

Available online at www.sciencedirect.com

ScienceDirect

Biomedical Journal

journal homepage: www.elsevier.com/locate/bj

Original Article

In silico prediction of T- and B-cell epitopes in PmpD: First step towards to the design of a *Chlamydia trachomatis* vaccine

Romina Cecilia Russi ^a, Elian Bourdin ^b, María Inés García ^a,
Carolina Melania I. Veaute ^{a,*}

^a Basic Immunology Laboratory, Faculty of Biochemistry and Biological Sciences, National University of the Littoral, Santa Fe, Argentina

^b Independent professional, C1425BME, Buenos Aires, Argentina

ARTICLE INFO

Article history:

Received 2 February 2017

Accepted 25 April 2018

Available online 21 May 2018

Keywords:

Chlamydia trachomatis

Vaccine

B-cell epitope

T-cell epitope

Molecular modeling

Epitope prediction

ABSTRACT

Background: *Chlamydia trachomatis* is the most common sexually transmitted bacterial infection globally. Currently, there are no vaccines available despite the efforts made to develop a protective one. Polymorphic membrane protein D (PmpD) is an attractive immunogen candidate as it is conserved among strains and it is target of neutralizing antibodies. However, its high molecular weight and its complex structure make it difficult to handle by recombinant DNA techniques. Our aim is to predict B-cell and T-cell epitopes of PmpD.

Method: A sequence (Genbank AAK69391.2) having 99–100% identity with various serovars of *C. trachomatis* was used for predictions. NetMHC and NetMHCII were used for T-cell epitope linked to MHC I or MHC II alleles prediction, respectively. BepiPred predicted linear B-cell epitopes. For three dimensional epitopes, PmpD was homology-modeled by Raptor X. Surface epitopes were predicted on its globular structure using DiscoTope.

Results: NetMHC predicted 271 T-cell epitopes of 9–12aa with weak affinity, and 70 with strong affinity to MHC I molecules. NetMHCII predicted 2903 T-cell epitopes of 15aa with weak affinity, and 742 with strong affinity to MHC II molecules. Twenty four linear B-cell epitopes were predicted. Raptor X was able to model 91% of the three-dimensional structure whereas 57 residues of discontinuous epitopes were suggested by DiscoTope. Six regions containing B-cell and T-cell epitopes were identified by at least two predictors.

Conclusions: PmpD has potential B-cell and T-cell epitopes distributed throughout the sequence. Thus, several fragments were identified as valuable candidates for subunit vaccines against *C. trachomatis*.

* Corresponding author. Basic Immunology Laboratory, Faculty of Biochemistry and Biological Sciences, National University of the Littoral, University Campus S3000ZAA, Santa Fe, Argentina.

E-mail address: cveaute@fcb.unl.edu.ar (C.M.I. Veaute).

Peer review under responsibility of Chang Gung University.

<https://doi.org/10.1016/j.bj.2018.04.007>

2319-4170/© 2018 Chang Gung University. Publishing services by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

At a glance commentary

Scientific background on the subject

Chlamydia trachomatis infection has a high global prevalence and is associated with serious consequences on reproductive health. Antibiotic therapy is not successful at all, thus, vaccine development is strongly needed. One immunogen candidate is Polymorphic Membrane Protein D, a surface protein, highly conserved among serovars, and target of neutralizing antibodies.

What this study adds to the field

The identification of T- and B-cell epitopes on PmpD allows the selection of several regions that may be used to design subunit vaccines, potentially inducing both humoral and cellular immune responses. Bioinformatics presents powerful tools to the characterization of proteins favoring the rational design of vaccines.

Chlamydia trachomatis (Ct) is an intracellular bacterium that is an important cause of sexually transmitted infections (STI) with significant impact on public health. The World Health Organization (WHO) estimates that Ct is responsible of almost 106 million of the 500 million new cases of STI reported worldwide annually [1]. Ct includes three human biovars composed of different serovars [2] that can infect various cell types in humans. Serovars A-C are responsible for ocular infections that result in trachoma leading to blindness [3]. Serovars D-K causes sexually transmitted diseases such as cervicitis and pelvic inflammatory disease (PID), and globally are an important infectious cause of infertility, ectopic pregnancy [4] and chronic pelvic pain [5] in women. In men it is associated with urethritis, epididymitis and orchitis [6]. Moreover, serovars D-K cause urethritis and neonatal pneumonia [2]. The lymphogranuloma venereum (LGV) serovars L1–L3 not only cause sexually transmitted disease but can also infiltrate local lymph nodes, which ultimately results in systemic infection [7,8]. Ct infections can be controlled by antibiotic therapy but the lack of compliance with treatment, the persistence of the infection even after a complete treatment, together with the high prevalence of asymptomatic cases [4] leading to severe reproductive complications strongly support the development of an effective *Chlamydia* vaccine. Currently, there are no vaccines available against Ct genital infection despite the many efforts that have been made throughout the years to develop a protective one. A failure of several vaccine designs may be attributed at least in part to the fact that protective immune response may result harmful for the host and the assumption that complete microorganisms could have components that induce both a protective and an immunopathogenic response. Safety concerns may be overcome by using subunit vaccines, but they require a thorough design in order to be efficient.

Among the antigen candidates that have been studied, members of the Polymorphic Membrane Protein family (Pmp

A-I) have shown to be promising as vaccine components as they are dominant antigenic targets for cellular immune responses [9–11]. Pmps are a group of membrane bound surface exposed chlamydial proteins [12–19] that have been characterized as autotransporter adhesins. These proteins are involved in the delivery of virulence factors involved in the initial phase of chlamydial infection [2], disease progression and immune evasion [18]. As typical type V autotransporters [20], all Pmps are characterized by containing conserved GGA (I, L, V) and FxxN tetrapeptide motifs, with an amino-terminal (N-terminal) dependent leader sequence, followed by a passenger domain and a carboxy-terminal (C-terminal) β -barrel [2,21,22]. The C-terminal region is incorporated into the outer membrane, forming a pore and allowing the translocation of the N-terminal passenger domain to the bacterial surface [22]. Pmp proteins may also undergo complex infection-dependent post-translational proteolytic processing [8,17,18,23]. These proteins mediate *in vitro* chlamydial attachment to human epithelial and endothelial cells [8,16].

PmpD is the second highest conserved Pmps demonstrating a 99.1% of amino acid identity among *C. trachomatis* serovars [24], and it is a target of broadly cross-reactive neutralizing antibodies [16]. The structural features of PmpD are relatively large size (1530 aa, 160.5 kDa), integrin-binding RGD motif (aa 698 to 670), and a putative nuclear localization signal (NLS; aa 783 to 798) [18] besides of N-terminal GGA(I/L/V) and FxxN tetrapeptide repeats as in all Pmps.

Due to its conserved nature, surface localization, and immunological importance PmpD is an attractive vaccine candidate for the prevention of human infections [16]. However, it has a high molecular weight with a complex structure which makes it difficult to handle by recombinant DNA techniques. Thus, an in-depth study of the molecule is needed in order to make a rational choice of the immunogen taking into consideration the critical epitopes to induce the appropriate immunological reaction. Several authors suggest that immunity against *C. trachomatis* requires CD4⁺ T cells, (mainly Th1) with INF- γ production and, to a lesser degree CD8⁺ T cells. Besides, neutralizing antibodies are now the focus of immune protection and vaccine development [16]. One of the major difficulties in developing an effective chlamydial vaccine is identifying the B-cell epitopes and the MHC-bound chlamydial protein epitopes that are recognized by T-cells.

In this context, bioinformatic approaches can contribute to the design of epitope-based vaccines. Therefore, the aim of this study was to perform *in silico* prediction of B and T epitopes within the amino acid sequence of PmpD for the design of a subunit vaccine against *C. trachomatis*.

Materials and methods

Datasets

An amino acid sequence of *C. trachomatis* serovar L2 PmpD available from NCBI (Genbank AAK69391.2) [25] was used for computational prediction. PmpD genes are one the most conserved between *C. trachomatis* serovars (with 99.1%

similarities at the amino acid sequence level [24]. This was confirmed by Clustal Omega, an algorithm of multiple amino acid sequence alignment available at <http://www.ebi.ac.uk/Tools/msa/clustalo/>.

3D structure prediction

The online tool integrated at Raptor X (<http://raptorx.uchicago.edu/>) was used for homology modeling of the PmpD in order to predict B cell epitopes in 3D conformation. Raptor X is a protein structure prediction server for protein sequences without close homologs in the Protein Data Bank (PDB) (Källberg et al., 2012). The server uses PSI-BLAST to find homologue templates and model the 3D structure on the basis of solvent accessibility, secondary and tertiary structures, and disordered regions. The following confidence scores are used to indicate the quality of a predicted 3D model: P-value for the relative global quality and uGDT (un-normalized GDT) for the absolute global quality.

Moreover, programs for structural analysis were used to check the 3D model. ANOLEA was used for energy calculations at the atomic level in protein structure [26] and RAMPAGE to generate the Ramachandran plot, showing protein residues in the favored, allowed and outliers regions [27].

B-cell epitope prediction

Linear B-cell epitope prediction

BepiPred was used to determine B cell linear epitope of PmpD. It is a method made by combining the predictions of a hidden Markov model and the propensity scale by Parker et al. [28,29]. BepiPred assigns a score value to each protein residue. The method's specificity and sensitivity depend on the score threshold. For this analysis the threshold of 0.35 was employed. BepiPred is available at www.cbs.dtu.dk/services/BepiPred.

Discontinuous B-cell epitopes

DiscoTope was employed to predict the surface epitopes on a globular structure. The method is based on amino acid statistics, spatial information, and surface accessibility considering

Parker's hydrophilicity value [29]. These analyses are carried out taking into account a compiled data set of discontinuous epitopes determined by X-ray crystallography of antibody/antigen protein complexes [30,31]. The updated version of DiscoTope is available at www.cbs.dtu.dk/services/DiscoTope-2.0.

T-cell epitope prediction

NetMHC [32–34] version 3.4 and NetMHCII [35,36] version 2.2 were used for T-cell epitope linked to mayor histocompatibility complex (MHC) I or MHC II alleles prediction, respectively. The predictions are based on artificial neural networks (ANN). The servers are free to use and available at: <http://www.cbs.dtu.dk/services/NetMHC> and <http://www.cbs.dtu.dk/services/NetMHCII>.

Results

3D structure prediction

PmpD modeling was performed using Raptor X, which makes alignments between a target sequence and one or multiple distantly related template proteins (especially those with sparse sequence profiles) and employs a nonlinear scoring function and a probabilistic-consistency algorithm [37].

The Raptor X software modeled 1398 residues of 1530 (91%) of the PmpD sequence and identified four domains. To model domains 1 and 4, Raptor X used the passenger domain of the *Escherichia coli* autotransporter EspP (PDB entry: 3sze A) as a template, rhamnogalacturonase A from *Aspergillus aculeatus* (PDB entry: 1rmg A) as a template of domain 2, and autotransporter EstA from *Pseudomonas aeruginosa* (PDB entry: 3kvn A) as a template of domain 3. The software predicts that the protein has β sheets (31%), some α helices structure (9%) and several loops (58%) (Fig. 1A). According to the prediction, a 34% of PmpD is exposed to the surface while a 22% of PmpD is half exposed and a 43% is buried. The model seems to be good quality as suggested by confidence scores assigned by the software: P-value = 2.59×10^{-17} , and uGDT = 409.

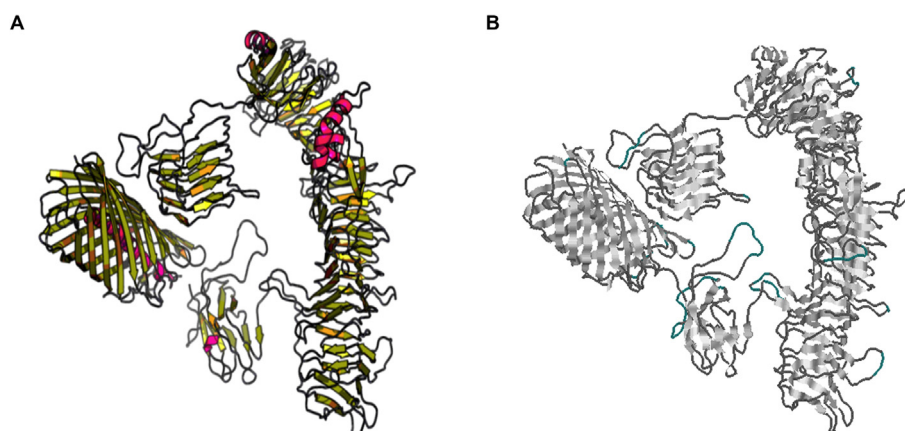


Fig. 1 3D structure of PmpD. (A) Molecular model obtained with Raptor X. (B) Conformational B-cell epitopes predicted by DiscoTope (cyan regions).

Additional checking was performed with ANOLEA and RAMPAGE. ANOLEA analysis indicated that 45% of amino acids modeled by Raptor X have high energy. Ramachandran plot showed that 91.3% residues in the PmpD model were within the most favored regions, 5.8% residues in the allowed region and 2.9% residues in the outlier region. Homology modeling was also performed with Phyre [38] and Swiss-Model [39] softwares. However, unlike Raptor X, these only predicted the β barrel forming a single domain (data not shown).

B-cell epitope prediction

Linear B-cell epitope prediction

BepiPred predicted 48 B-cell linear epitopes (that have more than four residues) with a sensitivity and a specificity of 75% and 50%, respectively. As can be seen in Fig. 2, these epitopes are abundant and are homogeneously distributed throughout the protein sequence. Forty motifs characteristic of Pmps (21 GGA and 19 FxxN) are present in the amino acid sequence, ten of them are included in epitope regions (sequences are highlighted in bold (FxxN) and underlined (GGA) in Fig. 2). Larsen et al. found that BepiPred had the highest prediction accuracy on the test data set, and it was shown to perform significantly better than all other methods tested on the validation data set [28,40].

Discontinuous B-cell epitopes

DiscoTope identified 57 residues, as part of epitope regions. A threshold of -3.7 for epitope identification was selected; therefore B-cell discontinuous epitopes were predicted with 0.75 specificity and 0.47 sensitivity. As the theory predicts, conformational epitopes were observed in loops, mainly in the regions not included in the β -barrel domain (C-terminal) (Fig. 1B).

Although confidence scores assigned by Raptor X indicated a good quality model, the reliability of the 3D model was

further analyzed at the localization of the predicted conformational epitopes. The analysis performed by ANOLEA indicated that 85% of the 57 residues have high energy. The Ramachandran plot showed that 86% of the epitope residues are included in favorable regions, 12.3% in permitted regions and 1.7% in outlier regions.

T-cell epitope prediction

The ability of an epitope for binding to MHC molecules is expressed as IC50 value (nM). Strong binding peptides (SB) have an IC50 value below 50 nM, and weak binding peptides (WB) an IC50 value between 50 and 500 nM. The location of T-cell epitopes with strong and weak affinity for human MHC-I and MHC-II alleles was predicted using IEDB resource. As can be seen in Table 1, NetMHC predicted 271 T-cell epitopes of different lengths (9aa, 10aa, 11aa, 12aa) with weak affinity to molecules of MHC I and 70 T-cell epitopes of different lengths (9aa, 10aa, 11aa, 12aa) with strong affinity to molecules of MHC I (Table 1A). On the other hand, NetMHCII predicted 2903 T-cell epitopes of 15aa with weak affinity and 742 T-cell epitopes of 15aa with strong affinity to molecules of MHC class II for the most prevalent MHC alleles in Argentina according to Perusco et al. [40] (Table 1B). These epitopes are abundant and are homogeneously distributed throughout the protein sequence. Multiple alleles of MHC Class II are able to bind the same epitope. This was observed with several epitopes.

Vaccine candidates

Giving that potential B- and T-cell epitopes were found to be distributed throughout the whole sequence, numerous fragments could be selected as candidates for the design of a vaccine. Nevertheless, it is necessary to define both B- and T-cell epitopes as candidates, to thereby generate a vaccine inducing both humoral and cellular response. Therefore, a

```

MSSEKDIKSTCSKFSLSVVAAILASVSGLASCVDLHAGGQSVNELVYVGPQAVLLLDQIRDLFVGSKD
SQAEGQYRLIVGDPSSSQEKDADTLPGKVEQSTLFSVTNPVVFQGVDDQDQVSSQGLICSFTSSNLDL
PRDGEFSLGIAFVGDSSKAGITLTDVKASLSGAALYSTEDLIFEKIKGGLEFASCSSLEQGGACAAQSIL
IHDCQGLQVKHCTTAVNAEGSSANDHLGFGGGAFFVTGSLSGEKSLYMPAGDMVANCDDGAIFEG
NSANFANGGAIAASGKVLFVANDKKTSFIENRALSGGAIAASSDIAFQNCAELVFKGNCAIGTEDKG
SLGGGAISSLGTVLLQGNHGITCDKNESASQGGAIFGKNCQISDNEGPVVFRDSTACLGGAIAAQEI
VSIQNNQAGISFEGGKASFGGGIACGSFSSAGGASVLTIDISKNLGAISFSRTLCTTSDLGQMEYQGG
GALFGENISLSENAGVLTFKDNIVKTFASNGKILGGGAILATGKVEITNNSGGISFTGNARAPQALPTQ
EEFPLFSKKEGRPLSSGYSGGGAILGREVAILHNAAVFEQNRLQCSEEEATLLGCCGGGAVHGMDST
SIVGNSSVRFGNNYAMGQGVSSGALLSKTVQLAGNGSVDFSRNIASLGGGALQASEGNCELVDNGY
GGAIFTGSLREEDKLDGQIPEVLISGNAGDVVFSGNSSKRDEHLPHTGGGAICTQNLISQHTGNVLF
YNNVACSGGAVRIEDHGNVLEAFGGDIVFKGNSSFRAQGSDAIYFAGKESHITALNTEGHAIVFHD
ALYFENLKERKSAEVLINSRENPGYTGSIRFLEAESKVPQCIHVQGSLELLNGATLCSYGFKQDAG
AKLVLAAGSKLKILDSGTPVQGHAISKPEAEISSSEPEGAHSLWIAKNAQTTVPMDIHTISVDLASE
SSSQEGTVEAPQVIVPGGSYVRSGELNLELVNTTGTGYENHALLKNEAKVPLMSFVASDEASEAIS
NLSVSDLQIHVATPEIEEDTYGHMGDWSEAKIQDGTLVINWNPTGYRLDPKAGALVFNALWEGAV
LSALKNARFAHNLTAQRMEFDYSTNVWGFAFGFRTLSAENLVAIDGYKGAYGGASAGVDIQLMED
FVLGVSGAAFLGKMDSQKFDAEVSRKGVVGSVYTGLAGSWFFGQYSLGETQNDMKTRYGVLGE
SSASWTSRGVLADALVEYRSLVGPVRPTFYALHFNPYVEVSYASMKFPGFTEQGREARSFEDASLTNI
TPLGMKFELAFIKGQFSEVNSLGISYAWEAYRKVEGGAVLLEAGFDWEGAPMDLPRQELRVALEN
NTEWSSYFSTVLGLTAFCGGTSTDSKLGYEANAGLRLIF

```

Fig. 2 PmpD amino acid sequence. BepiPred predicted 48 B-cell linear epitopes (highlighted in gray). FxxN motifs are indicated in bold, GGA sequences are underlined.

Table 1 (A) Number of T-cell epitopes of different lengths (9aa, 10aa, 11aa, 12aa) predicted with Strong Binding (SB) and Weak Binding (WB) to MHC class I. (B) Number of T-cell epitopes of 15aa predicted with Strong Binding (SB) and Weak Binding (WB) to MHC class II.

A								
MHC class I Prediction	9aa		10aa		11aa		12aa	
	SB	WB	SB	WB	SB	WB	SB	WB
HLA-A*02:01	14	25	4	20	1	15	7	21
HLA-B*35:01	9	33	1	17	10	26	10	22
HLA-B*35:03	1	0	0	1	0	0	0	1
HLA-C*07:01	2	15	3	11	3	12	4	6
HLA-C*07:02	0	10	0	10	1	12	0	14

B		
MHC class II Prediction	15aa	
	SB	WB
HLA-DRβ1*01:01	375	619
HLA-DRβ1*04:01	39	479
HLA-DRβ1*04:04	40	442
HLA-DRβ1*04:05	47	365
HLA-DRβ1*07:01	99	402
HLA-DRβ1*11:01	20	264
HLA-DRβ1*13:02	122	332

more exhaustive analysis was performed to identify regions comprising both types of epitopes. The analysis consisted of simultaneously comparing the results obtained by the predictors. As shown in Table 2, residues of several epitopes have been identified simultaneously by four predictors: BepiPred (linear B-cell), DiscoTope (conformational B-cell), NetMHC (T-cell for MHC I) and NetMHCII (T-cell for MHC II). The data

indicate that there are six regions with epitopes predicted by at least two predictors (Table 2, Fig. 3). An additional analysis of sequence conservancy showed that these regions are conserved among serovars, with the exception of aa 782 (V in serovars A-C, and A in serovars D-L). Regarding other proteins from the PmpD family, there is a variable degree of similarity among the selected fragments and PmpA, -B, -C, -E, -F, and -H (0–55%), and with Pmp21 from *Chlamydia pneumoniae* (15–55%). No significant similarity was observed with human proteins.

Discussion

C. trachomatis is the leading cause of bacterial sexually transmitted disease worldwide. Despite the fact that antibiotic treatment is effective in eliminating the pathogen, up to a 70% of all infections are asymptomatic. Untreated *Chlamydia* infections can lead to severe reproductive problems. Several *C. trachomatis* vaccine designs have been tested, but unfortunately there is no vaccine for *Chlamydia* despite the many efforts that have been made throughout the years. In part, this is due to an incomplete understanding of the immune response to *Chlamydia* urogenital infection [41] and of the critical immune factors required for resolving infections and protecting against reinfection [42,43]. Many researchers suggest that both humoral and cell mediated immunity for a successful *Chlamydia* vaccine are required. Olsen et al. [44] describe that the primary role of neutralizing antibodies will be to reduce initial infectious load once intra-cellular, remaining bacteria can be targeted by a bactericidal cell mediated immune (CMI) response. For this reason, a rationally designed *Chlamydia* vaccine requires identification of optimal cell T- and B-cell

Table 2 B- and T-cell epitopes that include residues predicted simultaneously by at least two predictors among BepiPred (B-cell linear), DiscoTope (B-cell conformational), NetMHC I (T-cell for MHC I) and NetMHCII (T-cell for MHC II) are highlighted in bold. Core amino acids of MHC II peptides are underlined.

Regions	Linear B-cell epitopes	Conformational B-cell epitopes	T-cell epitopes for MHC II (SB)	T-cell epitopes for MHC I (WB)
1	PSSFQEKDADTLPGKVE (82–98)	FQEKD (85–89)	HLA-DRβ1*13:02 GKVEQSTLFSVTNPV (95–109)	HLA-B*35:01 LIVGDPSSF (77–85), HLA-B*35:01 LPGKVEQSTLF (93–103) HLA-A*02:01 KVEQSTLFSV (96–105) HLA-C*07:02 SFQEKDADTL (84–93)
2	NNQAGISFEGGKASFG (412–427)	GIS (416–418) E (420)	HLA-DRβ1*01:01 AGISFEGGKASFGG (415–429)	HLA-B*35:01 VSIQNNQAGISF (408–419)
3	FASEDGDLSPESISSEELA (763–782)	SEELAL (778–783)	HLA-DRβ1*01:01 FITAAANQALFAS EDG (754–768)	HLA-B*35:01 FASEDGDLSPES (763–774)
4	LREEDKLDGQ (825–834)		HLA-DRβ1*13:02 GQIPEVLISGNAGDV (833–848)	HLA-A*02:01 KLDGQIPEVLI (830–840)
5	PEIEEDTYGHMGDWSEAKIQ (1171–1190)	IHVATPEI (1166–1173)	HLA-DRβ1*01:01 WSEAKIQDGLVINW (1184–1198)	HLA-B*35:01 TPEIEEDTY (1170–1178)
6	LGESSASWT (1353–1361)		HLA-DRβ1*01:01 RYGVLGESSASWTSR (1349–1363)	

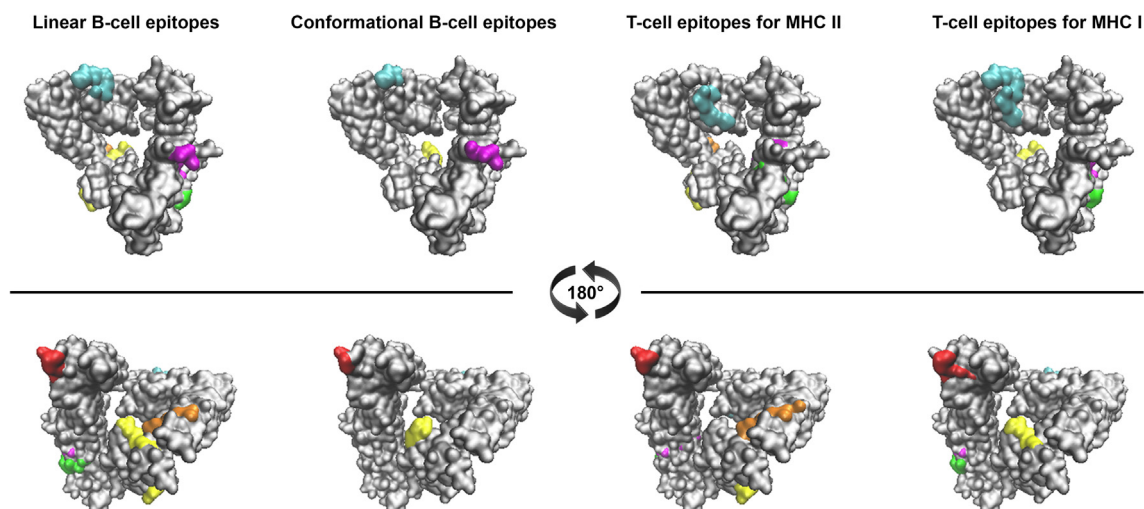


Fig. 3 Regions containing epitopes predicted simultaneously by more than one predictor. B- and T-cell epitopes. Regions are highlighted in different colors: 1: cyan, 2: red, 3: magenta, 4: green, 5: yellow, and 6: orange.

antigens. Considering that the immunoinformatics is a discipline which can predict *a priori* whether an antigen is immunogenic through epitope prediction, our aim was to perform *in silico* prediction of B- and T-cell epitopes within the amino acid sequence of PmpD for the design of a subunit vaccine against *C. trachomatis*.

In the present study, different tools were used to analyze antigenic epitopes for PmpD of *C. trachomatis*. The complete amino acid sequence of PmpD was acquired from the data base of NCBI (<http://www.ncbi.nlm.nih.gov/>) corresponding to serovar L2. PmpD is one of the most conserved protein among *C. trachomatis* serovars (with 99.1% similarity at the amino acid sequence level) [24]. We confirmed that serovars A to L share 99–100% identity of amino acid sequence by Clustal Omega (Multiple Alignment).

Considering that PmpD plays a critical role in chlamydial biology and pathogenesis, particularly at early host–cell interaction [45], it is very important to determine the presence and localization of B-cell epitopes in the molecule. In this study, B-cell epitopes with a sensitivity and a specificity of 75% and 50%, respectively were predicted employing two different algorithms, BepiPred (linear epitopes) and DiscoTope (conformational epitopes). BepiPred predicted 48 B-cell linear epitopes and DiscoTope identified 57 residues which could account for 12 B-cell conformational epitopes. Regarding linear epitopes, ten of the GGA and FxxN motifs are included in epitope regions. These motifs can modulate the protein conformation implicated in adhesion or can modulate adhesion interaction with the eukaryotic receptors [2].

Considering that the conformational epitopes are predicted using a 3D model, the reliability of the model must be taken into account, particularly in the predicted regions as conformational B-cell epitopes. Quality parameters given by Raptor X suggest an overall good quality model and RAMPAGE indicated that a high proportion (86%) of the residues predicted to be part of B-cell epitopes are in favored regions. On the other hand, ANOLEA analysis indicated that 45% of amino acids modeled by RaptorX and 85% of the residues in

conformational epitopes have high energy. This is not surprising giving that most of the conformational epitopes were predicted in loops, which are a big challenge in protein modeling because they often represent insertions or deletions in homologue proteins, they are less confined by alignment derived restraints [26], and they are usually structures with positive energy as predicted by ANOLEA or other servers [46,47].

Simultaneously, BepiPred and DiscoTope predicted 9 Epitopes. The analysis performed using a bioinformatic approach contributed to identify numerous B-cell epitopes, and this correlates well with several previous studies where PmpD was able to elicit antibody production. Human serum studies have found anti-Pmps antibodies in *C. trachomatis*-infected patients [2]. Moreover, Tan et al. indicate that some Pmps, including PmpD, may be more abundantly expressed or specifically exposed at the chlamydial surface to elicit a relatively stronger antibody response in adolescents with and without pelvic inflammatory disease (PID) [11]. Several researchers have also shown the neutralizing ability of anti-PmpD antibodies [15,16,18]. *In vitro* neutralization assays showed that serovars Ba, D, E, and L2 were more efficiently neutralized than serovars A, C, F, G and K by antibodies against recombinant PmpD [16]. Although the high percentage of similarity among sequences, these differences could be related to little changes in specific amino acid. Then, defining epitope containing regions which are simultaneously conserved may allow choosing an immunogen able to induce neutralizing antibodies against all serovars. In this work, we identified six regions containing B- and T-cell epitopes, which additionally are conserved among serovars A-L. Besides being useful for vaccine design, predicting B-cell epitopes can contribute to the design of various diagnostic methods.

An adequate humoral response depends upon appropriate CD4+ T-cell response. Moreover, vaccines that induce cellular immune responses are essential for infections caused by intracellular pathogens. One of the major impediments in developing such vaccines includes difficulty in identifying

relevant T-cell antigens that elicit protective cellular immunity [48]. CD4+ T-cell-mediated immunity is a major component of host defense against *C. trachomatis* infection [49] and to a lesser degree CD8+ T cells that recognize specific chlamydial antigens on MHC molecules [2]. The identification of epitopes presented by MHC class I and II molecules should enable the development of an effective T-cell response against *C. trachomatis*. In this study, NetMHC and NetMHCII were employed for T-cell epitope prediction. Regarding the prediction of T-cell epitopes in the PmpD sequence performed, NetMHC predicted 271 T-cell epitopes of different lengths (9aa, 10aa, 11aa, 12aa) with weak affinity to MHC I molecules and 70 T-cell epitopes of different lengths (9aa, 10aa, 11aa, 12aa) with strong affinity to MHC I molecules (Table 1A). Moreover, NetMHCII predicted 742 (Table 1B) T-cell epitopes of 15aa with strong affinity and 2903 T-cell epitopes of 15aa with weak affinity to MHC class II molecules for the most prevalent MHC alleles in Argentina according to Perusco et al. [40], as well as in North America and Europe (<http://www.allelefrequencies.net/default.asp>). These epitopes are abundant and homogeneously distributed throughout the protein sequence. Multiple alleles of MHC Class II are able to bind the same epitope. This was observed with several epitopes.

CD4+ and CD8+ T-cell immune responses in humans and mice are induced by *Chlamydia* infection, but the role of CD8+ T-cells in protective immunity is not clear. In the murine model, CD8+ T-cells contribute significantly to upper genital tract pathology [50,51] and infertility [52] but CD8+ T-cell are not necessary for clearing genital tract infection or protecting against reinfection [53,54]. Therefore, the results obtained in this paper are promising due to the fact that we found more epitopes able to link with strong affinity to MHC class II with respect to epitopes predicted to bind to MHC class I, which are less and with weak affinity. The results allow choosing a PmpD sequence containing strong affinity to molecules of MHC class II, which would favor a desired response of CD4+ type against *C. trachomatis*.

Comparative analyses performed by simultaneous predictors identifies six regions that have lineal and conformational B-cell epitopes and affinity to MHC I (WB) and MHC II (SB) molecules. This allows selecting sequences that may activate T-cells and induce cytokine and antibody production which contribute in the resolution of a chlamydial infection. Nevertheless, experimental assessment of the immunogenic ability of such regions both *in vitro* and *in vivo* should be done to verify this prediction or to select the best fragment.

Conclusion

PmpD has potential B-cell and T-cell epitopes distributed throughout the whole sequence. Numerous fragments could be selected as candidates for the design of a vaccine against *C. trachomatis*. To this end, in this study six regions containing both B-cell and T-cell epitopes were identified that may be used to generate both humoral and cellular response. These antigenic peptides can be valuable candidates for diagnostic as well as for therapeutic and preventive vaccines against *C. trachomatis*. Moreover, for vaccine design it would be important to select those regions that have the structural features of

PmpD such as N-terminal GGA(I/L/V) and FxxN tetrapeptide repeats, an integrin-binding RGD motif, and a putative nuclear localization signal besides both B-cell and T-cell epitopes.

Selected fragments of PmpD that contains B- and T-cell epitopes in conjunction with effective adjuvants and/or delivery systems may contribute to the development of an effective chlamydial vaccine that elicits a Th1-mediated and humoral protective immune response that does not induce adverse immunopathologies. The data obtained will be utilized to delineate a series of immunization strategies. Among them: the synthesis of single epitopes that could be administered as unique or in combination, the design of a multi-epitope chimeric molecule or the selection of fragments containing several immunogenic regions.

The results of our study provide computational data for the identification and screening of epitopes in PmpD, and may be used for the development of epitope vaccines that have an enhanced safety and efficacy.

Conflicts of interest

No conflicts of interests.

Acknowledgements

This work was supported by Universidad Nacional del Litoral (CAI+D Orientado 2014 to C.V.)

REFERENCES

- [1] WHO. Global incidence and prevalence of selected curable sexually transmitted infections-2008. Geneva, Switzerland: World Health Organization; 2012.
- [2] Vasilevsky S, Stojanov M, Greub G, Baud D. Chlamydial polymorphic membrane proteins: regulation, function and potential vaccine candidates. *Virulence* 2016;7:11–22.
- [3] Vogel G. Infectious diseases. Tackling neglected diseases could offer more bang for the buck. *Science* 2006;311:592–3.
- [4] Peipert JF. Clinical practice. Genital chlamydial infections. *N Engl J Med* 2003;349:2424–30.
- [5] Westrom L, Joesoef R, Reynolds G, Hagdu A, Thompson SE. Pelvic inflammatory disease and fertility. A cohort study of 1,844 women with laparoscopically verified disease and 657 control women with normal laparoscopic results. *Sex Transm Dis* 1992;19:185–92.
- [6] Redgrove KA, McLaughlin EA. The role of the immune response in *Chlamydia trachomatis* infection of the male genital tract: a double-edged sword. *Front Immunol* 2014;5:534.
- [7] Schachter J. Chlamydial infections (third of three parts). *N Engl J Med* 1978;298:540–9.
- [8] Becker E, Hegemann JH. All subtypes of the Pmp adhesin family are implicated in chlamydial virulence and show species-specific function. *Microbiologyopen* 2014;3:544–56.
- [9] Karunakaran KP, Yu H, Jiang X, Chan Q, Moon KM, Foster LJ, et al. Outer membrane proteins preferentially load MHC class II peptides: implications for a *Chlamydia trachomatis* T cell vaccine. *Vaccine* 2015;33:2159–66.

- [10] Johnson RM, Yu H, Kerr MS, Slaven JE, Karunakaran KP, Brunham RC. PmpG303-311, a protective vaccine epitope that elicits persistent cellular immune responses in *Chlamydia muridarum*-immune mice. *Infect Immun* 2012;80:2204–11.
- [11] Tan C, Hsia RC, Shou H, Haggerty CL, Ness RB, Gaydos CA, et al. *Chlamydia trachomatis*-infected patients display variable antibody profiles against the nine-member polymorphic membrane protein family. *Infect Immun* 2009;77:3218–26.
- [12] Molleken K, Schmidt E, Hegemann JH. Members of the Pmp protein family of *Chlamydia pneumoniae* mediate adhesion to human cells via short repetitive peptide motifs. *Mol Microbiol* 2010;78:1004–17.
- [13] Montigiani S, Falugi F, Scarselli M, Finco O, Petracca R, Galli G, et al. Genomic approach for analysis of surface proteins in *Chlamydia pneumoniae*. *Infect Immun* 2002;70:368–79.
- [14] Vandahl BB, Pedersen AS, Gevaert K, Holm A, Vandekerckhove J, Christiansen G, et al. The expression, processing and localization of polymorphic membrane proteins in *Chlamydia pneumoniae* strain CWL029. *BMC Microbiol* 2002;2:36.
- [15] Wehrl W, Brinkmann V, Jungblut PR, Meyer TF, Szczepek AJ. From the inside out—processing of the *Chlamydial* autotransporter PmpD and its role in bacterial adhesion and activation of human host cells. *Mol Microbiol* 2004;51:319–34.
- [16] Crane DD, Carlson JH, Fischer ER, Bavoi P, Hsia RC, Tan C, et al. *Chlamydia trachomatis* polymorphic membrane protein D is a species-common pan-neutralizing antigen. *Proc Natl Acad Sci U S A* 2006;103:1894–9.
- [17] Kiselev AO, Skinner MC, Lampe MF. Analysis of pmpD expression and PmpD post-translational processing during the life cycle of *Chlamydia trachomatis* serovars A, D, and L2. *PLoS One* 2009;4:e5191.
- [18] Swanson KA, Taylor LD, Frank SD, Sturdevant GL, Fischer ER, Carlson JH, et al. *Chlamydia trachomatis* polymorphic membrane protein D is an oligomeric autotransporter with a higher-order structure. *Infect Immun* 2009;77:508–16.
- [19] Tan C, Hsia RC, Shou H, Carrasco JA, Rank RG, Bavoi PM. Variable expression of surface-exposed polymorphic membrane proteins in in vitro-grown *Chlamydia trachomatis*. *Cell Microbiol* 2010;12:174–87.
- [20] Henderson IR, Navarro-Garcia F, Desvaux M, Fernandez RC, Ala'Aldeen D. Type V protein secretion pathway: the autotransporter story. *Microbiol Mol Biol Rev* 2004;68:692–744.
- [21] Grimwood J, Stephens RS. Computational analysis of the polymorphic membrane protein superfamily of *Chlamydia trachomatis* and *Chlamydia pneumoniae*. *Microb Comp Genomics* 1999;4:187–201.
- [22] Henderson IR, Lam AC. Polymorphic proteins of *Chlamydia* spp.—autotransporters beyond the Proteobacteria. *Trends Microbiol* 2001;9:573–8.
- [23] Saka HA, Thompson JW, Chen YS, Kumar Y, Dubois LG, Moseley MA, et al. Quantitative proteomics reveals metabolic and pathogenic properties of *Chlamydia trachomatis* developmental forms. *Mol Microbiol* 2011;82:1185–203.
- [24] Carrasco JA, Tan C, Rank RG, Hsia RC, Bavoi PM. Altered developmental expression of polymorphic membrane proteins in penicillin-stressed *Chlamydia trachomatis*. *Cell Microbiol* 2011;13:1014–25.
- [25] Kiselev AO, Stamm WE, Yates JR, Lampe MF. Expression, processing, and localization of PmpD of *Chlamydia trachomatis* Serovar L2 during the chlamydial developmental cycle. *PLoS One* 2007;2:e568.
- [26] Melo F, Devos D, Depiereux E, Feytmans E. ANOLEA: a www server to assess protein structures. *Proc Int Conf Intell Syst Mol Biol* 1997;5:187–90.
- [27] Lovell SC, Davis IW, Arendall 3rd WB, de Bakker PI, Word JM, Prisant MG, et al. Structure validation by Calpha geometry: phi, psi and Cbeta deviation. *Proteins* 2003;50:437–50.
- [28] Larsen JE, Lund O, Nielsen M. Improved method for predicting linear B-cell epitopes. *Immunome Res* 2006;2:2.
- [29] Parker JM, Guo D, Hodges RS. New hydrophilicity scale derived from high-performance liquid chromatography peptide retention data: correlation of predicted surface residues with antigenicity and X-ray-derived accessible sites. *Biochemistry* 1986;25:5425–32.
- [30] Ponomarenko JV, Van Regenmortel MH. B-cell epitope prediction. In: Gu J, Bourne PE, editors. *Structural bioinformatics*. John Wiley and Sons, Inc.; 2009. p. 849–79.
- [31] Haste Andersen P, Nielsen M, Lund O. Prediction of residues in discontinuous B-cell epitopes using protein 3D structures. *Protein Sci* 2006;15:2558–67.
- [32] Nielsen M, Lundegaard C, Worning P, Lauemoller SL, Lamberth K, Buus S, et al. Reliable prediction of T-cell epitopes using neural networks with novel sequence representations. *Protein Sci* 2003;12:1007–17.
- [33] Lundegaard C, Lund O, Nielsen M. Accurate approximation method for prediction of class I MHC affinities for peptides of length 8, 10 and 11 using prediction tools trained on 9mers. *Bioinformatics* 2008;24:1397–8.
- [34] Lundegaard C, Lamberth K, Harndahl M, Buus S, Lund O, Nielsen M. NetMHC-3.0: accurate web accessible predictions of human, mouse and monkey MHC class I affinities for peptides of length 8-11. *Nucleic Acids Res* 2008;1:509–12.
- [35] Nielsen M, Lund O. NN-align. An artificial neural network-based alignment algorithm for MHC class II peptide binding prediction. *BMC Bioinf* 2009;10:296.
- [36] Nielsen M, Lundegaard C, Lund O. Prediction of MHC class II binding affinity using SMM-align, a novel stabilization matrix alignment method. *BMC Bioinf* 2007;8:238.
- [37] Kallberg M, Wang H, Wang S, Peng J, Wang Z, Lu H, et al. Template-based protein structure modeling using the RaptorX web server. *Nat Protoc* 2012;7:1511–22.
- [38] Kelley LA, Sternberg MJ. Protein structure prediction on the Web: a case study using the Phyre server. *Nat Protoc* 2009;4:363–71.
- [39] Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, et al. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res* 2014;42(Web Server issue):W252–8.
- [40] Perusco A, Gamba C, Galarza P, Rojas F, Kalapodis A, Onofri A, et al. HLA antigenic and haplotype frequencies estimated in hematopoietic progenitor cell donors from Argentina. *Transplant Proc* 2014;46:3064–7.
- [41] Li LX, McSorley SJ. A re-evaluation of the role of B cells in protective immunity to *Chlamydia* infection. *Immunol Lett* 2015;164:88–93.
- [42] Morrison SG, Morrison RP. A predominant role for antibody in acquired immunity to chlamydial genital tract reinfection. *J Immunol* 2005;175:7536–42.
- [43] Johansson M, Schon K, Ward M, Lycke N. Genital tract infection with *Chlamydia trachomatis* fails to induce protective immunity in gamma interferon receptor-deficient mice despite a strong local immunoglobulin A response. *Infect Immun* 1997;65:1032–44.
- [44] Olsen AW, Follmann F, Erneholt K, Rosenkrands I, Andersen P. Protection against *Chlamydia trachomatis* infection and upper genital tract pathological changes by vaccine-promoted neutralizing antibodies directed to the VD4 of the major outer membrane protein. *J Infect Dis* 2015;212:978–89.

- [45] Kari L, Southern TR, Downey CJ, Watkins HS, Randall LB, Taylor LD, et al. Chlamydia trachomatis polymorphic membrane protein D is a virulence factor involved in early host-cell interactions. *Infect Immun* 2014;82:2756–62.
- [46] Figueroa M, Oliveira N, Lejeune A, Kaufmann KW, Dorr BM, Matagne A, et al. Octarellin VI: using rosetta to design a putative artificial (beta/alpha)8 protein. *PLoS One* 2013;8:e71858.
- [47] Chodanowski P, Grosdidier A, Feytmans E, Michielin O. Local alignment refinement using structural assessment. *PLoS One* 2008;3:e2645.
- [48] Karunakaran KP, Yu H, Foster LJ, Brunham RC. Development of a Chlamydia trachomatis T cell vaccine. *Hum Vaccine* 2010;6:676–80.
- [49] Brunham RC, Rey-Ladino J. Immunology of Chlamydia infection: implications for a Chlamydia trachomatis vaccine. *Nat Rev Immunol* 2005;5:149–61.
- [50] Vlcek KR, Li W, Manam S, Zanotti B, Nicholson BJ, Ramsey KH, et al. The contribution of Chlamydia-specific CD8(+) T cells to upper genital tract pathology. *Immunol Cell Biol* 2015;94:208–12.
- [51] Murthy AK, Li W, Chaganty BK, Kamalakaran S, Guentzel MN, Seshu J, et al. Tumor necrosis factor alpha production from CD8+ T cells mediates oviduct pathological sequelae following primary genital Chlamydia muridarum infection. *Infect Immun* 2011;79:2928–35.
- [52] Igietseme JU, He Q, Joseph K, Eko FO, Lyn D, Ananaba G, et al. Role of T lymphocytes in the pathogenesis of Chlamydia disease. *J Infect Dis* 2009;200:926–34.
- [53] Morrison SG, Morrison RP. Resolution of secondary Chlamydia trachomatis genital tract infection in immune mice with depletion of both CD4+ and CD8+ T cells. *Infect Immun* 2001;69:2643–9.
- [54] Morrison RP, Feilzer K, Tumas DB. Gene knockout mice establish a primary protective role for major histocompatibility complex class II-restricted responses in Chlamydia trachomatis genital tract infection. *Infect Immun* 1995;63:4661–8.