B (MenB),[3] however, has structural homology with a polysialic acid ($\alpha(2-8)N$-acetyl-neuraminic acid) present in human cells thus preventing the use of the strategy adopted for the development of conjugate vaccines (5). To date, the only vaccines proven to be efficacious against serogroup B are based on outer membrane vesicles purified from the bacterium. These vaccines, termed "tailor made," can be highly effective for epidemic control (6). The only limitation of such vaccines is their strain specificity due to the high variability of PorA, the main antigenic component of the vesicles. To overcome these limitations and develop a vaccine potentially able to cover the MenB strain diversity, we used a genome-based approach, commonly named "reverse vaccinology," which allowed the identification of novel surface-exposed proteins. These molecules constituted the basis of a multicomponent MenB vaccine (4CMenB), which has just completed phase III clinical trials in infants (7).

[3] The abbreviations used are: MenB, meningococcus B; fHBP, factor H-binding protein; TbpB, transferrin-binding protein B; NHBA, *Neisseria* heparin binding antigen; DPC, dodecylphosphocholine; HSQC, heteronuclear single quantum coherence; PDB, Protein Data Bank.

The three main antigens composing the MenB vaccine are NadA, fHBP, and NHBA (8). NadA (Neisseria adhesin A) is an outer membrane protein with trimeric structure that belongs to the OCA protein family (oligomeric coiled-coil adhesins), a group of autotransporters that mediates oligomerization involved in adhesion to host cells (9).

fHBP (factor H-binding protein) is a bactericidal surface-exposed lipoprotein able to bind human factor H. Structural studies have shown that fHBP contains two domains each folded in a "handled β-barrel," consisting of a classical β-barrel that is preceded by an all-β motif, which packs against the barrel. The protein is connected to the membrane by a flexible linker, which leaves it mostly solvent exposed (10, 11). Immunological and structural studies showed that most of the bactericidal epitopes of fHBP are located in the carboxyl-terminal β-barrel (8, 9, 12–19). The structure of the fHBP in complex with factor H was solved by x-ray crystallography and shows that interaction between the two proteins involves residues from both fHBP domains (20). Although these main allergens are at least partially characterized from the structural and immunological point of view, little is known about the third one.

NHBA (*Neisseria* heparin binding antigen, also named GNA2132) is a surface-exposed lipoprotein that is present in all *N. meningitidis* strains so far analyzed. The antigen induces cross-bactericidal activity in humans and animals and is recognized by sera of patients after meningococcal disease (21–24). The protein contains an arginine-rich region that mediates heparin binding *in vitro*, a property that correlates with increased survival of *N. meningitidis* in human serum (24). This evidence suggested the hypothesis that heparin could either mimic other important negatively charged partners or mediate the interaction with factor H and other complement components. NHBA is the target of two proteases: the meningococcal NalP and human lactoferrin, with the proteolytic sites being, respectively, located immediately upstream and downstream to the arginine-rich motif. This implies that both the host and the pathogen are able to produce fragments containing the highly conserved C terminus, which could then have an independent functional role also when not anchored to the membrane (24).

Here, we present a structural characterization of the C-terminal region of NHBA based on sequence analysis, circular dichroism (CD), and nuclear magnetic resonance (NMR) studies. We show that although the N terminus of NHBA is annotated as an intrinsically unfolded region by the most common prediction methods, the C terminus has all the features of an independently folded globular domain. Accordingly, we have solved the solution structure of the C-terminal domain and shown that it consists of an 8-stranded β-barrel. Based on a number of different structural parameters, we propose that NHBA is evolutionarily related to the other MenB vaccine antigen fHBP and to transferrin-binding protein B (TbpB), although likely with a divergent function. Our findings add new insights for understanding the structural properties of this important vaccine antigen and constitute the basis for future immunological studies.

MATERIALS AND METHODS

Sequence Analysis—Sequence homology was searched by BLAST2. The NHBA sequence was submitted to several World Wide Web servers for prediction of intrinsically unfolded regions: DisEMBL, GlobPlot, DISOPRED2, IUPred, and PONDR. Secondary structure prediction was achieved by standard bioinformatic tools using JPred3, PHD, and PSIPRED freely available on the World Wide Web.
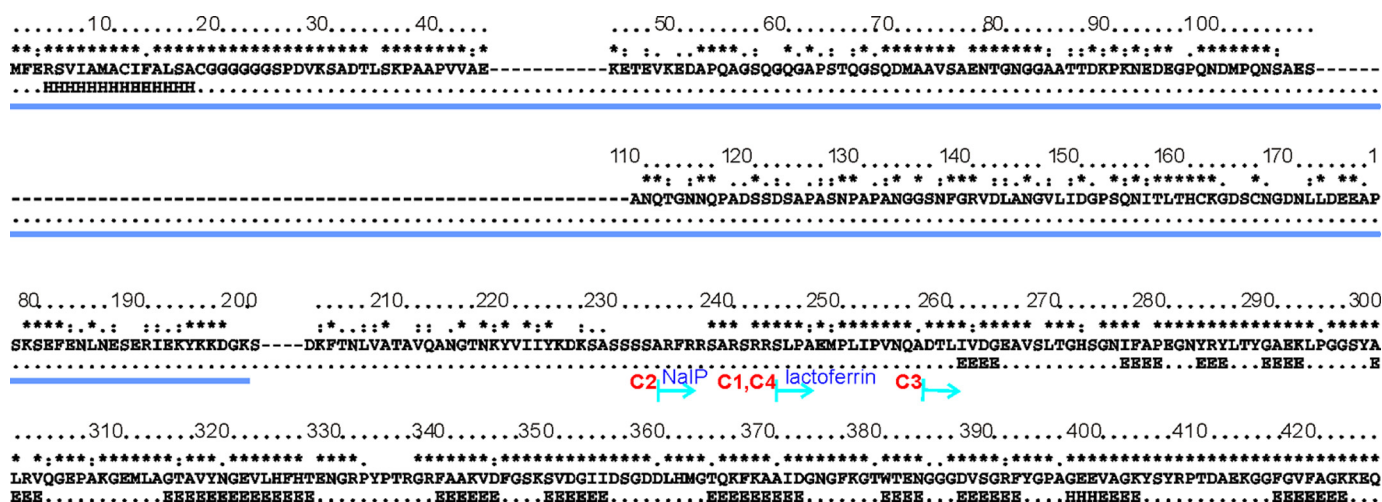
Sample Production—The production of recombinant NHBA and its fragments has been described elsewhere (25). In short, the *NHBA* gene, cloned in a pET21b+ vector (Novagen), was expressed in *Escherichia coli* BL21(DE3)-pLysS. Uniformly $^{15}N,^{13}C$-labeled NHBA samples were produced by growing cells at 37 °C in M9 minimal medium, enriched with 1 g/liter of ^{15}N-labeled ammonium sulfate and 4 g/liter of [U-$^2H,^{13}C$]glucose (Cambridge Isotope Laboratories). The proteins were induced with 1 mM isopropyl β-D-thiogalactopyranoside and expressed for 4 h. The cell pellets were suspended in 50 mM sodium phosphate buffer at pH 8 and lysed by lysozyme. Because all constructs were fused with a His$_6$ tag (usually C-terminal except for C4 in which it was N-terminal and cleavable), they were purified in two steps by affinity chromatography, using nickel-nitrilotriacetic acid and cation exchange columns equilibrated in 50 mM sodium acetate, pH 5.5. Sample purity and identity were checked by SDS-PAGE and electrospray mass spectrometry, respectively.

Circular Dichroism Measurements—Far UV CD spectra were recorded on a Jasco J-715 spectropolarimeter in the spectral range 195–250 nm. The near UV CD spectra of proteins were recorded in the spectral range between 250 and 340 nm. Thermal denaturation was followed by monitoring the far UV CD signal at 222 nm while changing the temperature from 2 to 95 °C at 1 °C/min. All measurements were performed at 20 °C in 20 mM Tris-HCl, pH 6.8, 100 mM NaCl. Dodecylphosphocholine (DPC) (up to 10 mM) was added to the solutions when exploring the effect of detergents.

NMR Experiments—NMR samples typically contained 0.5 mM protein concentrations (95% H_2O, 5% D_2O) in 20 mM Tris-HCl, pH 7, and 50 mM NaCl. Different detergents (CHAPS, lauryl dimethylamine oxide, dimyristoylphosphatidylcholine, and DPC, all purchased from Aldrich) were added to these solutions for screening purposes. Structure determination was achieved in solutions to which 30 mM DPC was added. The spectra were recorded at 40 °C on Bruker Avance spectrometers operating at 600 and 700 MHz and equipped with a cryoprobe. Data were processed with NMRPipe/NMRDraw (26) and analyzed with Sparky (51). Intermolecular interactions between NHBA and DPC were identified using a three-dimensional $^{13}C,^{15}N$-filtered ^{13}C-edited NOESY-HSQC ($^{13}C,^{15}N$-labeled NHBA, unlabeled DPC) experiment (adapted from Ref. 27).

Identification of Structure Restraints—Interproton distance restraints were derived from ^{15}N and ^{13}C NOESY-HSQC spectra with a mixing time of 100 ms recorded on a Bruker Avance spectrometer operating at 700 MHz ^1H frequency. 118 φ and 120 ψ dihedral angles were obtained by using the backbone torsion angle prediction package TALOS+ (28). Amide protec-

```
.......10........20.......30........40....        ....50.......60.......70.......80.......90......100........
**:.********* .*****************.***. ******.*     *: :. ..****.: *.*.: :*.****** *******; ::*:*;*;** .******:
MFERSVIAMACIFALSACGGGGGGSPDVKSADTLSKPAAPVVAE-----------KETEVKEDAPQAGSQGQGAPSTQGSQDMAAVSAENTGNGGAATTDKPKNEDEGPQNDMPQNSAES------
...HHHHHHHHHHHHH...........................
```

```
                                                  110........120.......130........140.......150.......160.......170......1
                                                  **: :** ..*.:. .:;** ..:* * ;*** ::.*.: :*. *;*:*;****** *.   .:* **.
-----------------------------------------------------ANQTGNNQPADSSDSAPASNPAPANGGSNFGRVDLANGVLIDGPSQNITLTHCKGDSCNGDNLLDEEAP
```

```
80......190.......200     .....210.......220.......230.......240.......250.......260......270.......280.......290......300
 ****:.*.: ::.:****        :*..::* :: .* *:* *:* :..       *** *****:*:******* .****:***** ***; **** ************** .*****
SKSEFENLNESERIEKYKKDGKS----DKFTNLVATAVQANGTNKYVIIYKDKSASSSSARFRRSARSRRSLPAEMPLIPVNQADTLIVDGEAVSLTGHSGNIFAPEGNYRYLTYGAEKLPGGSYA
...................................................................EEEE...........EEEE...EEE...EEEE......E
                                               C2 NalP    C1,C4 lactoferrin   C3
```

```
.......310.......320.......330......340......350.......360.......370.......380.......390......400......410......420.....
* *;:***;*******:*** .******:*      *******.*********.****.******* .*****.*:***; ****;.*****;**** **************************
LRVQGEPAKGEMLAGTAVYNGEVLHFHTENGRPYPTRGRFAAKVDFGSKSVDGIIDSGDDLHMGTQKFFKAAIDGNGFKGTWTENGGGDVSGRFYGPAGEEVAGKYSYRPTDAEKGGFGVFAGKKEQ
EEE..........EEEEEEEEEEEEE..........EEEEEE....EEEEEE.......EEEEEEEEE.......EEEE...EEEEEE...HHHEEE..........EEEEEEE...
```

FIGURE 1. Analysis of the NHBA sequence. For simplicity, only the sequence of the 2996 strain is reported but gaps are inserted according to the multiple alignment with other strains reported under supplemental Fig. S1. Marked on the *top* is the degree of conservation over the sequences of 18 different strains. *Asterisks* indicate complete conservation, *colons* and *dots* indicate positions conserved in 50 and 25% of the sequences, respectively. The consensus secondary structure as predicted by different servers is indicated *below* the sequence. *H* and *E* indicate helical and extended regions, respectively. The consensus region predicted as intrinsically unfolded by three different algorithms (Globplot, Disopred, and PONDR) as shown under supplemental Fig. S2 is *underlined in blue*. The starting points of the different constructs analyzed experimentally are indicated. They all run up to the end of the sequence. NHCA_C1 and NHCA_C4 differ by the presence of a C-terminal histidine tag in the former and cleaved off in the latter. The N termini of the NHCA_C1 and NHCA_C2 constructs coincide with the lactoferrin and NalP cleavage sites, respectively.

tion was inferred from deuterium exchange measurements performed at 40 °C on a freeze-dried, ^{15}N-labeled sample, re-dissolved in a 20 mM Tris-HCl, pH 7, 50 mM NaCl, and 30 mM DPC D$_2$O solution. A total of 50 slowly exchanging protons were identified as observable signals in the ^{1}H,^{15}N-HSQC spectra recorded immediately after redissolving the freeze-dried protein. A hydrogen bond restraint involving a slowly exchanging proton was added only if a hydrogen bond was also consistently observed in at least 50% of the initial structures inspected.

Structure Calculations—Structure calculations were performed by using the ARIA program (version 1.2) (29). Twenty structures were calculated by simulated annealing at each ARIA iteration using the standard CNS protocol (30). Floating assignment for prochiral groups and correction for spin diffusion during iterative NOE assignment was applied (31, 32). At the end of each iteration, the best seven structures in terms of the lowest global energy were selected and used for the assignment of additional NOEs during the following iteration. In the final ARIA run, the number of structures generated in iteration 8 was increased to 100. After refinement by molecular dynamics simulation in water (33), the 20 lowest energy structures were selected as representative and used for statistical analysis. The ARIA input used to generate the final bundle of structures consisted of about 50% manually assigned NOE cross-peaks from ^{15}N- and ^{13}C NOESY-HSQC spectra, manually picked to exclude noise and artifacts deriving from spin diffusion effects, along with ϕ and ψ dihedral angles, hydrogen bonds, and a chemical shift assignment. The root mean square deviation calculated by superimposing the backbone atoms of the structured regions was 0.54 ± 0.10 Å. In the final iteration, 1637 unambiguous and 116 ambiguous NOEs were assigned. Among the 1753 total NOEs, 638 were intraresidue, 430 sequential, 75 and 610 medium and long range distances, respectively. Molecular visualization and graphics were prepared with MolMol (34). The coordinates are available from the PDB data base with accession name 2lfu.

Heparin Titrations—Samples of NHBA_C1 and NHBA_C2 (0.5 μM) were titrated with heparin (average molecular mass of 5 kDa, Fisher Bioreagents) at molar ratios of 0.01, 0.02, 0.05, 0.1, 0.2, 0.25, 0.3, 1.0, and 2.0. The titrations were followed by recording NMR spectra on a 600 MHz Bruker Avance at 40 °C.

RESULTS

The N Terminus of NHBA Contains Potential Intrinsically Unstructured Regions—The sequence of NHBA differs only slightly among different *N. meningitidis* strains (Fig. 1 and supplemental Fig. S1). The maximal variability is observed in the N terminus of the protein, which also contains several insertions/deletions. We selected the NHBA amino acid sequence of the 2996 strain for this study (23) and analyzed it with various bioinformatics software. Independent methods all suggest that the N terminus of NHBA, up to approximately residue 200–250, is mostly unstructured (supplemental Fig. S2): this region is rich in small hydrophilic and hydrophobic amino acids (*e.g.* alanines, asparagines, and serines) and contains different low complexity sequences (*e.g.* polyglycine and polyserine motifs). Among the latter motifs, it should be mentioned that the arginine-rich region present around residues 235–245 is known to mediate interaction with heparin (24). The cleavage sites of NalP and lactoferrin are directly contiguous to the arginine-rich motif. Thus, although we cannot exclude that some shorter stretches could be structured in the context of the full-length protein, the N-terminal region has the most features of an intrinsically unfolded region (35). The C terminus of the protein has instead strong β-propensity and shares distant but clear homology with the sequence of TbpB and with fHBP. Based on this study we decided to characterize structurally protein constructs spanning only the C terminus of the protein.
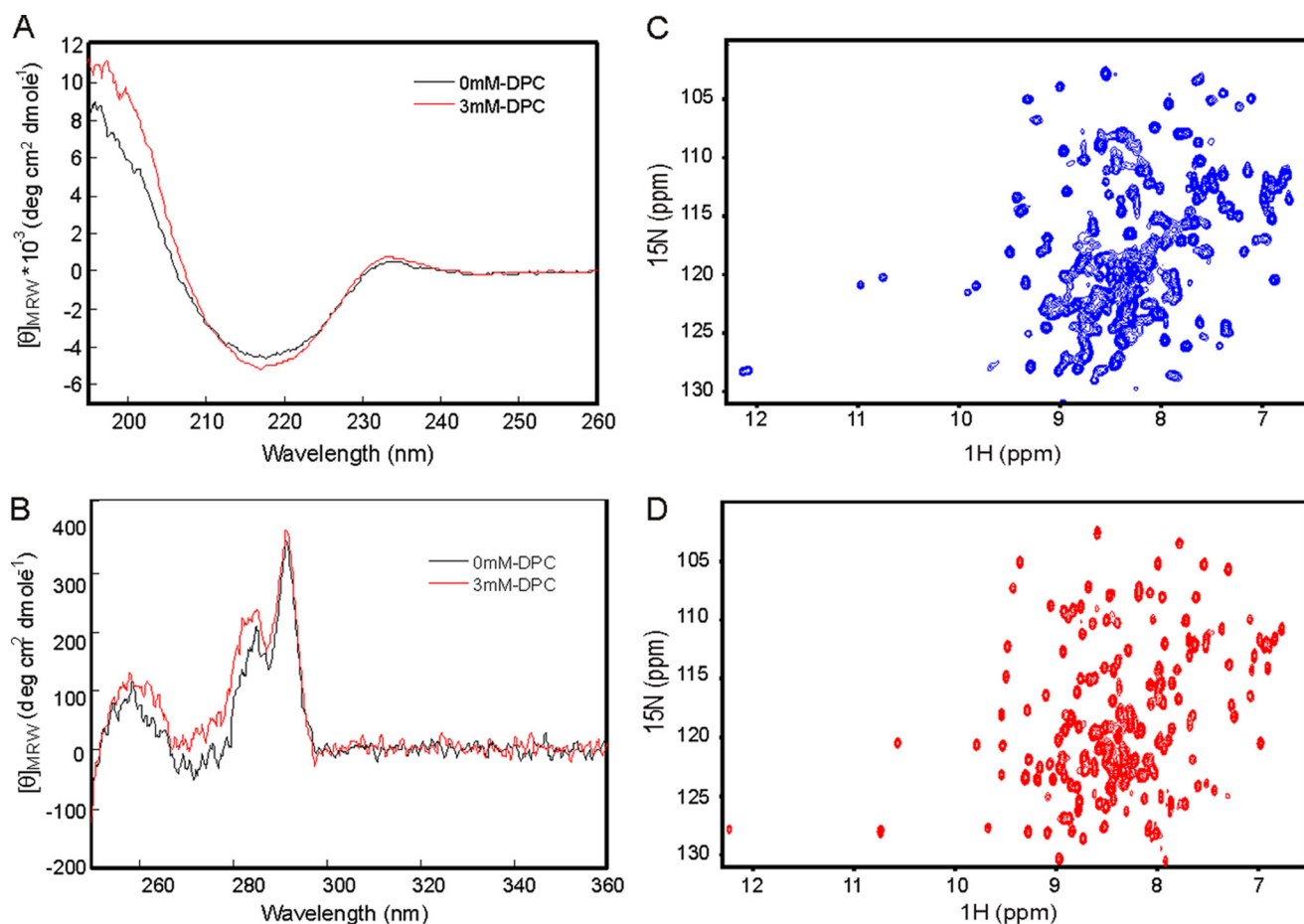
FIGURE 2. **CD and NMR spectra of NHBA_C1.** *Left panels*, comparison of the far UV (*A*) and near UV (*B*) CD spectra recorded at 35 °C and in the absence (*black trace*) and presence (*red trace*) of 3 mM DPC. *Right panels*, comparison of the ^{15}N-HSQC spectrum of NHBA_C1 recorded at 40 °C and 600 MHz in the absence (*C*) and presence (*D*) of 30 mM DPC.

*The C Terminus of NHBA Forms a Domain with an All-β-fold—*We characterized a construct of the 2996 strain that starts at the ferrolactin cleavage site at Ser-245 (SLPAE) and covers up to the end of NHBA (hereafter indicated as NHBA_C1) followed by a noncleavable C-terminal histidine tag (LEHHHHHH). The secondary and tertiary structures of this construct were, respectively, investigated by far and near UV CD. The far UV CD spectrum of NHBA_C1 has features typical of a β-rich protein with a minimum around 218 nm and a maximum at 230 nm (Fig. 2*A*). A positive band around this wavelength has been attributed to packing/stacking of aromatic residues (36). The near UV CD spectrum confirms that NHBA_C1 is folded having an intense positive ellipticity in the 280–300 nm range, which is indicative of Tyr and Trp residues embedded in a compact, well defined tertiary structure (Fig. 2*B*) (37). A lower positive ellipticity in the Phe region (250–270 nm) of the spectra was also observed. Thermal unfolding curves followed by both near and far UV CD indicate a high stability because the spectra remains invariant up to temperatures as high as 90 °C (supplemental Fig. S3).

*NHBA Constructs Are Heterogeneous in Solution—*The NMR spectrum of NHBA_C1 confirmed that the construct is an independently folded domain with an overall excellent dispersion of the resonances (Fig. 2*C*). However, we realized very early on that the number of resonances present in the ^{15}N-HSQC spec-

trum, which represents the protein fingerprint, contained at least 10–20% resonances more than expected, suggesting the presence in solution of different forms. Particularly diagnostic was, for instance, the resonance that resulted that corresponds to residue 287 (12.1 and 128.5 ppm), which was clearly doubled.

To understand the origin of this heterogeneity, we first ensured that the protein preparations were chemically homogeneous. We tested this hypothesis both by mass spectrometry and amino acid analysis. Both techniques evidenced the presence of some N-terminal degradation with 1–3 amino acids cleaved off (data not shown). However, the degree of this spontaneous proteolysis did not correlate with the relative populations observed by NMR. A second possibility to explain our data were a different degree of mono-/polydispersity. To test this hypothesis we studied the size exclusion chromatography profile of NHBA_C1 but found no clear evidence for different molecular weight species (supplemental Fig. S4).

We then explored whether the choice of the domain boundaries could have an effect on the NMR spectrum. Different constructs were produced by extending or truncating the N terminus of the NHBA_C1 construct (Fig. 1). One of the constructs (NHBA_C4) was designed with a cleavable N-terminal His tag so that, once the tag was cleaved off, the protein would have no added residues. Another construct (NHBA_C3) excluded the proline-rich stretch present between residues 245 and 254, thus

reducing the danger of cis/trans isomerization. A longer construct (NHBA_C2) coincided with the NaIP proteolytic fragment and included 14 additional N-terminal amino acids that had been shown to be important for heparin binding (24). However, despite these efforts, all constructs invariably contained at least two distinct species.

After lengthy screening in which the effect of different experimental parameters (*e.g.* pH, ionic strength, metal binding, buffer composition, etc.) was probed, we reasoned that in nature the protein is close to the lipid membrane of the bacterium or may interact with the lipid membrane of the host once released by proteolytic cleavage. The NHBA fold could therefore be influenced by the presence of detergents. We found that addition of an excess of DPC (30 mM) or other similar detergents (*e.g.* CHAPS, dimyristoylphosphatidylcholine, lauryl dimethylamine oxide) was sufficient to obtain a spectrum with only one set of resonances, thus leading to a significant simplification of the NMR spectrum whereas leaving the far-UV CD spectra essentially invariant (Fig. 2, *A* and *D*). The detergents caused a small enhancement of the positive ellipticities in the Phe-Tyr region of the near UV CD spectrum (260–280 nm), suggesting that Tyr and Phe exposures are affected by detergent binding (Fig. 2*B*). No significant changes were revealed in the near UV Trp region (290 nm). We used these conditions in all subsequent structural studies.

NHBA_C1 Has a β-Barrel Fold and Belongs to the Lipocalin Superfamily—Once it was shown that DPC was able to abolish NHBA_C1 polydispersity, we determined the structure of this protein construct. Overall, the resonance line widths in DPC were slightly increased as compared with those in aqueous buffer. This is in agreement with the increased molecular weight caused by the presence of the detergent. The spectra were thus recorded at 40 °C to improve spectral resolution. As discussed elsewhere (25), the first 33 residues could be barely identified in the spectra and for most of them no assignment was possible presumably because of line broadening due to conformational or chemical exchange. Of the remaining 153 residues more than 94% of the backbone resonances and more than 87% of the side chain resonances were unambiguously assigned. Exceptions were Met^{363}, Lys^{367}, Phe^{368}, and Tyr^{407}-Glu^{413}. These residues are all in loop regions that can therefore be expected to be more flexible.

Structure determination was achieved by standard methods. The NHBA_C1 structure is a β-barrel with 8 anti-parallel strands stabilized by a hydrogen bond network that links the β-strands (Fig. 3, *A* and *B*). This fold is very common in nature. It occurs, among others, in the lipocalin superfamily, a class of proteins that transport small hydrophobic molecules such as steroids, bilins, retinoids, and lipids (38, 39).

Dynamical Properties of NHBA—The dynamical properties of NHBA_C1 were probed by NMR relaxation studies. The T_1, T_2, and NOE values are relatively flat over the range from residue 315 to the end with some minor variability in correspondence to loops (supplemental Fig. S5). Also there are a few unassigned residues in the loops, suggesting that they have quite different dynamic properties and are probably in a conformational exchange. The T_1 and T_2 values of residues 279–314 present larger variability. Among this range, residue 287 is by
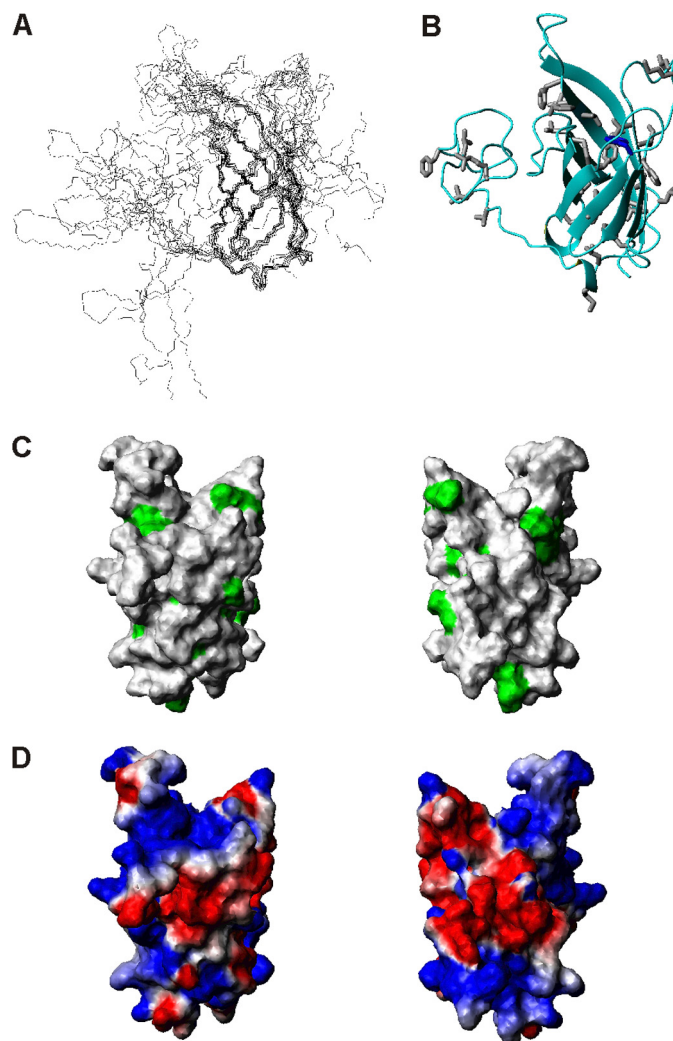
FIGURE 3. **Structure determination of the C-terminal domain of NHBA.** *A*, bundle of the 10 best structures refined in water obtained by superimposing the backbone heavy atoms (N, Cα, and C′) of residues in structured regions (β1, 316–328; β2, 337–345; β3, 350–355; β4, 367–372; β5, 376–381; β6, 388–393; β7, 400–405; β8, 419–425). *B*, ribbon drawing of the mean structure of NHBA_C1. Structure averaging was obtained according to Ref. 50. *C*, exposed hydrophobic residues are marked in *green* on the surface of the NHBA domain. Only the structured region (residues 305 to 426) is shown for clarity. The two views differ by a 180 degree rotation around the *y* axis. *D*, electrostatic surface using the same orientations as in *C*.

far the largest outlier. The heteronuclear NOE values also support a higher mobility of the N-terminal tail, although they remain all positive with the only exception for residue 324.

Structural Homology with Other Proteins of N. meningitidis—A DALI search for structural similarities with other proteins (40) identified the B domain of the TbpB from the porcine pathogens *Actinobacillus pleuropneumoniae* and *suis* (PDB 3hol and 3pqs) as the two highest hits, having *Z* scores of 10.7–10.2 and root mean square deviations of 3.6–5.5 Å (depending on the subunit) (41, 42). These proteins contain two nonidentical repeats likely created by gene duplication. Each repeat folds in a handled β-barrel and is disposed at approximately a right angle from the other repeat. Among the highest hits (*Z* score 10.2–9.9) is also the *N. meningitidis* fHBP in both its free and factor H-bound forms (10, 11, 43). This protein, previously identified as GNA1870, has architectural features similar to

```
NHBA   QGEPAKGEMLA--GTAVYNGEVLHFHT----------------------ENGRPYPT-------RGRFAAKVDFGSKSVDGIIDSGDDLHMGTQ
3KVD   AGEHTSFDKLPEGGRATYRGTAFGS-------------------------DDA--------GGKLTYTIDFAAKQGNGKIEHLK-SPELNV
3HOLN  GNSPAKELPVN--QLLTYTGSWDFTSNANLNNEEGRPNYLNDDYYTKFIGKRVGLVSGDAKPAKHKYTSQFEVDFATKKMTGKLS---DKEKTIY
3HOLC  QGERTATDKMPKGGNYKYIGTWDAQVS---------------------KEN-NWVATADDDRKAGYRTEFDVDFGNKNLSGKLFDKNG--VNPV
NHBA          HHH    EEEEEEEEEE                          EE          EEEEEEE     EEEEEEE
3KVD          HHH    EEEEEEEEEE                          EE          EEEEEEEEE    EEEEEEE       HHH
3HOLN          EEE    EEEEEEEEEEE     HHH                  HHH          EEEEEEE      EEEEEEE    E   EEEE
3HOLC          HHH    EEEEEEEEEEEE                         EEEE         EEEEEE       EEEEEE


NHBA   KFK-AAIDG-----NGFKGTWTE----------------NGGG-DVSGRFYGPAGEEVAGKYS-YRPTDAEKGGFGVFAGKKEQDLE------
3KVD   DLAAADIKPDGKRHAVISGSVLYNQ---------------AEKG-SYSLGIFGGKAQEVAGSAEVKTVN--GIRHIGLAAKQ-----------
3HOLN  TVN-ADIRG-----NRFTGAATASDKNKGKGESYNF----FSADSQSLEGGFYGPKAEEMAGKFV-AND----KSLFAVFSAKHNG-SNVDTVRI
3HOLC  FTVDPKIDG-----NGFTGKAKTSDAGFALDSGSSRYENVKFNDV-AVSGGFYGPTAAELGGQFH-HKSE--NGSVGAVFGAKQQVKK------
NHBA         EEE       EEEEE                        EEEEEE HHHH EEEEE              EEEEE
3KVD         EEEEEEEEE   EEEEEEEEE                    EEEE EEEEEEE    EEEEEEEEEE    EEEEEEEEEE
3HOLN  EEE EEE      EEEEEEE     HHH                EEEEEEE      EEEEEEE           EEEEEEEEELL     EEE
3HOLC  EEEEEEEE     EEEEEE                          E EEEEEEE    EEEEEEE EEE    EEEEEEEEEEE
```

FIGURE 4. **Structural comparison.** *Top*, ribbon representation of proteins with significant structural similarity to NHBA as picked up and aligned by a Dali search (40). From left to right, the C-terminal domain of NHBA, x-ray structure of fHBP (PDB 3kvd), and the N- and C-terminal repeats of TbpB (PDB 3hol) and NspA (PDB 1p4t). The side chains of residues involved in factor H binding (20) or forming a conformational epitope (11) are indicated in *red* and *green*, respectively, on the fHBP structure. The side chains of residues involved in transferrin binding are indicated in *magenta* on TbpB. *Bottom*, structural sequence alignment of the subfamily containing NHBA, TbpB, and fHBP based on the indicated Protein Data Bank structures. The residues involved in immune response and in factor H and transferrin binding are indicated using the same color coding adopted in the *top panels*.

those of NHBA with a C-terminal domain consisting of an 8 β-strand anti-parallel β-barrel. Another interesting hit, albeit more distant, is the *N. meningitidis* NspA, a β-barrel outer membrane protein belonging to the Opa family of lipocalins that mediates adhesion to host cells (44).

NHBA, fHBP, and TbpB Are Likely to Share a Common Ancestor—In agreement with the DALI scores, detailed structural analysis suggests maximal similarity between NHBA, fHBP, and both β-barrels of TbpB (Fig. 4, *top panel*). NspA has instead longer strands overall and is more compact, also having shorter and more rigid loops. This is in agreement with the assumption that, *in vivo*, NspA is totally embedded in the outer membrane, whereas the other three proteins are surface-exposed. It has already been noticed that NspA, as other canonical lipocalins, has a more hydrophobic surface thus being able to establish hydrophobic interactions with the aliphatic chains of membrane phospholipids (43, 44). Accordingly, the crystal structure of NspA contains four detergent molecules bound to the barrel surface.

NHBA, fHBP, and TbpB share not only the β-barrel fold but also other more subtle structural features. These include the presence of a one turn N-terminal helix, which packs against one of the barrel extremities and a semiconserved cluster of glycines and small amino acids on the barrel facing the "handle" domain. Based on these observations, we propose that NHBA,

fHBP, and TbpB do not only share the same fold but form a distinct subfamily that has evolved from a common ancestor.

We then wondered whether the proteins could share similar functions and/or immunologic properties. A sequence alignment based on the structural similarity of only the β-barrel regions leads to sequence alignments with 22–28% homologies (Fig. 4, *bottom panel*). We used this alignment to assess the degree of conservation of residues of fHBP and TbpB important for interaction or immune response. Interaction to factor H involves an extended surface of fHBP, which involves both the handle and barrel moiety (20) (Fig. 4). None of these residues are conserved in NHBA. Substantial sequence diversity is also observed at positions that form a conformational epitope on fHBP (45). Likewise, transferrin binds the N-terminal repeat of its receptor and interacts with the tip of the barrel as demonstrated by mutations of Phe[58], Lys[61], and Phe[171], which are sufficient to abolish or drastically reduce binding (41). Of these residues only the latter is in the barrel and is absent in NHBA. We must therefore conclude that, despite the common origin, the three proteins have significantly diverged and gained functional specificity.

Exploring the Role of Detergents on the NHBA Fold—We then wondered about the role of lipids in stabilizing the NHBA fold. To get further clues on the matter, ^{15}N-labeled NHBA_C1 was titrated with different detergents to identify the region(s)
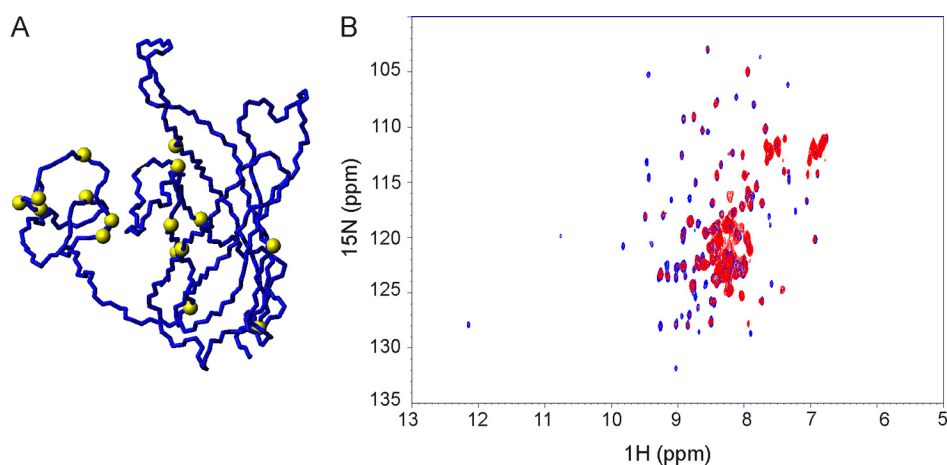
FIGURE 5. **Titration of NHBA constructs.** *A*, residues involved in detergent binding map on the surface of the NHBA_C1 structure as identified by chemical shift perturbation. The Cα carbons of the residues affected are indicated by *yellow spheres*. *B*, overlay of the HSQC of ^{15}N-labeled NHBA_C2 with no heparin (*blue*) and with 1:10 heparin:protein molar equivalents (*red*). The spectra were acquired in DPC at 40 °C and 600 MHz. The titration was stopped when no additional changes were observed in two consecutive experiments.

involved in interaction. DPC titration produced the maximal chemical shift perturbation at the resonances corresponding to residues Ile[277], Phe[278], Ala[279], Gly[282], Arg[285], Tyr[286], Ser[298], Gly[315], Leu[324], Asp[352], Gly[391], Ala[397], Glu[400], Gly[403], Gly[418], and Gly[422]. These data, together with ^{15}N,^{13}C filtered experiments using double labeled protein and unlabeled DPC, confirmed that detergent binds at the interface between the β-barrel and N terminus (Fig. 5*A*). The regions affected by DPC binding correlate well with charged regions on the protein surface. In agreement with these observations is the small enhancement of the positive ellipticities observed in the Phe-Tyr region of the near UV CD spectrum (Fig. 2*B*). This effect suggests that the exposure of these aromatic residues is affected by detergent binding.

These results confirm that DPC does not insert into the barrel cavity as has been observed in other lipocalins (38, 39) and suggests that its role in the full-length protein could be that of stabilizing the disordered N terminus and promoting the adoption of a similar structure to that observed in the other handled barrels of the subfamily.

Understanding the Involvement of NHBA in Heparin Binding— Finally, we tested our NHBA constructs for heparin binding. We used low ionic strength to enhance any effect. The spectrum in DPC of NHBA_C1 remains invariant up to an excess of a 5:1 heparin:protein ratio, in agreement with previous plasmon resonance studies suggesting that heparin binding requires the arginine-rich region around residues 235–245 (24). These residues are absent in NHBA_C1. Addition of even small amounts of heparin to NHBA_C2 causes instead considerable reduction of the peak intensities as expected upon formation of an increased molecular weight complex and disappearance of several resonances (Fig. 5*B*). The effect reaches a plateau already at about 1:10 heparin:protein molar equivalents. These results confirm previous studies and show that the arginine-rich region is necessary and sufficient for binding.

DISCUSSION

The standing threat constituted by meningitis keeps imposing the need of new vaccines able to immunize children and young adults against all serogroups of the *Neisseria* pathogen. Although effective vaccines are now available against meningitis serogroups A, C, W135, and Y, MenB is still an unsolved medical problem (4). MenB vaccines have been recently proposed for clinical development (46, 47). A multicomponent 4CMenB vaccine shown to be safe and immunogenic in adults, adolescents, and infants contains three main vaccine antigens, NadA, fHBP, and NHBA in combination with outer membrane vesicles (21). The genes coding for these antigens have been analyzed in a large panel of strains collected worldwide and the antigens studied for their expression, functional, and immunogenic properties, although only the structure of fHBP has been so far available (48, 49).

Here, we have structurally characterized the C-terminal region of NHBA in the long term attempt of understanding the structure/ immunogenicity relationship of this protein and, more in general, the factors that make a protein a good potential vaccine candidate. Through sequence analysis, we show that NHBA contains two distinct regions: an N terminus with features of an intrinsically unfolded region, and a structured C-terminal globular domain with detectable homology with other *Neisseria* proteins. Interestingly, whereas the N terminus shows some sequence variability, the domain is well conserved among MenB strains.

We have experimentally tested these predictions by characterizing constructs of different lengths spanning the C-terminal domain and confirmed that it is stably folded in an 8-stranded β-barrel. This fold is typical of lipocalins and is shared by other *N. meningitidis* proteins, such as the homologous fHBP and TbpB proteins and the outer membrane protein NspA. Is this similarity due to divergent or convergent evolution, and which role does a β-barrel structure play for the cellular role of these proteins and their immunogenicity? Subtle but clear indications indicate that NHBA closely shares its detailed structural features with fHBP and TbpB suggesting a common ancestor. NspA and other similar outer membrane proteins of *N. meningitidis*, on the other hand, seem to form a distinct family: the barrel of NspA is more compact, with shorter and more rigid loops and, as already previously noticed, more exposed hydro-

phobic residues. Structural similarity of NHBA with this protein is therefore suggestive of a convergent evolution. Nonetheless, despite the similarity, no common functional behavior is likely to be shared between NHBA, fHBP, and TbpB: the surfaces involved in partner recognition in fHBP and TbpB are not conserved in NHBA so that this protein must have evolved to fulfill a distinct role.

Could the difficulties encountered in structure determination of NHBA reflect its functional role? All tested NHBA constructs were mixtures of different populations regardless of the experimental conditions and the chosen domain boundaries. This is strongly at variance with fHBP, which posed no difficulties in an analogous study (10, 43). It was only by studying NHBA in a lipid environment that we could finally overcome sample heterogeneity and obtain a single species. Our data thus suggest that when expressed on the meningococcus surface the protein binds a yet unidentified co-factor that, in the recombinant protein, does not quantitatively resist purification thus leading to different protein populations. This possibility and the hypothesis that the co-factor is a lipid are well in agreement with the evidence that NHBA is a lipoprotein attached to the bacterial lipid surface but otherwise surface exposed. Furthermore, NHBA is immunogenic in patients recovering from the disease thus suggesting that it is expressed during infection. Therefore, we cannot exclude that interaction with lipids might also be important for its functionality within the host. Identification of the co-factor will be valuable to shed further light on the structure-function relationship of this protein. Another possibility in no way mutually exclusive with the previous one is that, in the full-length protein, distal regions of the N terminus fold by packing against the C-terminal domain and stabilize the fold.

Also interesting will be understanding whether the NHBA proteolytic fragments have any activity in the interaction with the host cells. NHBA is able to induce bactericidal activity against homologous and heterologous strains, suggesting the presence of multiple epitopes. Because previous studies have indicated that cleavage of the C-terminal domain does not influence susceptibility of the strain to bacterial killing (24), many of the bactericidal epitopes are likely to be located in the N terminus of the protein. However, it cannot be excluded that NHBA cleavage and full-length synthesis are dynamic processes so that the full-length protein is always present on the bacterial surface as a target for the bactericidal antibodies. Further analysis to map the bactericidal epitopes is currently ongoing in our laboratories.

In conclusion, by adding a new member to the family of lipoproteins with a β-barrel fold we have provided new evidence about the importance of this fold in the immunogenic properties of *N. meningitidis* proteins. Future studies, which can now be guided by the structural characterization of NHBA presented here, will bring new insights into the relationship between structural, functional, and immunological properties of this important antigen

REFERENCES

1. Stephens, D. S., Greenwood, B., and Brandtzaeg, P. (2007) *Lancet* **369,** 2196–2210
2. Caugant, D. A., and Maiden, M. C. (2009) *Vaccine* **27,** Suppl. 2, B64–70
3. Virji, M. (2009) *Nat. Rev. Microbiol.* **7,** 274–286
4. Pace, D. (2009) *Expert Rev. Vaccines* **8,** 529–542
5. Finne, J., Bitter-Suermann, D., Goridis, C., and Finne, U. (1987) *J. Immunol.* **138,** 4402–4407
6. Holst, J., Martin, D., Arnold, R., Huergo, C. C., Oster, P., O'Hallahan, J., and Rosenqvist, E. (2009) *Vaccine* **27,** Suppl. 2, B3–12
7. Bai, X., Findlow, J., and Borrow, R. (2011) *Expert Opin. Biol. Ther.* **11,** 969–985
8. Giuliani, M. M., 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., and Pizza, M. (2005) *Infect. Immun.* **73,** 1151–1160
9. Desvaux, M., Parham, N. J., and Henderson, I. R. (2004) *Curr. Issues Mol. Biol.* **6,** 111–124
10. Cantini, F., Veggi, D., Dragonetti, S., Savino, S., Scarselli, M., Romagnoli, G., Pizza, M., Banci, L., and Rappuoli, R. (2009) *J. Biol. Chem.* **284,** 9022–9026
11. Mascioni, A., Bentley, B. E., Camarda, R., Dilts, D. A., Fink, P., Gusarova, V., Hoiseth, S. K., Jacob, J., Lin, S. L., Malakian, K., McNeil, L. K., Mininni, T., Moy, F., Murphy, E., Novikova, E., Sigethy, S., Wen, Y., Zlotnick, G. W., and Tsao, D. H. (2009) *J. Biol. Chem.* **284,** 8738–8746
12. Capecchi, B., Adu-Bobie, J., Di Marcello, F., Ciucchi, L., Masignani, V., Taddei, A., Rappuoli, R., Pizza, M., and Aricò, B. (2005) *Mol. Microbiol.* **55,** 687–698
13. Litt, D. J., Savino, S., Beddek, A., Comanducci, M., Sandiford, C., Stevens, J., Levin, M., Ison, C., Pizza, M., Rappuoli, R., and Kroll, J. S. (2004) *Infect Dis.* **190,** 1488–1497
14. 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
15. Fletcher, L. D., Bernfield, L., Barniak, V., Farley, J. E., Howell, A., Knauf, M., Ooi, P., Smith, R. P., Weise, P., Wetherell, M., Xie, X., Zagursky, R., Zhang, Y., and Zlotnick, G. W. (2004) *Infect. Immun.* **72,** 2088–2100
16. Jacobsson, S., Mölling, P., and Olcén, P. (2009) *Vaccine* **27,** 5755–5759
17. 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
18. Schneider, M. C., Exley, R. M., Chan, H., Feavers, I., Kang, Y. H., Sim, R. B., and Tang, C. M. (2006) *J. Immunol.* **176,** 7566–7575
19. 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.(2009) *J. Mol. Biol.* **386,** 97–108
20. Schneider, M. C., Prosser, B. E., Caesar, J. J., Kugelberg, E., Li, S., Zhang, Q., Quoraishi, S., Lovett, J. E., Deane, J. E., Sim, R. B., Roversi, P., Johnson, S., Tang, C. M., and Lea, S. M. (2009) *Nature* **458,** 890–893
21. 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. U.S.A.* **103,** 10834–10839
22. Donnelly, J., Medini, D., Boccadifuoco, G., Biolchi, A., Ward, J., Frasch, C., Moxon, E. R., Stella, M., Comanducci, M., Bambini, S., Muzzi, A., Andrews, W., Chen, J., Santos, G., Santini, L., Boucher, P., Serruto, D., Pizza, M., Rappuoli, R., and Giuliani, M. M. (2010) *Proc. Natl. Acad. Sci. U.S.A.* **107,** 19490–19495
23. 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., 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,** 5459, 1816–1820

24. Serruto, D., Spadafina, T., Ciucchi, L., Lewis, L. A., Ram, S., Tontini, M., Santini, L., Biolchi, A., Seib, K. L., Giuliani, M. M., Donnelly, J. J., Berti, F., Savino, S., Scarselli, M., Costantino, P., Kroll, J. S., O'Dwyer, C., Qiu, J., Plaut, A. G., Moxon, R., Rappuoli, R., Pizza, M., and Aricò, B. (2010) *Proc. Natl. Acad. Sci. U.S.A.* **107,** 3770 –3775

25. Esposito, V., Musi, V., Veggi, D., Pastore, A., and Pizza, M. G. (2010) *Biomol. NMR Assign.* **4,** 107–109

26. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR.* **6,** 277–293

27. Zwahlen, C., Legault, P., Vincent, S. J. F., Greenblatt, J., Konrat, R., and Kay, L. E. (1997) *J. Am. Chem. Soc.* **119,** 6711– 6721

28. Shen, Y., Delaglio, F., Cornilescu, G., and Bax, A (2009) *J. Biomol. NMR* **44,** 213–223

29. Linge, J. P., O'Donoghue, S. I., and Nilges, M. (2001) *Methods Enzymol.* **339,** 71–90

30. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. D Biol. Crystallogr.* **54,** 905–921

31. Folmer, R. H., Hilbers, C. W., Konings, R. N., and Nilges, M. (1997) *J. Biomol. NMR.* **9,** 245–258

32. Linge, J. P., Habeck, M., Rieping, W., and Nilges, M. (2004) *J. Magn. Res.* **167,** 334 –342

33. Linge, J. P., Williams, M. A., Spronk, C. A., Bonvin, A. M., and Nilges, M. (2003) *Proteins* **50,** 496 –506

34. Koradi, R., Billeter, M., and Wüthrich, K. (1996) *J. Mol. Graph.* **14,** 51–5, 29 –32

35. Tompa, P. (2005) *FEBS Lett.* **579,** 3346 –3354

36. Vuilleumier, S., Sancho, J., Loewenthal, R., and Fersht, A. R. (1993) *Bio-*

37. Strickland, E. H. (1974) *CRC Crit. Rev. Biochem.* **2,** 113–175

38. Flower, D. R. (2000) *Biochim. Biophys. Acta* **1482,** 46 –56

39. Flower, D. R., North, A. C., and Sansom, C. E. (2000) *Biochim. Biophys. Acta* **1482,** 9 –24

40. Holm, L., and Sander, C. (1996) *Science* **273,** 595– 603

41. Moraes, T. F., Yu, R. H., Strynadka, N. C., and Schryvers, A. B. (2009) *Mol. Cell* **35,** 523–533

42. Calmettes, C., Yu, R. H., Silva, L. P., Curran, D., Schriemer, D. C., Schryvers, A. B., and Moraes, T. F. (2011) *J. Biol. Chem.* **286,** 12683–12692

43. 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

44. Vandeputte-Rutten, L., Bos, M. P., Tommassen, J., and Gros, P. (2003) *J. Biol. Chem.* **278,** 24825–24830

45. Cendron, L., Veggi, D., Girardi, E., and Zanotti, G. (2011) *Acta Crystallogr. F* **67,** 531–535

46. Snape, M. D., and Kelly, D. F. (2011) *Lancet* **377,** 1809 –1810

47. Lewis, S., Sadarangani, M., Hoe, J. C., and Pollard, A. J. (2009) *Expert Rev. Vaccines* **8,** 729 –745

48. Bambini S, Muzzi, A., Olcen, P., Rappuoli, R., Pizza, M., and Comanducci, M. (2009) *Vaccine* **27,** 2794 –2803

49. Lucidarme, J., Comanducci, M., Findlow, J., Gray, S. J., Kaczmarski, E. B., Guiver, M., Kugelberg, E., Vallely, P. J., Oster, P., Pizza, M., Bambini, S., Muzzi, A., Tang, C. M., and Borrow, R. (2010) *J. Clin. Microbiol.* **47,** 3577–3585

50. Thomas, D., and Pastore, A. (2005) *Acta Crystallogr. D* **61,** 112–116

51. Lee, W., Westler, W. M., Bahrami, A., Eghbalnia, H. R., and Markley, J. L. (2009) *Bioinformatics* **25,** 2085–2087

chemistry **32,** 10303–10313