INFECTION AND IMMUNITY, Mar. 1991, p. 1141–1147
0019-9567/91/031141-07$02.00/0

Vol. 59, No. 3

Antigenic Determinants of the Chlamydial Major Outer Membrane Protein Resolved at a Single Amino Acid Level

GUANGMING ZHONG AND ROBERT C. BRUNHAM*

Department of Medical Microbiology, Faculty of Medicine, University of Manitoba, Winnipeg, Manitoba, Canada R3E 0W3

Received 17 September 1990/Accepted 22 December 1990

Antigenic determinants were identified from seven chlamydial major outer membrane proteins by using overlapping hexapeptides and polyclonal antisera. Sixty-one determinants were detected, and 30 were surface exposed on the native organisms. The two negatively charged residues, aspartic acid and glutamic acid, were found most often in determinants. Thirteen antigenic sites were further characterized by alanine substitution. Differences in fine specificities of these linear determinants were observed in alanine substitution profiles. Five determinants had adjacent critical residues, while eight had critical residues alternated with noncritical residues. Complete replacement analysis of two antigenic determinants provided more detailed information for elucidating the structural basis of the specificity of antigen-antibody interaction and suggested a correlation between sequence conservation and tolerance to amino acid substitution for antigenic sites subject to intense immune selection pressure.

The major outer membrane protein (MOMP) of *Chlamydia trachomatis* is a transmembrane protein which is speculated to have four surface-exposed regions and five transmembranous segments (4, 16). The four surface-exposed regions contain virtually the entire amino acid sequence variation noted among different *C. trachomatis* serologic variants (4, 18). The sequence-variable regions are also the major sites of antigenicity (21, 22). Several of these epitopes elicit neutralizing antibodies as assayed both in cell cultures and in animal models (14, 19, 20). The interesting immunological features of this protein make it attractive for detailed antigenic characterization.

Antigenic sites can be divided into two structural categories: sequence-continuous and sequence-discontinuous epitopes (5). The best way to characterize antigenic sites and to study antigen-antibody (Ag-Ab) interaction is by X-ray crystallography. X-ray crystallographic studies have resolved the three-dimensional structure of a number of Ag-Ab complexes and shown that epitopes occupy areas of about 600 to 700 Å and are composed of 15 to 22 amino acid residues (1, 2, 6, 7, 15). The antigen-Fab combining site interface is characterized by exclusion of water, high complementarity of fit, and hydrogen bond formation (15). A smaller subset of five to six contacting amino acid residues contributes the majority of binding energy, with the remaining residues providing complementarity (13). Critical amino acids can be assembled from distant regions in the primary sequence or be sequential in the primary sequence as in surface-exposed loops seen with lysozyme (7). Thus, interacting residues constituting an antigenic determinant might be detectable with synthetic peptides as long as the synthetic peptides cover all the critical residues (3) and have proper folding (10). Compared with the difficulties inherent in the production of large numbers of single-site protein mutants and the complexity of detailed X-ray studies, peptide synthesis is a relative simple way to carry out studies of this type.

We previously employed solid-phase peptide synthesis, as described by Geysen et al. (9), to comprehensively map antigenic sites on MOMPs from two different *C. trachomatis* serovars (21, 22). We now extend analysis to five additional *C. trachomatis* serovars. MOMP was chosen as a model system for systematic antigenic analysis for two reasons: (i) antigenic structural analysis of MOMPs might provide information relevant to vaccine development, and (ii) the regionally discrete segments of antigenicity are an interesting system in which to correlate sequence variation with antigenic variation. These antigenic determinants contain segments of amino acid conservation and variation and can serve as a useful model for study of the specificity of Ag-Ab interactions.

MATERIALS AND METHODS

Polyclonal antiserum production. Antisera to elementary bodies from *C. trachomatis* serovars A (clinical isolate in our own laboratory, isolated from the eye of a child with ocular trachoma and which we designate as UM1/OT), H (UW43/Cx), I (UW12/Ur), J (UW36/Cx), and L_2 (434/Bu) were raised in New Zealand White female rabbits (weight, 2 to 2.5 kg) (Roger Tessier, St. Pierre-Jolys, Manitoba, Canada) as previously described (22). Two rabbits were immunized with each *C. trachomatis* serovar. Antisera to serovar B and C elementary bodies were previously produced (21, 22).

Synthesis of support-coupled peptides. Peptides were synthesized on solid polyethylene rods by using a commercially available kit (Cambridge Research Biochemicals, Cambridge, United Kingdom) (9) as described previously (22). Successful synthesis was ensured by the simultaneous synthesis of positive and negative control pins and the comparison of the reactivity with test monoclonal antibodies to pins supplied by the manufacturer on which control peptides had been synthesized. Three control hexapeptides (ACDEFG, HIKLMN, and PQRSTV) containing difficult coupling residues were synthesized with each synthesis schedule and subjected to amino acid analysis to evaluate the success of the synthesis. Peptide sets of three types were synthesized. The first consisted of all possible overlapping peptides in a length of 6-mer, homologous with the amino acid sequences of entire serovar A MOMP (375 residues), entire serovar L_2

* Corresponding author.

MOMP (372 residues), and regions covering four sequence-variable domains (VDs) from serovar H, I, and J MOMPs (18). The second type of peptide set comprised alanine-substituted analogs of the parent peptide which differed from the parent by a single alanine substitution at each position. The third peptide set comprised peptides in which each residue was sequentially replaced with each of the 19 alternative amino acids.

Detection by ELISA of antibody binding to rod-coupled peptides. The three sets of immobilized peptides were assayed by enzyme-linked immunosorbent assay (ELISA) with appropriate antisera as described earlier (22). The results were expressed in optical density (OD) values at 405 nm for the first set of overlapping hexapeptides, and for the second and third sets of peptides OD values were further calculated as percentages of the parent control peptides. The solid-phase peptides were reused after the rods were freed of bound antibody by sonication for 30 min at 60°C in a solution containing 1% sodium dodecyl sulfate, 0.1% 2-mercaptoethanol, and 0.1 M sodium phosphate (pH 7.2).

Absorption of rabbit antisera. Aliquots of antisera used for the first set of overlapping hexapeptides were also absorbed with native organisms from homologous serovars before assay. Comparison of the bindings of nonabsorbed and absorbed antisera allowed determination of surface-exposed epitopes on the native organism (22).

RESULTS

Mapping antigenic determinants of MOMPs with overlapping hexapeptides. Complete overlapping hexapeptides covering the entire MOMP sequences or regions of MOMP sequences were used to determine peptide-antibody interactions in peptide ELISA. The hexapeptide immunoreactivity is shown in Fig. 1. Most of the detectable antigenic sites were localized within the four VDs of MOMPs irrespective of the origin of the MOMP sequence. Only three exceptions were found. Peptide 94 in the constant region of serovar B MOMP and peptides 210 and 212 in constant regions of serovar L_2 MOMP were found to react with corresponding antisera. All four VDs were not equally immunoreactive. Peptides from VDI and VDIV were generally most immunoreactive, although VDI of serovar B MOMP was an exception.

Because of their immunodominance, we focused attention on the four VDs of MOMPs. Since the hexapeptides completely overlap with a single amino acid shift, we determined critical residues for each antigenic site. To better differentiate adjacent antigenic determinants, we used individual antiserum as well as pooled antisera to map overlapping hexapeptide sets. Antigenic determinants mapped within the four VDs from all seven tested MOMPs are listed in Table 1. A total of 61 different antigenic determinants were identified in the 28 regions of MOMPs. Antigenic determinants ranged from two to six residues in length, with an average of five residues. Thirty determinants were antibody accessible on the surface of native organisms, as determined by absorption experiments. Generally the determinants in VDI, VDIII, and VDIV C termini of C serogroup MOMPs are surface exposed, whereas the determinants in VDII and VDIV N termini of B serogroup MOMPs are surface exposed. Serovar L_2 MOMP has the most surface-exposed determinants. A total of 615 amino acids are included in the analyzed sequences, and 218 of the amino acids were found in an antigenic determinant at least once. The frequency of appearance for a given amino acid in the antigenic determinant

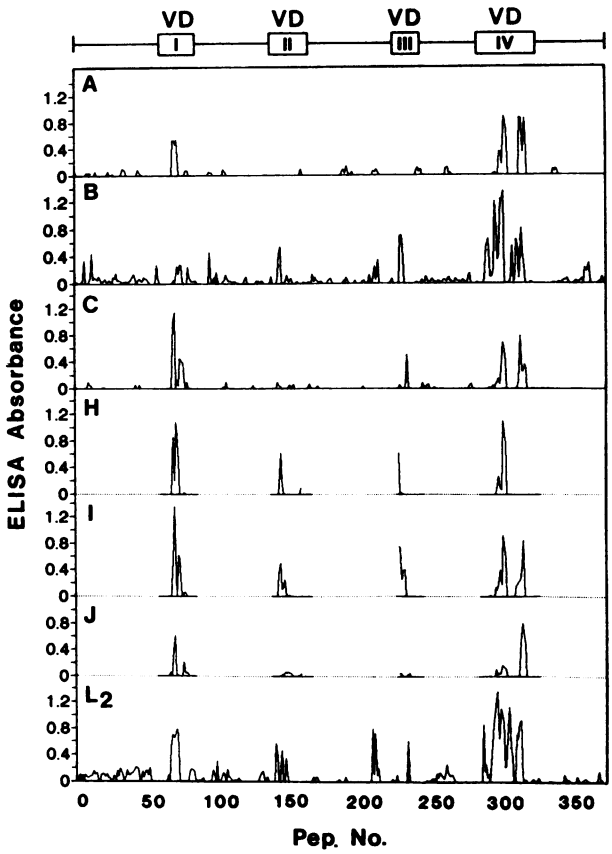
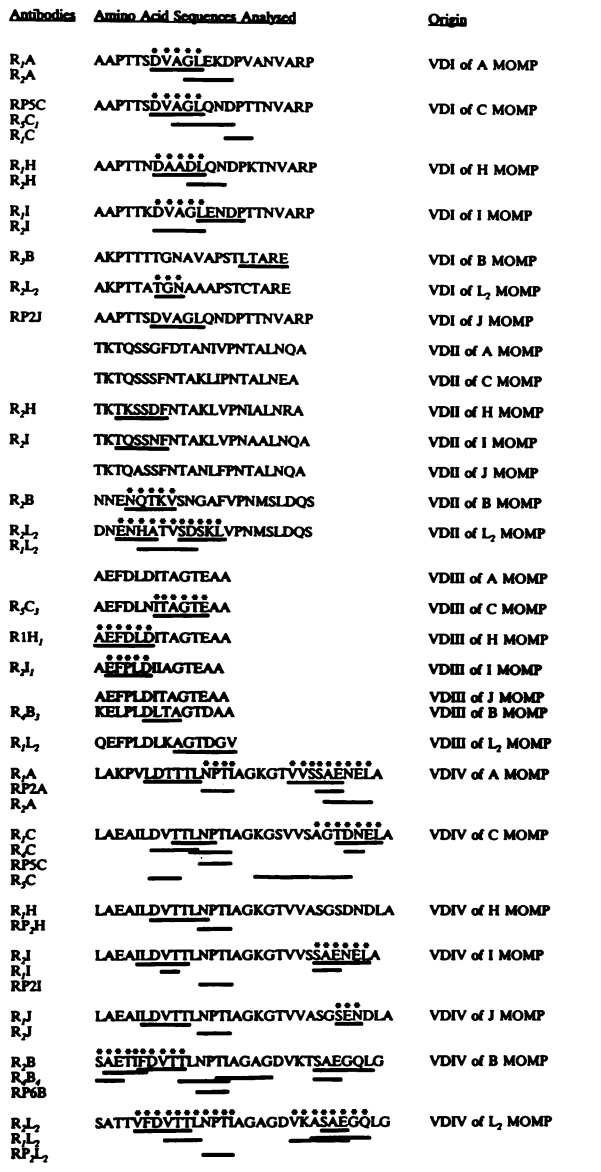


FIG. 1. Profiles of antigenic responses of serovar A, B, C, H, I, J, and L_2 MOMPs (as indicated in the figure) as a function of corresponding hexapeptides. For comparison, results for serovar B and C MOMPs published previously (21, 22) were also included in this figure. Pooled antisera (at equal ratios) from seven rabbits for serovar B MOMP, five rabbits for C MOMP, and two rabbits for A, H, I, J, and L_2 MOMPs were used. All antisera were used at a dilution of 1:200. Regions corresponding to four sequence-variable domains in MOMPs are indicated above the figure. Along the *y* axis, ELISA absorbance (in OD values) is displayed. Hexapeptides numbered at their first residue position in corresponding MOMPs are shown along the *x* axis. All MOMP sequences with different numbers of residues are aligned to match their maximum homology, and peptides covering the four variable domains from serovar H, I, and J MOMPs are aligned to corresponding regions of serovar C MOMP (18).

relative to its overall frequency in the four VDs was calculated according to a Chou-Fasman calculation of residue frequency appearing in secondary structure of a protein (data not shown). Residues found to occur most often in the antigenic region, relative to their overall occurrence in the sequences, were the negatively charged, hydrophilic amino acids aspartic acid and glutamic acid, the aromatic amino acid phenylalanine, and neutral amino acids, including the polar residues serine and asparagine and the hydrophobic residues isoleucine and valine. Those occurring least often were two positively charged amino acids, lysine and arginine, and the neutral amino acids proline and alanine. Not analyzed were tryptophan, tyrosine, cysteine, histidine, and methionine, because their overall occurrence in the sequences was too infrequent. The relative frequency of amino acid appearance varied by only fivefold, and thus, even the

TABLE 1. Antigenic determinants identified in MOMP segments[a]



Antibodies	Amino Acid Sequences Analyzed	Origin
R_1A R_2A	AAPTTSDVAGLEKDPVANVARP	VDI of A MOMP
RP5C R_5C_3 R_1C	AAPTTSDVAGLQNDPTTNVARP	VDI of C MOMP
R_1H R_2H	AAPTTNDAADLQNDPKTNVARP	VDI of H MOMP
R_1I R_2I	AAPTTKDVAGLENDPTTNVARP	VDI of I MOMP
R_3B	AKPTTTTGNAVAPSTLTARE	VDI of B MOMP
R_3L_2	AKPTTATGNAAAPSTCTARE	VDI of L_2 MOMP
RP2J	AAPTTSDVAGLQNDPTTNVARP	VDI of J MOMP
	TKTQSSGFDTANIVPNTALNQA	VDII of A MOMP
	TKTQSSSFNTAKLIPNTALNEA	VDII of C MOMP
R_2H	TKTKSSDFNTAKLVPNIALNRA	VDII of H MOMP
R_2I	TKTQSSNFNTAKLVPNAALNQA	VDII of I MOMP
	TKTQASSFNTANLFPNTALNQA	VDII of J MOMP
R_2B	NNENQTKVSNGAFVPNMSLDQS	VDII of B MOMP
R_2L_2 R_1L_2	DNENHATVSDSKLVPNMSLDQS	VDII of L_2 MOMP
	AEFDLDITAGTEAA	VDIII of A MOMP
R_5C_3	AEFDLNITAGTEAA	VDIII of C MOMP
$R1H_1$	AEFDLDITAGTEAA	VDIII of H MOMP
R_2I_1	AEFPLDIIAGTEAA	VDIII of I MOMP
	AEFPLDITAGTEAA	VDIII of J MOMP
R_3B_1	KELPLDLTAGTDAA	VDIII of B MOMP
R_1L_2	QEFPLDLKAGTDGV	VDIII of L_2 MOMP
R_1A RP2A R_2A	LAKPVLDTTTLNPTIAGKGTVVSSAENELA	VDIV of A MOMP
R_1C R_2C RP5C R_5C	LAEAILDVTTLNPTIAGKGSVVSAGTDNELA	VDIV of C MOMP
R_1H RP_2H	LAEAILDVTTLNPTIAGKGTVVASGSDNDLA	VDIV of H MOMP
R_2I R_1I RP2I	LAEAILDVTTLNPTIAGKGTVVSSAENELA	VDIV of I MOMP
R_2J R_1J	LAEAILDVTTLNPTIAGKGTVVASGSENDLA	VDIV of J MOMP
R_2B R_1B_1 RP6B	SAETIFDVTTLNPTIAGAGDVKTSAEGQLG	VDIV of B MOMP
R_2L_2 R_1L_2 RP_2L_2	SATTVFDVTTLNPTIAGAGDVKASAEGQLG	VDIV of L_2 MOMP

[a] Sequences from four VDs of all seven scanned MOMPs. Residues composing the minimum antigenic determinant are underlined, and sequences which are surface exposed on native organisms are marked with stars above the corresponding residues. All antisera were raised with the whole organisms of corresponding serovars, and individual rabbit serum (for example, R_2L_2 represents antiserum from rabbit number 2 immunized with serovar L_2 organisms) as well as pooled serum (for example, RP_2L_2 represents pooled serum from two rabbits immunized with serovar L_2 organisms) was used at a 1:200 dilution in peptide ELISA. Only the antigenic determinants identically recognized by all of the individual sera are underlined along with pooled antiserum.

least frequently found amino acid still played a key role in some antigenic determinants.

Analysis of antigenic determinants by using alanine-substituted analogs. Thirteen antigenic sites identified by overlapping hexapeptides as sequence continuous and unique at the amino acid level were selected from Table 1 to generate complete series of alanine substitution analogs. Each position was separately replaced with alanine. Fine specificity differences among these antigenic determinants were observed. Eight determinants were found to have their critical residues alternated by alanine-replaceable amino acids within the bounds of the antigenic determinant. Five determinants had their critical residues adjacent to each other. Typical results for critical residue-adjacent sites are shown in Fig. 2, and results for critical residue-separated sites are shown in Fig. 3. R_1H antiserum required DAADL residues for binding in the overlapping hexapeptide profile (Fig. 2A). With alanine substitution analogs (Fig. 2B), only the first aspartic acid was critical for R_1H binding; when this aspartic acid was substituted with alanine, binding decreased to <10% of that of the parent peptide. Whether the two central alanines are required for R_1H binding is unknown, since substitutions to these residues were not generated. R_2I antiserum required EFPLD for binding as identified by overlapping hexapeptides (Fig. 2C). Alanine substitution analog profiles indicated that only FPLD are required for R_2I antibody binding. Overall, both antigenic determinants (DAADL and EFPLD) are sequence continuous in overlapping hexapeptide and alanine substitution analog profiles, although fewer residues were identified as critical by the alanine substitution analog analysis.

Figure 3 shows examples of critical residue-separated determinants characterized by alanine substitution analog analysis. By overlapping hexapeptides, the critical residues for R_1C antiserum binding are residues DPT (Fig. 3A; Table 1). Alanine substitution analogs of peptide LQNDPTTN revealed the same general determinant region covering DPT, as identified by overlapping hexapeptides, but with a significantly more detailed profile (Fig. 3B). Substituting the proline residue of the antigenic determinant with an alanine residue did not alter the ability of the antibody to recognize the analog, indicating that the proline residue serves as a spacer in the determinant. Side chains of aspartic acid and threonine may determine the specificity of the peptide-antibody interaction, since substitution of either of these two residues with alanine completely eliminated binding. R_5C antiserum recognizes antigenic determinant AGTD in an overlapping hexapeptide profile (Fig. 3C). In alanine substitution profiles, threonine was replaceable with alanine without influencing binding, whereas substitution of glycine significantly decreased binding (30% of control) and substitution of aspartic acid completely eliminated binding (Fig. 3D). Aspartic acid and glycine are the key amino acids in this antigenic determinant, whereas threonine functions as a spacer. Whether alanine is important in this antigenic determinant was not determined since substitution was not made (Fig. 3D). R_5C_3 antiserum recognizes antigenic determinant ITAGTE (Fig. 3E). Figure 3F shows the alanine substitution analog analysis of this determinant. Substitution of glycine with alanine completely eliminated binding, and substitution of isoleucine, threonine (position 3), and glutamic acid significantly decreased binding. The second threonine at position 6 is replaceable with alanine without reducing binding. These observations indicate that isoleucine at position 2, threonine at position 3, glycine at position 5, and glutamic acid at position 7 are all important for antibody binding, whereas threonine at position 5 is not required. The role of alanine at position 4 in binding in this determinant is not known, since substitution of this residue was not performed. The R_5C antiserum requires six residues for binding to KGSVVS (Fig. 3G). By alanine substitution analysis, lysine at position 1, the two serines at positions 3 and 6, and

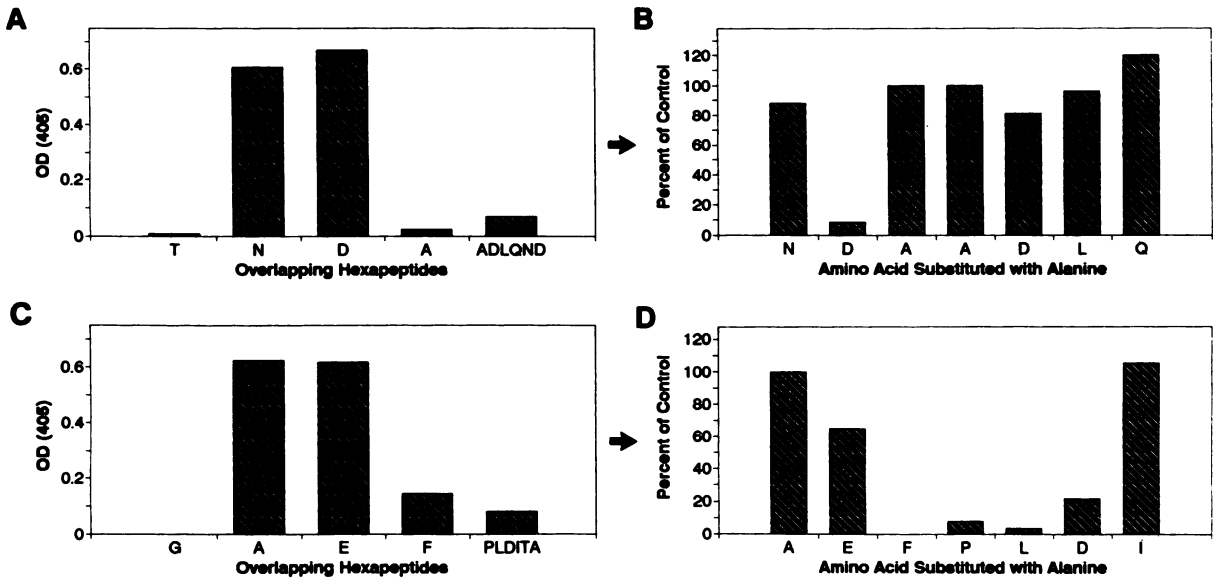


FIG. 2. Overlapping hexapeptide (A and C) and alanine substitution analog (B and D) profiles for determinant DAADL with antiserum R_1H (A and B) and for determinant EFDLD with antiserum R_2I (C and D). Overlapping hexapeptides are indicated with their first residues along the *x* axes, and the binding affinities of antibody to hexapeptides are expressed as OD values displayed along the *y* axes (A and C). The values displayed along *y* axes of panels B and D represent the percent binding of analogs relative to the parent peptide. All antisera were used at a 1:200 dilution.

the valine at position 5 were replaceable with alanine. Glycine at position 2 and the valine at position 4 were not substitutable. Therefore, the specificity of this determinant is mainly determined by glycine at position 2 and valine at position 4.

Complete substitution analysis of two antigenic determinants. To define with greater precision the role of individual amino acids in each antigenic determinant, complete amino acid replacement analogs were synthesized for two selected antigenic determinants. One antigenic site demonstrated marked amino acid variation among different *C. trachomatis* serovars, and the other demonstrated marked amino acid conservation. Each analog differs in sequence from its parent peptide by a single amino acid change. All possible analogs were synthesized by substituting the 19 different amino acids at each position. All analogs of the two antigenic determinants were separately tested for binding by antiserum homologous to the parent sequence. R_2I antibody recognizes antigenic determinant DVAGL in the overlapping hexapeptide profile (Fig. 4A). The origin of this determinant is VDI of serovar I MOMP. This region is sequence variable among different chlamydial serovars (Table 1). A complete replacement set for the 7-mer peptide which includes the 5-mer antigenic determinant DVAGL in the middle was synthesized (Fig. 4B). The N-terminal lysine (K) and C-terminal glutamic acid (E) are substitutable with most amino acids; however, histidine and lysine cannot substitute for glutamic acid. This was expected, since these residues were not critical for the antibody binding, as shown in overlapping hexapeptide profiles (Fig. 4A). Residue leucine at position 6 is also replaceable with most of the other 19 amino acids, except with proline, glycine, alanine, lysine, and aspartic acid. Residues DVAG are highly specific and not substitutable with any other amino acid, except asparagine substituting for the aspartic acid at position 2. Overall, this antigenic determinant has its critical residues adjacent to each other, and residues in the determinant are highly intolerant to substitution.

The second antigenic site chosen for complete replacement study is LNPTI (Fig. 5B). This site is conserved among all *C. trachomatis* MOMP sequences. Antibody R_1A recognizes NPTI as critical residues in the overlapping hexapeptide profile (Fig. 5A). The N-terminal leucine (L) is replaceable with any other amino acid except histidine, which was expected since the L was not critical to R_1A binding. Proline at position 3 could be replaced with most of other amino acids, except with histidine and lysine. Substitution of asparagine at position 2, threonine at position 4, and isoleucine at position 5 was much more limited because only the aromatic amino acids phenylalanine, tryptophan, and tyrosine substituted for asparagine; serine substituted for threonine; and phenylalanine, leucine, valine, and tyrosine substituted for isoleucine. Antigenic determinant LNPTI has its critical residues in a sequence-discontinuous array, although within a short region, and it is more tolerant to amino acid replacement than is determinant DVAGL.

DISCUSSION

The location and characterization of sequential (continuous) epitopes have been simplified by the development of techniques for rapid, concurrent synthesis of large numbers of peptides in a reuseable and convenient format for direct testing by ELISA (9, 22). We previously employed this technology to comprehensively scan the antigenicity of chlamydial MOMPs from two serovars (21, 22). In the present study, we used this same technique to comprehensively define additional antigenic determinants on MOMPs, to correlate sequence information with antigenic variation and conservation, and to study the specificity of Ag-Ab interactions. We found that the immunogenicity of this membrane protein is mainly located within the four se-

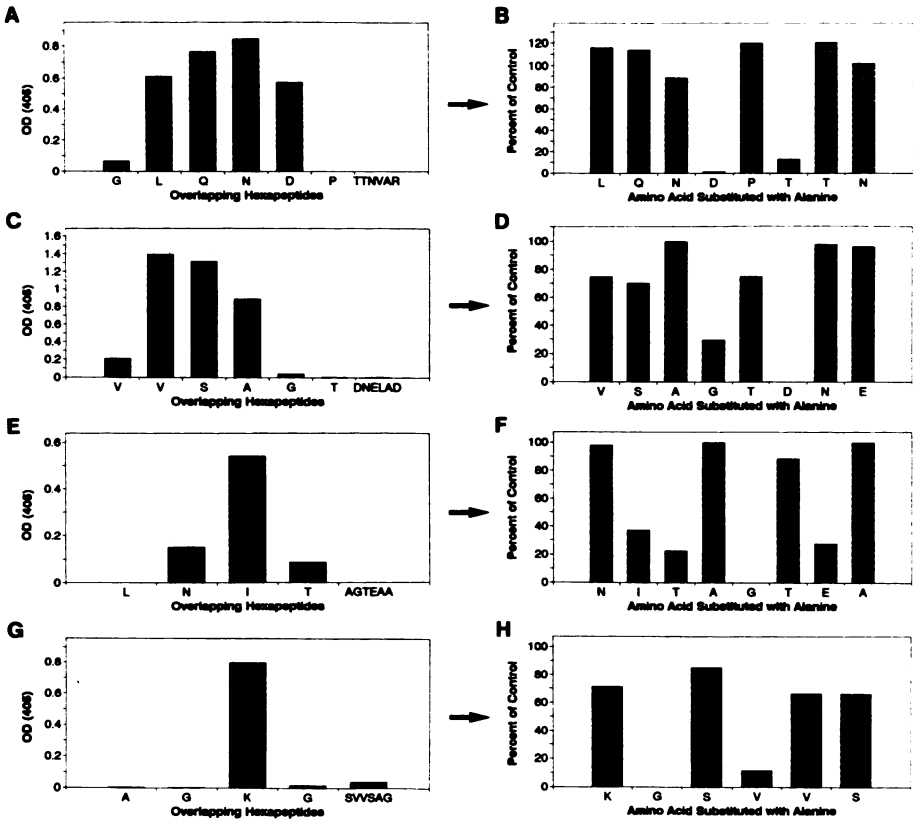


FIG. 3. Profiles of overlapping hexapeptide (A, C, E, and G) and alanine substitution (B, D, F, and H) analysis of four antigenic determinants: antigenic determinant DPT with antiserum R_1C (A and B), AGTD with R_5C (C and D), ITAGTE with R_5C_3 (E and F), and KGSVVS with R_5C (G and H). Details are described in the legend to Fig. 2.

quence-variable regions. This observation supports the concept that amino acid variability in this protein is due to selective pressure imposed by the immune system. It was also observed in many other proteins that surface-exposed regions with great sequence variability are generally more immunogenic. Complete overlapping hexapeptides covering the four VDs from seven different MOMPs (a total of 28 sequences) were assayed with rabbit antisera to identify the corresponding antigenic determinants. Among the identified 61 antigenic sites, the determinant length ranged from two to six residues with an average of five residues, although the maximum length was limited to six residues because hexapeptides were used. However, our observation is similar to what Appel et al. (3) found in the antigenicity study of oncogene products. The absorption experiments revealed distinctly different distribution patterns of surface-exposed determinants among MOMPs from different serovar. Generally, B serogroup MOMPs have VDII and N-terminal regions of VDIV surface exposed, whereas C serogroup MOMPs have VDI, VDIII, and C-terminal regions of VDIV surface exposed. The LGV serovar L_2 has the most surface-exposed regions, which might correlate with its high immunogenicity and invasiveness. These results may be important in selecting a vaccine candidate sequence, since the surface exposure of epitopes on the native organisms is a prerequisite for designing neutralizing antigenic targets. A total of 615 amino acids in 28 sequences were analyzed, and 218 amino acids were found to occur at least once in antigenic determinants. The two negatively charged residues aspartic

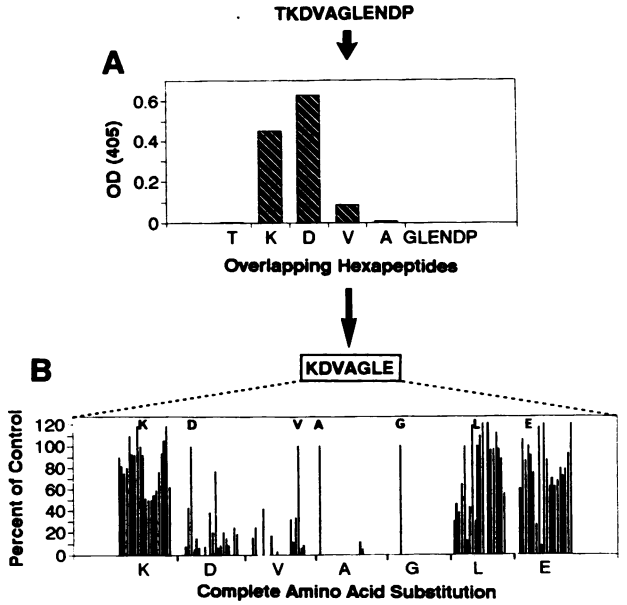


FIG. 4. Systematic approach for identifying the antigenic determinant and key binding residues for peptide KDVAGLE. (A) Overlapping hexapeptide mapping of antiserum R_2I to critical re-sidues DVAGL; (B) individual substitution analogs of peptide KDVAGLE in which each residue was individually replaced with the other 19 amino acids. Analogs for each position are in alphabetical order by using the single-letter amino acid code. Values represent the percent binding of the analog relative to that of the original control peptide.

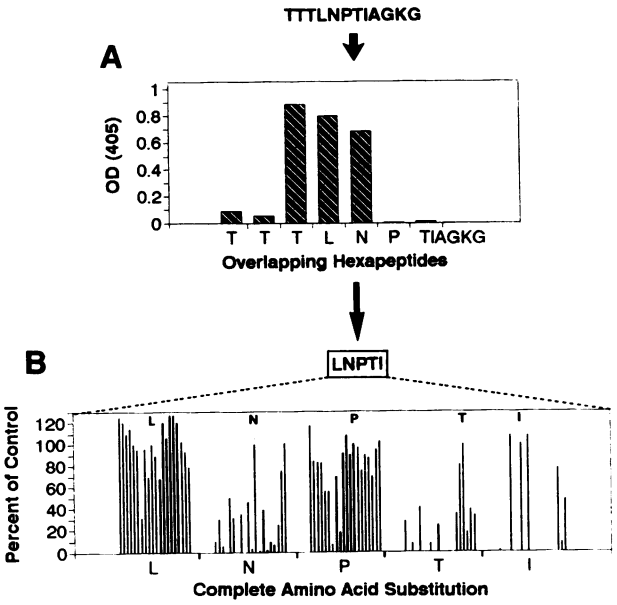


FIG. 5. Systematic approach for identifying the antigenic determinant and key binding residues for the peptide LNPTI with antiserum R_1A. Figures are shown in the same pattern described in the legend to Fig. 4.

acid and glutamic acid and the neutral residue serine appeared most often in the determinant regions. The two positively charged residues lysine and arginine and the neutral residues alanine and proline were found least often. Since the sample size in our study is limited, we cannot exclude the potential importance of less-common residues in constituting antigenic sites. Interestingly, Geysen et al. (8) had a similar finding in an independent system. They found that the highest overall propensity factor belongs to glutamic acid, whereas arginine has the lowest frequency of appearing in antigenic sites.

We used alanine substitution analogs to characterize antigenic determinants. Alanine has one methyl group as its side chain and does not exhibit preference with respect to the surface or interior location in a protein. Substituting alanine for each residue of the original peptide will maintain the main chain structure while minimizing potential conformational changes in the new analogs. Overlapping hexapeptide profiles showed that all 61 antigenic determinants were sequence-continuous determinants, in which each residue appeared to be essential to binding (Table 1). Alanine substitution analogs of 13 selected antigenic determinants produced different profiles, in which some residues could be replaced by alanine without loss of antibody recognition (Fig. 2 to 5). This suggests that amino acid residues at these positions may be necessary only to maintain the overall spacing and that the side chains of residues in these positions are neither specific nor strongly involved in antibody binding. Replacement of other residues in antigenic determinants with alanine resulted in complete loss of antibody binding. These results suggest that nonsubstitutable residues are very specific to antibody binding. Alanine substitution profiles provided finer resolution of antigenic determinants than did complete overlapping hexapeptide profiles.

Two types of antigenic determinants were identified with alanine substitution analog analysis. Five of 13 determinants had their critical residues in a sequence-continuous format; examples are determinant DAADL with antiserum R_1H and determinant EFPLD with R_2I (Fig. 2). Eight antigenic determinants had critical residues alternated with substitutable residues. Examples are DPT, AGTD, ITAGTE, and KGS VVS (Fig. 3). Therefore, alanine substitution provided a much more detailed analysis of antigenic epitopes than did overlapping peptides.

Two immunodominant epitopes were subject to complete amino acid replacement study. One (KDVAGLE) is from VDI of *C. trachomatis* serovar I MOMP, and another (TTLNPTIAG) is conserved in VDIV from all *C. trachomatis* serovars. Complete replacement analysis for both determinants showed that the replacement of residues outside the antigenic determinant as defined by overlapping hexapeptides failed to alter antibody recognition (12). In the antigenic determinant DVAGL, the C-terminal leucine was the least specific residue for peptide-antibody interaction, since 12 of 19 other amino acids could substitute for leucine without significantly affecting binding activity (Fig. 4B). However, leucine still remained important for antibody R_2I binding to this antigenic determinant, since the hexapeptide TKDVAG missing leucine failed to bind antiserum R_2I (Fig. 4A). This suggests that leucine may serve as a spacer for the peptide-antibody interaction. Conservative replacement of aspartic acid with asparagine resulted in analogs that retained about 80% of their binding ability, which suggests that the size and shape rather than the charge of the aspartic acid side chain contribute to the specificity of Ag-Ab interaction. Residues valine, alanine, and glycine are extremely specific in this antigenic determinant, since no other amino acid could substitute without reducing binding activity. Overall, this antigenic determinant has its critical residues adjacent to each other and consists mainly of small residues.

Determinant NPTI was generally less specific than DVAGL since each residue in this determinant could be replaced with several other residues. Proline was the least specific residue in this determinant since it could be substituted by any amino acid except the two positively charged residues, histidine and lysine. Proline may function as a spacer rather than offering specificity. Asparagine could be replaced with the three aromatic acids, phenylalanine, tryptophan, and tyrosine. It is unclear why aromatic residues rather than structurally similar residues such as aspartic acid could substitute asparagine without loss of binding ability. Perhaps this peptide-antibody interaction is dominated by hydrophobic forces. Threonine is only tolerant to conservative replacement with serine, which suggests that hydrogen bond formation by the threonine side chain might contribute to the specificity of the peptide-antibody interaction. Isoleucine could be replaced with the aromatic residues phenylalanine and tyrosine, again suggesting the dominance of hydrophobic forces in the peptide-antibody interaction. Isoleucine could be also replaced with the other two branched side chain residues leucine and valine, which suggests the importance of the rigidity of the isoleucine side chain in maintaining binding. Overall, complete replacement analysis revealed that antigenic determinant NPTI has its critical residues alternated by replaceable spacer residues. Hydrogen bond and hydrophobic interaction may be the main factors determining this peptide-antibody interaction specificity.

When comparing complete replacement profiles of determinant DVAGL with those of determinant NPTI, it is apparent that both determinants consist of small residues with a motif of rigid residues (V, I, or L) plus a small nonpolar residue plus a proline or glycine in which charge

formation may be less important than hydrogen bond formation and hydrophobic interactions in maintaining the peptide-antibody binding. Yet complete replacement profiles of these two determinants are distinctly different. DVAGL has its critical residues in a continuous sequence and is highly sensitive to substitutions, whereas NPTI has its critical residues distributed discontinuously and is more tolerant to residue substitutions. This suggests that there may be a correlation between sequence conservation and tolerance to amino acid substitution, since the highly conserved (conserved among only 15 *C. trachomatis* serovars) NPTI is much more tolerant to amino acid substitution than the relatively less conserved DVAGL determinant (conserved among only 4 serovars). This potentially explains the paradoxical observation of the high immunogenicity of the NPTI epitope despite its limited sequence diversification. Even within the determinant DVAGL, a correlation between substitution tolerance and amino acid conservation was observed. D and L are conserved among the seven serovars and are tolerant to at least some substitution, whereas V, A, and G are less conserved (with V found in six serovars and A and G found in five serovars) and are completely intolerant to substitution. Conservation of amino acid sequence in surface-exposed regions of MOMP seems to correlate with tolerance to amino acid substitution in antigenic analysis. As the degree of conservation grows, the sequence is more tolerant to substitution. The significance of this correlation is as yet unknown, but it may have evolutionary significance. Tolerant, conserved antigenic sequences may represent ancestral MOMP sequences nonselectable by the immune system; intolerant variable sequences may represent junior sequences more recently selected by the immune system. However, more sequences in a wider range of systems need to be analyzed in similar ways in order to confirm our hypothesis.

ACKNOWLEDGMENTS

This work was supported by a grant from the Medical Research Council. G.Z. is a recipient of a Manitoba Health Research Council Fellowship.

REFERENCES

1. **Alzari, P. M., M. B. Lascombe, and R. J. Poljak.** 1988. Three-dimensional structure of antibodies. Annu. Rev. Immunol. **6**:555–580.
2. **Amit, A. G., R. A. Mariuzza, S. E. V. Phillips, and R. J. Poljak.** 1986. Three dimensional structure of an antigen-antibody complex at 2.8 A resolution. Science **233**:747–753.
3. **Appel, R. J., C. Pinilla, H. Niman, and R. Houghten.** 1990. Elucidation of discontinuous linear determinants in peptides. J. Immunol. **144**:976–983.
4. **Baehr, W., Y. X. Zhang, T. Joseph, H. Su, E. F. Nano, E. D. K. Everett, and H. D. Caldwell.** 1988. Mapping antigenic domains expressed by *Chlamydia trachomatis* major outer membrane protein genes. Proc. Natl. Acad. Sci. USA **85**:4000–4004.
5. **Barlow, D. J., M. S. Edwards, and J. M. Thornton.** 1986. Continuous and discontinuous protein antigenic determinants. Nature (London) **322**:747–748.
6. **Colman, P. M., W. G. Laver, J. N. Varghere, A. T. Baker, P. A. Tulloch, G. M. Air, and R. G. Webster.** 1987. Three-dimensional structure of a complex of antibody with influenza virus neuraminidase. Nature (London) **326**:358–363.
7. **Davies, R. D., S. Sheriff, and E. A. Padlan.** 1988. Antibody-antigen complexes. J. Biol. Chem. **263**:10541–10544.
8. **Geysen, H. M., J. T. Mason, and J. S. Rodda.** 1988. Cognitive features of continuous antigenic determinants. J. Mol. Recognition **1**:32–41.
9. **Geysen, M. H., J. S. Rodda, J. T. Mason, G. Tribbick, and G. P. Schoofs.** 1987. Strategies for epitope analysis using peptide synthesis. J. Immunol. Methods **102**:259–274.
10. **Hirayama, A., Y. Takagaki, and F. Karush.** 1985. Interaction of monoclonal anti-peptide antibodies with lysosyme. J. Immunol. **134**:3241–3247.
11. **Hodges, R. S., R. J. Heaton, J. N. R. Parker, L. Molday, and R. S. Molday.** 1988. Antigen-antibody interaction: synthetic peptides define linear antigenic determinants recognized by monoclonal antibodies directed to the cytoplasmic carboxyl terminus of rhodopsin. J. Biol. Chem. **263**:11768–11775.
12. **Houghten, R. A., S. R. Hoffman, and H. L. Niman.** 1986. Relative importance of position and individual amino acid residues in peptide antigen-antibody interactions: implications in the mechanism of antigenic drift and antigenic shift, p. 21–25. *In* F. Brown, R. M. Canock, and R. A. Lerner (ed.), Vaccines 86: new approaches to immunization—developing vaccines against parasitic, bacterial, and viral diseases. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
13. **Laver, W. G., G. M. Air, R. G. Webster, and S. J. Smith-Gill.** 1990. Epitopes on protein antigens: misconceptions and realities. Cell **61**:553–556.
14. **Peterson, M. E., G. M. Zhong, E. Carlson, and L. M. de la Maza.** 1988. Protective role of magnesium in the neutralization by antibodies of *Chlamydia trachomatis* infectivity. Infect. Immun. **56**:885–891.
15. **Stanfield, L. R., T. M. Fieser, R. A. Lerver, and I. A. Wilson.** 1990. Crystal structures of an antibody to a peptide and its complex with peptide antigen at 2.8 A. Science **248**:712–719.
16. **Stephens, S. R., R. Sanchez-Pescador, E. A. Wagar, I. Inouye, and M. S. Urdea.** 1987. Diversity of *Chlamydia trachomatis* major outer membrane protein genes. J. Bacteriol. **169**:3879–3885.
17. **Todd, E. E. P., I. J. East, and S. J. Leach.** 1982. The immunogenicity and antigenicity of proteins. Trends Biochem. Sci. **7**:212–216.
18. **Yuan, Y., Y. Zhang, N. G. Watkins, and H. D. Caldwell.** 1989. Nucleotide and deduced amino acid sequences for the four variable domains of the major outer membrane proteins of the 15 *Chlamydia trachomatis* serovars. Infect. Immun. **57**:1040–1049.
19. **Zhang, Y. X., S. Stewart, T. Joseph, M. R. Taylor, and H. D. Caldwell.** 1987. Protective monoclonal antibodies recognize epitopes located on the major outer membrane protein of *Chlamydia trachomatis*. J. Immunol. **138**:575–581.
20. **Zhang, Y. X., S. J. Stewart, and H. D. Caldwell.** 1989. Protective monoclonal antibodies to *Chlamydia trachomatis* serovar- and serogroup-specific major outer membrane protein determinants. Infect. Immun. **57**:636–638.
21. **Zhong, G. M., and R. C. Brunham.** 1990. Immunoaccessible peptide sequences of the major outer membrane proteins from *Chlamydia trachomatis* serovar C. Infect. Immun. **58**:3438–3441.
22. **Zhong, G. M., R. E. Reid, and R. C. Brunham.** 1990. Mapping antigenic sites on the major outer membrane protein of *Chlamydia trachomatis* with synthetic peptides. Infect. Immun. **58**:1450–1455.