Structural comparison and epitope analysis of outer-membrane protein PIA from strains of Neisseria gonorrhoeae with differing serovar specificities

B. J. MEE,† H. THOMAS, S. J. COOKE, P. R. LAMBDEN and J. E. HECKELS*

Molecular Microbiology Group, University of Southampton Medical School, Southampton General Hospital, Southampton SO9 4XY, UK

(Received 17 May 1993; accepted 2 July 1993)

The sequences of the por genes, encoding outer-membrane protein PI, have been obtained from a number of strains of Neisseria gonorrhoeae that express PIA molecules with differing serovar specificities. The inferred amino acid sequences of the mature proteins each comprise 308 residues and show considerable homology, with the degree of sequence variation between PIA molecules being considerably less than seen previously with PIB, but more evenly distributed throughout the molecule. The positions of sequence variation are largely confined to the regions predicted to form one of eight surface-exposed loops, suggesting a more widespread distribution of potential antigenic diversity. The deduced amino acid sequences were used to synthesize peptides for epitope mapping experiments. Some epitopes responsible for serovar specificity or recognized by bactericidal monoclonal antibodies could be identified on the basis of their reactivity with simple linear peptides, whilst others recognized conformational epitopes. By comparison of sequence differences with mAb reactivity it was possible to identify regions that appear to contribute to such determinants, including separated regions of the molecule which together were required for the formation of the conformational epitopes. All the epitopes identified lie at or close to the apices of the predicted surface-exposed loops 1, 3, 6 or 8, focusing attention on these regions as accessible targets for immune attack.

Introduction

Protein I (PI) is the major protein present on the surface of Neisseria gonorrhoeae, and unlike the other major surface antigens, pili, PII and LPS, does not undergo antigenic variation during infection (Zak et al., 1984). PI appears to be essential for gonococcal viability, acting as an anion-selective porin (Douglas et al., 1981) and may also play an important role in pathogenesis, by virtue of its ability to interact directly with the cell membrane of eukaryotic cells (Blake, 1985). PI is also a major target for immune attack on the gonococcus. Monoclonal antibodies (mAbs) directed against PI have been shown to promote complement-mediated killing, to be opsonic for phagocytosis by polymorphonuclear leukocytes and to inhibit gonococcal interaction with epithelial cells (Virji et al., 1986, 1987). In addition, evidence suggests that anti-PI antibodies induced during gonococcal infection provide serovar-specific protection against gonococcal salpingitis (Buchanan et al., 1980) and partial immunity to uncomplicated infection (Plummer et al., 1989). Thus, by virtue of its stability and antigenicity PI is of considerable interest as a potential component of a gonococcal vaccine.

Although PI is stable within a strain, differences occur between strains. Biochemical and immunological studies have revealed two major structural classes of PI (PIA and PIB) which differ in their apparent molecular mass, susceptibility to proteolysis and antigenic reactivity (Sandstrom et al., 1982). Each class can be further subdivided into a number of serovars on the basis of
reactivity with panels of either PIA- or PIB-reacting mAbs (Knapp et al., 1984). Considerable information about the structure and immunochemistry of PIB has been obtained following the sequencing of the PI-encoding genes (por) from strains with different serovar specificity (Carbonetti & Sparling, 1987; Gotschlich et al., 1987; Butt et al., 1990a). Comparison of the inferred amino acid sequences revealed that despite serovar differences sequence variation was confined to two discrete regions of the molecule designated Varl and Var2 (Butt et al., 1990b). The sequence information was also used to generate a series of synthetic peptides which were used to map the epitopes recognized by protective, PIB-specific mAbs. In each case the epitopes were localized to the Var1 region. A model of the organization of PIB within the outer membrane predicts the existence of eight surface-exposed loops with the protective epitopes in Var1 located at the apex of loop 5 (van der Ley et al., 1991). A synthetic peptide corresponding to this region has been synthesized and used for immunization after coupling to a carrier protein. The resulting antibodies promoted complement-mediated killing of the homologous strain (Heckels et al., 1990).

Although studies with mAbs have demonstrated a protective effect against PIA (Virji et al., 1987), studies equivalent to those with PIB-expressing strains have been hampered by the lack of detailed structural and immunochemical information for PIA since the sequence from only one strain has been published so far (Carbonetti & Sparling, 1987; Elkins et al., 1992). In this paper we report the sequencing of the por genes from a number of strains with differing serovar specificities, analysis of the structure of the encoded PIA molecules and the use of this information to construct synthetic peptides for use in epitope mapping experiments.

Methods

Bacterial strains and growth conditions. Fresh local isolates of Neisseria gonorrhoeae strains were obtained from sources described previously (Zak et al., 1984), strain PI-35 was obtained from Dr C. Ison, St Mary's Hospital Medical School, London, and strain FA19 was from Professor P. F. Sparling, University of North Carolina, USA. All bacteria were grown on proteose peptone agar at 37 °C in 5% (v/v) CO₂ (Fletcher et al., 1986). Serotyping was performed by Dr C. Ison.

Monoclonal antibodies (mAbs). mAbs SM100, SM101, SM102 and SM103 which react with PIA from a number of gonococcal strains have been described in detail previously (Virji et al., 1987). Other mAbs used were standard serotyping antibodies (Knapp et al., 1984). The reactivity of mAbs with gonococcal strains was determined by dot blotting (Butt et al., 1990b).

Sequence analysis of PIA. The sequences of the por genes encoding PIA were determined by direct sequencing after selective amplification of the gene by the polymerase chain reaction (PCR) in which pairs of biotinylated and unmodified oligonucleotide primers were used (McGuinness et al., 1993). Three to five gonococcal colonies were picked, resuspended in 50 μl water, lysed by addition of 100 μl 0.25 M KOH and the samples were denatured at 100 °C for 5 min. The lysate was neutralized by the addition of 100 μl 0.5 M-Tris/HCl (pH 7.5), clarified by centrifugation and 25 μl taken for PCR.

The oligonucleotides used were primer 13 (5'-CCTTTGCTACGAG-3') coding for the region just upstream of the por gene start codon and antisense primer 14 (5'-GGCATAGGAAATTGTTGGTG-3'), beginning seven nucleotides from the termination codon. To obtain biotinylated primers, oligonucleotides were synthesized with a free amino group at the 5'-terminus using Aminolink-2 (Applied Biosystems) and then biotinylated with N-biotinyl-L-amino caproic acid N-hydroxysuccinimide ester (Boehringer Mannheim). Each biotinylated primer was then used in combination with its respective unmodified oligonucleotide to generate the sense and anti-sense strands corresponding to the por gene.

The final reaction mixture contained 5 μg gonococcal DNA with 0.2 mm of each deoxynucleoside triphosphate (dNTP), 1.5 mm-MgCl₂, 600 ng of each of a pair of primers and 4-5 units of Taq polymerase (Cetus) in a final volume of 50 μl. Samples were amplified by 35 cycles of PCR (1 min at 94 °C, 2 min at 45 °C and 3 min at 70 °C); on the last cycle, the denaturation step was omitted and the extension step extended by 7 min. The amplified DNA was purified by affinity adsorption onto streptavidin-linked paramagnetic beads (Dynabeads M-280 Streptavidin, Dynal). After alkaline denaturation, the immobilized strand was sequenced directly by the di-deoxy chain-termination method using Sequenase version 2.0 (USB) according to the manufacturer's instructions. Sequencing of both strands of the por gene was accomplished with a set of custom-built oligonucleotide primers.

Multiple solid-phase peptide synthesis and epitope mapping. Multiple solid-phase peptide synthesis was carried out with the aid of an Epiguide (Labsystems) using pin technology with a commercially available kit (Cambridge Research Biochemicals) as described in detail previously (Virji & Heckels, 1989; McGuinness et al., 1993). Synthesis used pentafluorophenyl activated esters of N-α-(9-fluorenylmethoxycarbonyl) (FMOC)-L-amino acids with tert-butyl side-chain protection (Milligen), except for arginine (methoxytrimethylphenylsulphonyl) side-chain) and serine and threonine, which were used as oxabenzatrazine active esters. Each synthesis was performed in duplicate and on each occasion, peptide controls were synthesized and reacted with mAb SM300 which recognizes the sequence TKDTNNNLTL, but not TKNTNNNLTL (McGuinness et al., 1993).

Immunological reactivity of the peptides was assayed by ELISA as described previously (Virji & Heckels, 1989). The solid phase peptides were re-used after bound antibody was dissociated by sonication of the pins in 1% (w/v) SDS, 0.1% (v/v) 2-mercaptoethanol in 0.1 M-sodium phosphate buffer at 60 °C for 30 min. Immunological reactivity was always observed in duplicate peptides and in assays repeated on at least two occasions.

Sequence analysis. Sequence comparisons were carried out with the GCG sequence analysis programs (Devereux et al., 1984), using the SERC sequen facility.

Results

Amplification and sequencing of por genes

Primers 13 and 14 had been used previously to amplify the por gene from gonococcal strain P9 which expresses PIB (Butt et al., 1990b). With each of the PIA-expressing strains shown in Table 1, the use of these primers in PCR
resulted in the production of DNA fragments of approximately 1.0 kb compared with the 1.1 kb fragment obtained with strain P9. Oligonucleotides prepared from regions just inside the PCR primers were used in initial sequencing reactions, subsequent primers were constructed from the emerging sequences. Fig. 1 shows an alignment of the deduced protein sequences of the mature PIA molecules, lacking the signal peptide, together with the regions predicted to be surface-exposed according to the Neisseria porin model of van der Ley et al. (1991). The mature protein sequences each comprise 308 amino acid residues and show considerable homology, but with sequence differences located throughout the molecule. The quantitative measurement shown in

Table 1. Gonococcal PIA strains used and reactivity with mAbs

<table>
<thead>
<tr>
<th>Strain</th>
<th>Serovar</th>
<th>4A12</th>
<th>4G5</th>
<th>2F12</th>
<th>6D9</th>
<th>5G9</th>
<th>5D1</th>
<th>9D2</th>
<th>SM100</th>
<th>SM101</th>
<th>SM102</th>
<th>SM103</th>
</tr>
</thead>
<tbody>
<tr>
<td>SU89</td>
<td>IA-1</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>SU98</td>
<td>IA-1</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>SU70</td>
<td>IA-1</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>SU96</td>
<td>IA-2</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>SU99</td>
<td>IA-10</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>SU102</td>
<td>IA-6</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>SU100</td>
<td>IA-3</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>PI-35</td>
<td>IA-8</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>SU95</td>
<td>IA-4</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Fig. 1. Amino acid sequences of mature PIA proteins from different strains. Nucleotide sequences were determined as described in Methods. The sequence of strain FA-19 is from Carbonetti et al. (1988) as corrected recently by Elkins et al. (1992). Sequences were aligned using the UWGCGL PILEUP program (Devereux et al., 1984). Boxes show regions predicted as surface-exposed loops according to the model of van der Ley et al. (1991), shaded boxes show the position of epitopes recognized by mAbs determined by reaction with synthetic peptides and by comparative sequence analysis.
which are predicted to be surface-exposed according to the model of van der Ley et al. (1991).

Reactivity with mAbs

Each of the strains which had been subjected to sequencing were reacted in dot blots with two panels of mAbs; one set of cross-reacting antibodies had been shown previously to be bactericidal for PIA strains (Virji et al., 1987) and the other set is routinely used in the determination of PI serovar (Knapp et al., 1984). The results obtained on dot blots were in general agreement with those obtained by the co-agglutination technique used to determine serovar, although one significant difference was observed. Of three serovar IA-1 strains tested that react with all mAbs in co-agglutination, two (SU89 and SU98) failed to react with mAb 6D9, whilst the third (SU70) reacted strongly with 6D9 but only weakly with mAb 4A12. Comparison of the respective amino acid sequences of strains SU89 and SU98 showed only a single amino acid difference, whilst strains SU89 and SU70 showed a total of six differences, suggesting that one or more of these must be recognized by antibody 6D9 and confirming the difference between the IA-1 strains.

Use of synthetic peptides for epitope mapping

Since all mAbs reacted with either strain SU89 or SU96 the sequences of the PIA molecules from these two
strains were used in epitope mapping experiments. A series of solid-phase dodecapeptides were synthesized to correspond to the region of reactivity in loop 6 with reactivity in Fig. 3 were synthesized with adjacent peptides differing by a single amino acid residue. The peptides were synthesized to correspond to sequences from strains SU96 (a, c, e) and SU89 (b, d, f). The results shown are from mAbs SM103 (a, b) 6D9 (c, d) and 4G5 (e, f).

Fig. 4. Determination of epitopes within loop 6, recognized by PIA-reactive mAbs. A series of dodecapeptides corresponding to the region of reactivity in loop 6 with reactivity in Fig. 3 were synthesized with adjacent peptides differing by a single amino acid residue. The peptides were synthesized to correspond to sequences from strains SU96 (a, c, e) and SU89 (b, d, f). The results shown are from mAbs SM103 (a, b) 6D9 (c, d) and 4G5 (e, f).

Antibodies tested showed significant reactivity with any of the synthetic peptides.

To define more precisely the epitopes recognized, a series of dodecapeptides centred on the region 214-219 in both strains was synthesized with adjacent peptides differing by a single amino acid residue. Antibody SM103 reacted strongly with seven adjacent peptides from strain SU96 containing the common sequence 214LTWRND219, but not with the corresponding sequences from strain SU89 containing 214LTWRDN219 (Fig. 4a). An almost identical pattern of reactivity was also seen with the mAbs SM100, SM102 and 6D9 (data not shown). In contrast, mAb 4G5 reacted strongly with a large number of peptides throughout the PIA molecule (Fig. 3c, d), the only thing in common between the reacting sequences was that they all contained a glutamine residue. None of the other
both strains centred on the sequence $^{208}$QQQ$^{210}$ and all containing at least a single glutamine (Q) residue (Fig. 4e,f).}

**Discussion**

The porin proteins of pathogenic *Neisseria* form a structurally related family of proteins. Meningococci express one of two alternative forms of an anion-selective porin, class 2 and class 3 protein, corresponding respectively to gonococcal PIB and PIA, and an additional cation-selective porin, class 1 protein which has no gonococcal equivalent (Hitchcock, 1989). With the exception of PIA, amino acid sequences are available for each protein from a number of different strains. Comparison of the sequences has produced a working model for the organization of the proteins within the outer membrane (van der Ley et al., 1991). The model predicts a series of amphipathic transmembrane $\beta$-sheets which generate eight surface-exposed loops. The membrane-bound regions are highly conserved between the different porin classes, whilst the predicted loops show extensive variation in both length and amino acid composition. Within each porin class so far examined, sequence diversity is far more restricted. Studies with PIB and meningococcal class 1 protein have revealed considerable sequence homology within each protein, with one or two discrete regions of sequence variation encoding serological specificity. However, the location of these variable regions differed between the two proteins, being located in predicted loops 5 and 6 in PIB (Butt et al., 1990b) but in loops 1 and 4 in meningococcal class 1 protein (McGuinness et al., 1990, 1993). Such variable regions are of particular importance since they are responsible not only for generating antigenic diversity but are also targets for protective antibodies (McGuinness et al., 1990; Butt et al., 1990b). In the present study therefore, the location of sequence diversity within PIA molecules is of considerable interest. Comparison of the deduced amino acid sequences of the PIA from different gonococcal strains reveals a different pattern from that seen with PIB, with variations much more widely distributed throughout the molecule (Fig. 2). However, the variations are largely confined to the regions predicted to form the surface-exposed loops in the model of van der Ley et al. (1991), suggesting the possibility of more widespread distribution of potential antigenic diversity.

The epitopes recognized by the mAbs in Table 1 are of particular interest, since they either define targets for immune attack or represent antigenic determinants responsible for PIA serovar specificity. Previous studies with mAbs directed against PIB were able to define most of the epitopes recognized as short linear peptides present at or close to the apex of predicted loop 5 (Butt et al., 1990b). Both the mAbs and polyclonal antisera raised against the corresponding synthetic peptide promoted complement-mediated killing (Heckels et al., 1990). In the case of PIA, loop 5 appears to be considerably less surface-exposed, comprising only nine residues compared with 26 residues in PIB (van der Ley et al., 1991), and shows little variation. Accordingly, the linear epitopes that could be defined on PIA were located not in loop 5, but in loop 6. With the bactericidal antibodies SM100, SM102 and SM103, all reactive peptides contained the sequence $^{218}$WRND$^{219}$ located at the apex of loop 6, so defining this as the minimum reactive epitope. Closer inspection of the reactivity reveals a pattern unlike that seen previously with PIB and class 1 protein, in that the peptides at the centre of the reactive area showed lower reactivity than those on the extremes (Fig. 4a). This pattern of reactivity suggests that the epitope is not a simple continuous region, but composed of two domains separated by a region that does not make a positive contribution to the antibody–antigen interaction. The serotyping mAb 6D9 also showed similar variation in binding across the peptides corresponding to the apex of loop 6 but with an optimal epitope defined as $^{211}$DAKL$^{219}$TWRND$^{219}$. Given the discrepancy in reactivity of this antibody between dot blot reactivity and the co-agglutination reaction normally used to determine serovar, it was interesting to note that the occurrence of this sequence in any strain correlated precisely with 6D9 reactivity in dot blots, whilst no correlation was seen with co-agglutination reaction. The additional strain, SU70, which gave a positive reaction in co-agglutination but not in dot blots, contained the substitution $^{218}$DN$^{219}$ which was also present in the PIA-2 strain SU96 which failed to react in both co-agglutination and dot blotting. Thus the dot blot results, epitope mapping and sequence analysis are in complete agreement. Further studies with a number of strains which react as serovar PIA-1 has revealed that the anomalous co-agglutination reaction is not confined to strain SU70, as they can be divided into two groups according to their reactivity with 6D9 on dot blots (data not shown). The reason for the discrepancy in co-agglutination reaction is unclear but may be a consequence of denaturation of the protein by the boiling step before reaction with antibody in this test.

The epitopes recognized by a number of the PIA-specific mAbs could not be determined by the use of synthetic linear peptides, suggesting that they recognize conformational epitopes. Nevertheless, by comparison of sequence differences with mAb reactivity it is possible to identify regions which appear to contribute to such conformational determinants. Strains SU102 and SU100 which differ by a single amino acid at position 20, belong
to different serovars, IA-3 and IA-6, respectively, due to a difference in reactivity with mAb 5D1, Strain SU102, which contains the sequence \(18\text{VAYHG}^{22}\), reacts with mAb 5D1, as do four other strains with this sequence, whilst SU100, which contains the sequence \(18\text{VAHHG}^{22}\), fails to react, as do the other strains which show differences from the \(18\text{VAYHG}^{22}\) sequence. However, this sequence alone cannot define the 5D1 epitope since it is present in one strain, PI-35, which fails to react. This strain shows two other sequence differences with the reactive SU102, one at position 218 and the other at 293, the first of which shows no correlation with reactivity in other strains, whilst the other sequence, \(291\text{GTEKF}^{295}\), is also present in all reactive strains. Thus it would appear that the conformational epitope recognized by mAb 5D1 is formed from the sequences \(18\text{VAYHG}^{22}\) and \(291\text{GTEKF}^{295}\) which occur in loops 1 and 8, respectively. This is consistent with the observations of Carbonetti et al. (1988) who engineered gonococcal strains to express hybrid PIA/PIB molecules. Only those molecules with both N and C termini from PIA reacted with mAb 5D1. Similarly, considerations of antibody reactivity and sequence diversity also identify the loop 8 sequence \(291\text{GTEKF}^{295}\) as essential for reactivity with mAb 5G9. In contrast to the mAbs which did not react with linear peptides, mAb 4G5 reacted with a large number of peptides throughout the protein, all of which contained a glutamine residue. Whilst this apparent non-specificity of reactivity might appear surprising for an antibody used to define antigenic specificity, similar properties have been observed with one of the mAbs used to define meningococcal serosubtype. In that case it was suggested that the conformation of the native protein was such that conserved regions containing the potential epitope were not exposed, so that only variations within a surface-exposed loop were responsible for generating mAb reactivity (McGuinness et al., 1990). Similar conclusions can be drawn in the present case, since the single strain which fails to react with 4G5 contains a Q to R substitution at position 120 located in loop 3, suggesting that it is this region which is exposed for reaction in the native protein. Definition of the epitope recognized by mAb SM101 was of particular interest since this antibody recognizes approximately 96% of PIA-expressing strains and is protective in \textit{in vitro} assays of killing, opsonization and gonococcal-epithelial cell interactions. Whilst no reactivity was observed with synthetic peptides, suggesting involvement of conformational epitopes, sequence comparison again implicates the involvement of loop 8, since the two strains which fail to react are the ones which show variation from the consensus \(289\text{GKTEK}^{294}\) in this region.

The epitopes within PIA represent surface-exposed regions of the molecule which are accessible to the immune system. The epitopes located in the current study all lie in regions predicted to be surface-exposed loops by the model of van der Ley et al. (1991), confirming the relevance of the model and focusing attention on PIA loops 1, 3, 6 and 8 in particular. The epitopes are located not just in the regions predicted to be loops, but at or close to the centre of such regions, corresponding to apices of the loop (Fig. 1). Similar localization of epitopes at the apex of predicted loops has been observed in the studies of both PIB and meningococcal class 1 protein (McGuinness et al., 1993; Butt et al., 1990b), suggesting that these represent the regions most exposed to the immune system. Indeed with class 1 protein, a point mutation causing a single amino acid replacement at the apex of loop 4 causes profound changes to the antigenic properties of the protein. PIA differs from PIB and class 1 protein in both the location and the greater number of regions identified as epitopes. This may be due to the fact that the regions defined as epitopes in PIB (loop 5) and class 1 protein (loops 1 and 4) are substantially shorter in PIA, presumably resulting in exposure of regions which are masked in the other proteins.

The epitope mapping studies of PIB and class 1 protein have lead to the production of bactericidal antibodies following immunization with synthetic peptides containing the epitopes (Heckels et al., 1990; Christodoulides et al., 1993). A recent study involving immunization with a synthetic peptide containing part of loop 1 from PIA produced sera which were bactericidal for several strains when tested at relatively high concentrations (Elkins et al., 1992). The current study confirms the accessibility of this region to antibody attack and identifies additional potential immunogens. In addition, the localization of epitopes at the apices of the predicted loops suggests that, as has been demonstrated with meningococcal class 1 protein (Christodoulides et al., 1993), modification of peptides to mimic more accurately the loop structure of the native protein, may result in a considerably more effective immune response.

This work was supported by a Medical Research Council Project Grant. We are grateful to Dr C. Ison, St. Mary's Hospital Medical School, for the gift of strains and for performing serovar analysis.

References


