

Pathogen Proteins Eliciting Antibodies Do Not Share Epitopes with Host Proteins: A Bioinformatics Approach

Isaac Amela, Juan Cedano, Enrique Querol*

Institut de Biotecnologia i de Biomedicina and Departament de Bioquímica i Biologia Molecular, Universitat Autònoma de Barcelona, Barcelona, Spain

The best way to prevent diseases caused by pathogens is by the use of vaccines. The advent of genomics enables genome-wide searches of new vaccine candidates, called reverse vaccinology. The most common strategy to apply reverse vaccinology is by designing subunit recombinant vaccines, which usually generate an humoral immune response due to B-cell epitopes in proteins. A major problem for this strategy is the identification of protective immunogenic proteins from the surface of the pathogen. Epitope mimicry may lead to auto-immune phenomena related to several human diseases. A sequence-based computational analysis has been carried out applying the BLASTP algorithm. Therefore, two huge databases have been created, one with the most complete and current linear B-cell epitopes, and the other one with the surface-protein sequences of the main human respiratory bacterial pathogens. We found that none of the 7353 linear B-cell epitopes analysed shares any sequence identity region with human proteins capable of generating antibodies, and that only 1% of the 2175 exposed proteins analysed contain a stretch of shared sequence with the human proteome. These findings suggest the existence of a mechanism to avoid autoimmunity. We also propose a strategy for corroborating or warning about the viability of a protein linear B-cell epitope as a putative vaccine candidate in a reverse vaccinology study; so, epitopes without any sequence identity with human proteins should be very good vaccine candidates, and the other way around.

Citation: Amela I, Cedano J, Querol E (2007) Pathogen Proteins Eliciting Antibodies Do Not Share Epitopes with Host Proteins: A Bioinformatics Approach. PLoS ONE 2(6): e512. doi:10.1371/journal.pone.0000512

INTRODUCTION

Vaccination is the preventive method of choice to fight against microbial pathogens and presents the best cost/benefit ratio among current clinical and pharmaceutical practices. There are many reasons and serious threats which make the development of new advanced vaccines necessary, for example, avian flu the spread of antibiotic-resistant strains of pathogens [1].

The advent of genomics and high-throughput cloning/expression of large sets of genomic ORFs from pathogens makes genome-wide searches of new vaccine candidates possible. This systematic identification of potential antigens and virulence factors of a pathogen without the need for its cultivation has been termed “reverse vaccinology” [2,3]. The objective is to find proteins eliciting antibodies capable of binding to the bacterial surface, and through interaction with the complement system kill certain pathogen microorganisms. However, current studies show that only a small fraction of the pathogen proteins, most surface-exposed or secreted, appears to elicit antibodies with bactericidal activity [2,3]. It is generally considered that a bactericidal assay, that is, an antigen that elicits murine antibodies capable of triggering bacterial-cell death *in vitro* in a complement-dependent manner, is a good candidate for human vaccine development [4–7]. However, a major obstacle to reverse vaccinology, besides sequence and antigenic variability, is the difficulty to identify, from the pathogen proteome, those proteins that will generate a protective response. Typically, only a very small fraction of the antibodies raised in large-scale antigen-screening studies are bactericidal [3,8,9].

Epitope mimicry appears when a stretch of shared sequence, called “mimetope”, exists between a protein of a certain pathogen and a protein of its host. In some cases this may lead to auto-immune phenomena most of them related to several diseases. Different pathogens presenting epitope mimicry could cause auto-immune diseases like, for instance, *Borrelia burgdorferi*, leading to Lyme disease or neuroborreliosis; several Streptococci related to Rheumatic Fever, *Trypanosoma cruzi*, causing Chaga’s Disease; *Campylobacter jejuni*, causing Guillain- Barré Syndrome; a group of

viruses and also *Chlamydia pneumoniae*, related to Multiple Sclerosis, B3 coxsakieviruses, leading to myocarditis; B4 coxsakieviruses or cytomegaloviruses, leading to Type I Diabetes; HSV-1, causing herpetic stromal keratitis or some kind of auto-immunity caused by mimicry of *Chlamydia pneumoniae* outer-membrane proteins and myosin [10]. For several of these kinds of auto-immune diseases, we do not yet know what the possible causal agent is like, for example, with Primary Biliary Cirrhosis, Psoriasis, Scleroderma, Sjögren’s Syndrome or Lupus.

In this context, a method to quickly identify potential protective antigens from the pathogen proteome would be very helpful. Our initial hypothesis is that, given that one of the most important tasks of the immune system is the differentiation between self and non-self proteins, this system will discard eliciting antibodies against pathogen proteins sharing epitopes with host proteins. On the other hand, epitope mimicry could cause several auto-immune diseases as stated above [10]. Moreover, nowadays it is thought that humoral response against foreign proteins can distinguish between dangerous proteins and nondangerous ones [11]. For this reason we have focused our study only on pathogen proteins,

.....

Academic Editor: Winston Hide, University of the Western Cape, South Africa

Received April 11, 2007; Accepted May 4, 2007; Published June 6, 2007

Copyright: © 2007 Amela et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This research was supported by Grants LSHB-CT-2006-037325 funded under the Sixth Research Framework Programme of the European Union; BFU2004-06377-CO2-01 from the MEC (Ministerio de Educación y Ciencia, Spain); and by the Centre de Referència de R+D de Biotecnologia de la Generalitat de Catalunya. I. Amela is a predoctoral fellowship recipient from the UAB (Universitat Autònoma de Barcelona).

Competing Interests: The authors have declared that no competing interests exist.

* To whom correspondence should be addressed. E-mail: enric.querol@uab.es

probably dangerous, and not on human self-microbiota proteins, most of them nondangerous.

A sequence-based computational analysis has been carried out to exemplify what has been said above. We studied the complete set of exposed surface proteins –the surfome- of most of the sequenced respiratory human-pathogen microorganisms with the aim of elucidating the presence of stretches of a shared sequence, which have characteristics of putative T-dependent B-cell epitopes, between human and pathogen proteins. This group of bacterial proteins, recently labelled the surfome [7], is susceptible to generating antibody immune response via B-cell epitopes, and it has been reported that some of them effectively deal with this type of response. The paradigm analysed here could be used to describe how the bacterial proteins sharing linear B-cell epitopes with human proteins avoid immunoreactivity against the host if these sequence stretches tend to produce antibodies. Many questions like the following ones were formulated: Are these produced antibodies non-protective ones? Or are these proteins not producing antibody responses even though they were able to do it? Auto-immune diseases should appear if the exposed proteins of the bacterial pathogens share linear B-cell epitopes with any human protein, so this study may be useful to identify them [2,3]. These findings may also help us to identify proteins which should not be used as putative vaccine candidates in a reverse vaccinology study. Moreover, the existence of a mechanism to avoid auto-immunity could be proposed. On the other hand, we therefore studied the current, already-known linear B-cell epitopes to elucidate if these sequence stretches either share or do not share common regions with human proteins. This study may help us to say that novel epitope mimics could generate new auto-immune diseases and could reinforce the hypothesis of the existence of the mechanism to avoid auto-immunity proposed above.

A linear or continuous B-cell epitope is a specific region of an antigen to which an antibody binds and its constitutive residues are sequential in the primary sequence of the protein. In contrast, conformational or discontinuous B-cell epitopes are highly conformational-dependant and its constitutive residues are non-sequential in the primary sequence. That is why conformational B-cell epitopes were not included in our field of study, thus we restricted our study to the group of linear B-cell epitopes. The lack of many databases for conformational B-cell epitopes and the poorly developed method for predicting it from structure, enforced this decision. We have to take into account that only the protein sequences sharing certain sequence identity regions were considered if these two proteins are different and do not have anything in common in their annotation. This decision was assumed due to the fact that there exist many proteins of completely different organisms grouped in the same family or being considered as protein-like ones and obviously these proteins are condemned to share common regions, so they were not included in our field of study. The antibodies, in fact, recognise a small part of a big molecule. In this sense, two quite different proteins could be considered that would be recognised by the same antibody if both shared a small epitope of five to six amino acids, more or less, which is the usually accepted minimal length for a linear B-cell epitope.

As a brief comment of our results, we found that none of the well-known protective antigens under our study present putative, common linear B-cell epitopes with the host proteins, which demonstrates the existence of a system that tries to avoid autoimmunity that select linear B-cell epitopes not sharing sequence stretches with human proteins. On the other hand, we found several pathogen proteins which share sequence stretches with the host ones, so we recommend not to using these proteins as

putative vaccine candidates in a reverse vaccinology study. These proteins are a small group of the total surfome studied and they also elicit antibodies with great difficulty, suggesting the same immunotolerance effect that tries to avoid autoimmunity mentioned here [12].

MATERIALS AND METHODS

Generation of the Datasets

With the aim of proposing a method applicable in a previous reverse vaccinology study, databases were made of the exposed proteins from the up-to-date, sequenced main human bacterial respiratory pathogens, which are: *Neisseria meningitidis* serogroup B, *Legionella pneumophila* (Lens strain), *Streptococcus pneumoniae*, *Haemophilus influenzae*, *Pseudomonas aeruginosa*, *Streptococcus pyogenes* serotype M1, *Yersinia pestis*, *Bordetella bronchiseptica*, *Staphylococcus aureus* (COL strain), *Pasteurella multocida*, *Bordetella parapertussis*, *Bordetella pertussis*, *Chlamydia pneumoniae* (*Chlamydophila pneumoniae*) and *Mycoplasma pneumoniae*. Besides, to attempt to demonstrate our previous proposal that some selectivity exists to try to avoid the auto-immune response, we obtained a complete set of linear B-cell epitopes from the three most complete and current epitope databases.

On one hand, a surface-protein sequence database for each pathogen was created downloading the protein sequences from the HAMAP tool under the ExPASy web server [13,14]. Furthermore, exhaustive searches in the NCBI [www.ncbi.nlm.nih.gov/entrez], SCIRUS for scientific information [www.scirus.com] and in specialised journals helped us to find proteins with certain interest for our study because of their reported capacity to generate antibody immune response via B-cell epitope activation. These protein sequences were added to each pathogen database. A total of 2175 proteins sequences from the surfome [7] of the studied pathogens was analysed.

On the other hand, we developed different computational programmes to obtain the human pathogen section of the Bcipep database [15], the most complete and specific B-cell epitope database now available from which 2275 pathogen linear B-cell epitopes were obtained. In addition, we performed a similar procedure to collect the B-cell epitope portion of the IEDB database [16], an extensive immune epitope database appeared in early-2006 with a total of 2154 linear B-cell epitopes. Finally, we could computationally retrieve all the linear B-cell epitopes from the AntiJen database [17], an important immunological database that contains quantitative binding data for epitope peptides from which we could obtain 2924 entries. Thus a total of 7353 linear B-cell epitopes has been studied and most of them were formed by at least six or more amino acids.

Procedure

These database files, 14 consisting of the different pathogen-exposed proteins and 3 of the most powerful recent databases, including linear B-cell epitopes, were generated to analyse if there exist sequence identities between human proteins and the collected surface microorganism proteins or between the same human proteins and the complete set of linear B-cell epitopes obtained. This study may lead us to request a method based on sequence similarity, like the one obtained here, that ought to be applied in a previous step of a reverse vaccinology study. Furthermore, this should help us to discover putative pathogens enhancing autoimmunity after their infection or demonstrating that some selectivity exists to avoid the auto-immune response.

For each of the pathogens analysed, downloading was started from the HAMAP tool under the ExPASy web server of all protein

sequences whose annotation included the following keywords: Outer, membrane, lipoprotein, adhesin, surface, secreted or exposed. Then, as said above, extensive searches were carried out through the NCBI web server, the SCIRUS web site and several specialised journals, with the aim of finding scientific articles talking about proteins that generate antibody immune response via B-cell epitope activation or could produce autoimmunity and probably act as putative vaccine candidates. An initial sequence database (ID) was made as explained before. Additionally, this initial database file for each pathogen was used to run a BLASTP algorithm analysis [18] against the protein non-redundant database at the NCBI ftp site [www.ncbi.nlm.nih.gov/Ftp]. If we found that some of these proteins had stretches of shared sequence with human proteins, thanks to this preliminary BLASTP analysis, we constructed a more extensive sequence database (ED) for each pathogen using, once again, the HAMAP tool under the ExPASy web server, now including all protein sequences that contain, in their annotation, the hypothetical, probable, conserved or putative protein annotation statements also considering the above-described keywords. This ED file of protein sequences for each pathogen served as input for a more accurate BLASTP analysis.

The BLASTP output file for each pathogen analysed had been carefully scrutinised to obtain a last file (LF) including the alignments that comprise local, significant sequence similarity between the pathogen query protein and a subject human one if one exists. A stretch of shared sequence was considered as a putative mimotope if there were at least 5 sequential, equal amino acids or if it is a region to take into account because of its amino acid physico-chemical properties similarity. For the LF file of exposed protein databases of the 14 pathogens analysed, we checked to see if these similar stretches correspond to transmembrane regions because a B-cell epitope cannot exist in a transmembrane region. This was done by applying TransMem, a programme for predicting transmembrane domains in proteins [19]. We also checked to see if these stretches of shared sequence coincide with the signal peptide section of the protein, when these are close to the N-terminal extreme, using the SignalP 3.0 Server [20]. Lastly, it was checked to determine if these stretches are predicted as putative linear B-cell epitopes using prediction servers like ANTIGENIC [21], ABCpred [22], and BcePred [23].

B-Cell Epitope Database

Thanks to the Bcipep, IEDB and AntiJen databases, we could group together a total of 7353 linear B-cell epitopes, preparing them for the subsequent BLASTP analysis against the human protein, non-redundant database at the NCBI ftp site.

Exposed Protein Databases

Regarding *Neisseria meningitidis* serogroup B, the ID contained 62 protein sequences [5,9,24–28] and the ED was made up of 263 protein sequences. In the case of *Legionella pneumophila* (Lens strain), the ID had 17 protein sequences [29–31] and the ED consisted of 352 protein sequences. The *Streptococcus pneumoniae* study allowed us to obtain an ID that comprised 27 protein sequences [7,32–36] and an ED of 338 protein sequences. Regarding *Haemophilus influenzae*, we were able to generate an ID of 25 protein sequences [37–41]. In this case, we did not search for more protein sequences because we did not find any previous sequence-identity region between the proteins from the ID and human ones, so the database remained consisting of a total of 25 protein sequences. Concerning *Pseudomonas aeruginosa*, we were able to assemble an ID of 34 protein sequences [42–49] and a huge ED that contained

569 protein sequences. The *Streptococcus pyogenes* serotype M1 study allowed us to obtain an ID of 30 protein sequences [7,50–58] and an ED of 278 protein sequences. Regarding *Yersinia pestis*, we only created an ID of 32 protein sequences [59–61]. We did not increase this ID because we did not find any region with a sequence identity between the proteins from this database and human proteins in the preliminary BLASTP analysis. Concerning *Bordetella bronchiseptica*, we could group together 46 protein sequences in its ID [62–64]. Nothing else was added to this ID because we did not find stretches of shared sequence between the proteins from this database and human ones. In the case of *Staphylococcus aureus* (strain COL), we were able to generate an ID of 43 protein sequences [65–67] and we could extend it to 490 protein sequences, producing a large ED. Regarding *Pasteurella multocida* assembled an ID of 20 protein sequences [68]. Nothing else was added to this ID because we did not find any local sequence-identity regions between the sequences from this database and human proteins. We made the same analysis for the rest of the pathogens in our study, which are *Bordetella parapertussis*, *Bordetella pertussis*, *Chlamydia pneumoniae* (*Chlamydophila pneumoniae*) and *Mycoplasma pneumoniae*. In the cases of *Bordetella parapertussis* and *Bordetella pertussis*, we were able to generate an ID of 28 and 37 sequences respectively [69–73]. For *Chlamydia pneumoniae* (*Chlamydophila pneumoniae*), the ID was made up of 68 protein sequences [74–77]. Regarding *Mycoplasma pneumoniae*, the ID contained 198 protein sequences. These last four pathogens have not got any sequence-identity region between the proteins from the IDs and human ones, so the databases were not enlarged. Taking into account all of the pathogens analysed, we were able to study the sequence of a total of 2175 proteins.

RESULTS AND DISCUSSION

B-Cell Epitope Analysis

In terms of our study, after analysing the three database files of linear B-cell epitopes obtained via BLASTP [18], an exhaustive review of the output files obtained led us to the proposal that some selectivity really does exist to try to avoid autoimmunity. We concluded this because we could see that none of those 7353 linear B-cell epitopes shared any sequence identity region with human proteins capable of generating antibodies despite the already known epitopes that generate auto-antibodies and cause autoimmune diseases, the allergies caused by some epitopes, or certain sequence identities found between some artificial peptides and human proteins after their administration. We found that well-known protective antigens do not present putative common linear B-cell epitopes with the host proteins, so this fact reinforces what we have postulated before that some selectivity may exist to try to avoid autoimmunity. Moreover, pathogen proteins which share sequence stretches with the host ones elicit antibodies with great difficulty, demonstrating the immunotolerance effect [12].

Exposed Protein Analysis

The LF for each pathogen analysed, coming from the output file after the BLASTP analysis [18], which contains the alignments comprising local, significant sequence similarity between the pathogen query protein and a subject human one, have been carefully scrutinised. Despite the huge amount of exposed proteins and pathogens analysed, around 2000, we could only consider approximately 20 protein alignments, representing only 1% of the total of proteins analysed. Consequently, we can say that, due to the probable existing selectivity mechanism to avoid autoimmunity already mentioned, the finding of pathogen-exposed proteins that have a stretch of shared sequence, considered as

a putative linear B-cell epitope, with human proteins is very difficult. As said before, epitope mimicry between proteins may lead to auto-immune phenomena most of them related to several diseases, but also pathogen proteins which share sequence stretches with the host ones elicit antibodies with great difficulty demonstrating the immunotolerance effect. The way of trying to avoid this mimicry and the presence, in some cases, of this immunotolerance effect, is to reinforce the difficulty in finding regions of sequence identity between pathogen and host proteins.

a***Streptococcus pneumoniae:******** BLASTP SUBJECT:**

```
>gb|AAG49577.1| uveal autoantigen [Homo sapiens]
Length = 1416

Score = 46.6 bits (109), Expect = 0.008
Identities = 78/373 (20%), Positives = 148/373 (39%), Gaps = 37/373 (9%)

Query: 461 KQPEVNSETNKLTAIDALNVDKTELNNNTIADAKTKVREHYSRSRW--QNLQTEVTKAERK 518
K ++N E K+K + L + N I + EH + S Q+++
Sbjct: 889 KPEDINQEFVKIKDKNEILKRNLENQNQIKAESLAEHEAKMSSLSQSMMRKVQDSNAE 948

Query: 519 VAANIDAKQSEVN-----EAVEKLTTATIEKLVLESELKEPILTLTSTDKKILEREAVAKYT 572
+ AN Q E+ +A +K TI++ +++ PI++ ++K E K
Sbjct: 949 ILANYRKQEEIVTLHAEIKAQKKELEDTIQECIKVYAPIVSFEECERKPKATEKELKDQ 1008

Query: 573 LENQNKTIIKSIATELKKGEVINTVVLITDDKVTTETISAAFKNLEYYKEYTLSTTMIYD 632
L Q + K E+KK ++ +DR+ E + K+L K + + +
Sbjct: 1009 LSEQTQ-KYSVSEEEVKKNRQ-----ENDLRLKEIFITLQ-KDLRD-KTVLIEKSHEME 1058

Query: 633 RGNGEGETTLENNQNIQLDLKKVELKNIKRDLIYENGKETNESLTTIPDDKSNNYLLKI 692
R +T+ L Q L K E+KN+K + + EN K+T+E L K + L+
Sbjct: 1059 RALSRKTDLENLKQLKDLSQKYTEVKVVKV--EKLVEENAKQTSEILAVQNLLQKQHVPLEQ 1116

Query: 693 TSNNQKTTLLAVKNIEETTVNGTPVYKVTAIADNLVSRTADNK-----F 736
+K+ ++N+E + Y+ + + +N+ F
Sbjct: 1117 VEALKKSLNGTIENLKEELKSMQRCYEKEQQTVTKLHQQLLENQKNSSVPLAELHQIKEAF 1176

Query: 737 EEEYVHYIEKPKVHEDNVYNNFKELVEAIQNQPSKEYLQGQSMSARNVVPNGKSYITKEP 796
E+E V I+ ++ N E V +O++ + + + R VV K TK
Sbjct: 1177 EKE-VGIKASLREKEEESQNKMEEVSKLQSEVQNTKQALKLETREVVDSLKYKATKSD 1235

Query: 797 TGKLLSSEGKQFA 809
+SS ++ A
Sbjct: 1236 LETQISSLNEKLA 1248
```

***** BLASTP QUERY:**

```
>UniProt/Swiss-Prot|Q9L7Q2|ZMPB_STRPN Zinc metalloprotease zmpB precursor
```

b***Chlamydia pneumoniae:******** BLASTP SUBJECT:**

```
>gb|AAG49577.1| uveal autoantigen [Homo sapiens]
Length = 1416

Score = 40.4 bits (93), Expect = 0.068
Identities = 44/182 (24%), Positives = 81/182 (44%), Gaps = 23/182 (12%)

Query: 1 MEEVSEYLQQVENQLESCSKRLTKMETFALGVRLREAKEEIESII-----LSDVVNRFEV 54
MEEVSE+ +V+N ++ K+L E L K ++E+ I L+++ ++E
Sbjct: 1198 MEEVSKLQSEVQNTKQAL-KKLETRVVDSLKYKATKSDLETQISSLNEKLANLNRYEE 1256

Query: 55 LCRDI----DEMLSRVEEIERMLRMAELPLLPIKEALTAKAFVQ---HNSCKEKLTKVEP 106
+C ++ +S +E E + E + KE K+ +E + +E
Sbjct: 1257 VCEEVLHAKKKEISAKDEKELLHFSIEQEIKDQKERCDSKLTITELQRRIQESAKQIEA 1316

Query: 107 YFKESPAYLTSEERL-QSLNQTLQRAY-----KESQKVSGLESEVRACREQLKDQVRQ 158
+ L ERL Q+LN Q Y ++SQ + L+ +V++ +QL D RQ
Sbjct: 1317 KDNKITELLNDVERLKQALNGLSQLTYTSGNPTKRQSQLIDTLQHQVKSLEQQQLADADRQ 1376

Query: 159 FE 160
+
Sbjct: 1377 HQ 1378
```

***** BLASTP QUERY:**

```
>AAD18195.1 cpnel CPn0042
```

Here we show an example of a procedure, applied in this case only to the main human, bacterial respiratory pathogens, which we think might be genuinely helpful for the development of new vaccines when corroborating or advising the viability of a pathogen surfome protein [7] as a putative vaccine candidate in a reverse vaccinology study [2,3].

As stated above, only 20 proteins of a total of 2000 analysed share a significant sequence-identity region with human proteins, so we strongly advise against recommending these proteins as putative vaccine candidates. On the contrary, the rest of the

Figure 1. Example of mimotope identification. (a) Partial sequence alignment coming from the output file of the BLASTP algorithm analysis between the exposed-protein database from *Streptococcus pneumoniae* and the non-redundant protein database. A putative linear B-cell epitope is highlighted. (b) Partial sequence alignment coming from the output file of the BLASTP algorithm analysis between the exposed-protein database from *Chlamydia pneumoniae* and the non-redundant protein database. A putative linear B-cell epitope is highlighted.

doi:10.1371/journal.pone.0000512.g001

proteins should be good vaccine candidates. Although this situation has to be considered, we show below some of the alignments between the exposed-pathogen proteins and several human proteins that we think would be more significant to highlight as an example of a previous massive analysis in a reverse vaccinology study.

Streptococcus pneumoniae and Chlamydia pneumoniae: After the BLASTP output file analysis of these two pathogens, we could see that there exist several bacterial proteins sharing a significant sequence-identity region with human proteins, but we found a very interesting particular case on which we focused in more detail (Figures 1a, 1b). Two proteins of these pathogens have a stretch of shared sequence with a human protein that is considered to have a uveal auto-antigen [78]. Furthermore, the region of each of the two proteins is not included in the predicted transmembrane sections. These regions also are not part of a signal peptide and there are also certain references highlighting that these proteins may elicit antibody immune

response [32–36,74–77]. These two proteins are the Zinc metalloprotease zmpB precursor in the case of *Streptococcus pneumoniae* and the Cpn0042 protein for *Chlamydia pneumoniae*. Additionally, the stretch of shared sequence for the *Streptococcus pneumoniae* protein does not correspond to the signal peptide because it is not close to the N-terminal extreme, and the stretch of shared sequence for *Chlamydia pneumoniae* protein is not a predicted signal peptide either. Even though the epitope prediction servers already mentioned could not corroborate that these sequence sections correspond to putative linear B-cell epitopes, we considered this case to be an important one.

In our supposed preliminary analysis of a reverse vaccinology study, we should recommend, although these two proteins previously seem to be good putative vaccine candidates, not using these proteins for the development of new vaccines as they probably generate antibodies against the human protein with which they share a stretch of sequence. Moreover, we have to consider that the human protein has an auto-antigen, so we

a

Pseudomonas aeruginosa:

```
*** BLASTP SUBJECT:
>ref|NP_078789.1| FYVE and coiled-coil domain containing 1 [Homo sapiens]
emb|CAC33883.1| FYVE and coiled-coil domain containing 1 [Homo sapiens]
Length = 1478

Score = 36.6 bits (83), Expect = 1.2
Identities = 35/122 (28%), Positives = 57/122 (46%), Gaps = 19/122 (15%)

Query: 152 EKTLQKRNIDGRNELIDAKMQL-----GKTSDEDRKIVGKVWAADQVQDSVALRAAGRAFE 207
EK LQ N+ GRN+L++ K+QAI + I G + + + Q S+ R G E
Sbjct: 622 EKELQ-NVVRGNQNLLEGKLQALQADYQALQQRESAIQGSLASLEAEQASL--RHLIGDQME 678

Query: 208 SRNGALQVANTIVIQSFVQMANASVQRQGESQ-----ASAREGEVNVATIGSQ 255
+ A++ A +++ + A +Q ++GE Q A AR E+ A Q Q
Sbjct: 679 ASLLAVRKAKEAMKAQMAKEAILQSKEGCQQQLREEVEQCQQQLAEARHRELRALESQCQ 738

Query: 256 KQ 257
+Q
Sbjct: 739 QQ 740
```

*** BLASTP QUERY:

```
>UniProt/TrEMBL|Q9I323|Q9I323 PSEAE Translocator outer membrane protein PopD
```

b

Chlamydia pneumoniae:

```
*** BLASTP SUBJECT:
>ref|NP_003557.2| early endosome (membrane-bound compartment inside cells) antigen 1,
162kD [Homo sapiens]
sp|Q15075|BEAL_HUMAN Early endosome antigen 1 (Endosome-associated protein p162) (Zinc
finger FYVE (= P.aeruginosa) domain-containing protein 2)
emb|CAA55632.1| endosomal protein [Homo sapiens]
Length = 1411

Score = 44.3 bits (103), Expect = 0.009
Identities = 51/217 (23%), Positives = 97/217 (44%), Gaps = 23/217 (10%)

Query: 128 LKAAKDQLTILEAIEFRNENGNLKTAAED---LEEQVSKLSEQLEALERINQLIQANAGDA 184
L+ ++ L +I+A E L E L+EQV++L+E+L+ ++ Q N D
Sbjct: 542 LEKEREDLYAKIQAEGEGETAVLNQLQEKNHTLQEQTOLTEKLNQSESHKQAEQNLHD- 600

Query: 185 QEISSEKKLISGWDSKVVEQINTSIQALKVQLLGQEWVQEAQTHVKAMQEQIQLAQAEIL 244
++ + L + D V + TS+ L L E4+ V + QI+A +L
Sbjct: 601 -QVQEQQKAHLRAAQDR--VLSLETSVNELNSQLN----ESKEKVSQLDIQIKAKTELLL 652

Query: 245 GMHNQSTALQKSVENLL-----VQD---QALTRVVGELLESENKLS---QACSLRQEIE 293
TA + ++N L +QD Q L ++ +L + KL + CS L ++
Sbjct: 653 SAEAAKTAQRAIDLQNHLDTAQNALQDKQQEELNKITTQLDQVIAKLDQKQEHCSQSLESHLK 712

Query: 294 KLAQHETSLQRIDAMLAQEQLNLAEQVTALEMKQEA 330
+ + SL+Q+ ++ Q + L ++ K+A
Sbjct: 713 EYKEKYLSLEQKTEELEGQIKKLEADSLEVAKSKEQA 749
```

*** BLASTP QUERY:

```
>Q9Z828|Q9Z828_CHLPN Similarity to CT119 Inclusion membrane protein A (IncA)
```

Figure 2. Example of mimotope identification. (a) Partial sequence alignment coming from the output file of the BLASTP algorithm analysis between the exposed-protein database from *Pseudomonas aeruginosa* and the non-redundant protein database. A putative linear B-cell epitope is highlighted. (b) Partial sequence alignment coming from the output file of the BLASTP algorithm analysis between the exposed-protein database from *Chlamydia pneumoniae* and the non-redundant protein database. A putative linear B-cell epitope is highlighted.

doi:10.1371/journal.pone.0000512.g002

Streptococcus pyogenes:

*** BLASTP SUBJECT:

```
>gb|AAH51330.1| DOCK6 protein [Homo sapiens]
Length = 509
```

Score = 34.7 bits (78), Expect = 4.4
 Identities = 19/58 (32%), Positives = 29/58 (50%)

```
Query: 181 AQSASEGPWLLAEGLPTVEDHRHLPPIGLQVELMKAIGTIDNLISNQFISEEEELAACT 238
          AQ      L+AE L +EDHRHLP+G      +  ++  IS+ +S +E   C+
Sbjct: 79  AQCMVHAAALVAEYLALLEDHRHLPVGCVSFQNISSNVLEESAISDDILSPDEEGFCS 136
```

*** BLASTP QUERY:

```
>outer surface protein
```

Figure 3. Example of mimotope identification. Partial sequence alignment coming from the output file of the BLASTP algorithm analysis between the exposed-protein database from *Streptococcus pyogenes* and the non-redundant protein database. A putative linear B-cell epitope is highlighted.
 doi:10.1371/journal.pone.0000512.g003

propose that perhaps a previous infection with *Streptococcus pneumoniae* or *Chlamydia pneumoniae* could promote an auto-immune reaction at the uveal tract apart from producing infection by themselves.

Pseudomonas aeruginosa and Chlamydia pneumoniae: The precise analysis of the BLASTP output file for these two pathogens allowed us to identify two proteins containing a significant sequence-identity region with a human protein completely differently annotated from the pathogen one (Figures 2a, 2b). Additionally, these sections do not correspond to transmembrane regions and there are also certain references highlighting the antibody elicitation and therefore the consequential immune response [45–49,74–77]. Furthermore, the epitope-prediction servers already mentioned could corroborate that the two identity regions correspond to putative linear B-cell epitopes. These proteins are the Translocator outer membrane protein PopD, a constituent of the *Pseudomonas aeruginosa* Type III Apparatus, and the Inclusion membrane protein A in the case of *Chlamydia pneumoniae*. Obviously, the stretches of shared sequence do not correspond to the signal peptide because they are not close to the N-terminal extreme of the protein.

We considered this example a truly important one for showing why a protein should not be used as a vaccine candidate after a preliminary analysis in a reverse vaccinology study. As is seen in the above alignments, the two pathogen proteins share a significant sequence-identity region with human proteins, so they are candidates for generating an antibody response against the human protein with which they share this stretch of sequence. Moreover, these human proteins are part of the FYVE domain containing proteins and are required for the formation of early-endosomal membranes that are the main resource for phagocyte action [79–83]. For this reason, the probable antibody elicitation against these proteins could generate a serious auto-immune problem for the endosome-mediated action in the human defence system.

Streptococcus pyogenes: The final BLASTP output file gave us a remarkable result indicating that there exists only one pathogen protein sharing a certain sequence identity-region with a human protein differently annotated from the pathogen one

(Figure 3). In addition to this protein having all the characteristics to be a putative vaccine candidate and to elicit antibody immune response [55,56], the epitope-prediction servers already mentioned could corroborate that the sequence-identity region corresponds to a putative linear B-cell epitope and the transmembrane-prediction server shows that this sequence identity region does not coincide with the predicted transmembrane ones. Obviously, the stretches of shared sequence do not correspond to the signal peptide because they are not close to the N-terminal extreme of the protein.

As can be seen here, this pathogen protein has a stretch of shared sequence with a human protein called DOCK6. Consequently, the putative antibodies generated against this exposed-pathogen protein could attack DOCK6, leading to auto-immune effects. DOCK6 promotes neurite outgrowth [84], so this protein is necessary for neural development and we believe that a lack of DOCK6 could be quite dangerous for human health. Therefore, we also recommend not considering this outer-surface protein as a putative vaccine candidate.

Final Considerations

We have proposed a protein sequence analysis that may be applied before a reverse vaccinology study for which we have shown several examples. Our proposal may be used in this kind of massive analysis for the development of new vaccines when corroborating or warning about the viability of a linear B-cell epitope as a putative vaccine candidate. Therefore, epitopes without any sequence identity with human proteins should be very good vaccine candidates, and the other way round.

ACKNOWLEDGMENTS

Author Contributions

Conceived and designed the experiments: EQ JC. Performed the experiments: IA. Analyzed the data: IA JC. Contributed reagents/materials/analysis tools: IA. Wrote the paper: EQ IA JC.

REFERENCES

1. Borchardt JK (2004) The history of bacterial meningitis treatment. *Drug News Perspect* 17: 219–224.
2. Rappuoli R (2000) Reverse vaccinology. *Curr Opin Microbiol* 3: 445–450.
3. Rappuoli R, Covacci A (2003) Reverse vaccinology and genomics. *Science* 302: 602.
4. Goldschneider I, Gotschlich EC, Artenstein MS (1969) Human immunity to the meningococcus. I. The role of humoral antibodies. *J Exp Med* 129: 1307–1326.
5. Pizza M, Scarlato V, Masiagnani V, Giuliani MM, Arico B, et al. (2000) Identification of vaccine candidates against serogroup B meningococcus by whole-genome sequencing. *Science* 287: 1816–1820.

6. Welsch JA, Rossi R, Comanducci M, Granoff DM (2004) Protective activity of monoclonal antibodies to genome-derived neisserial antigen 1870, a *Neisseria meningitidis* candidate vaccine. *J Immunol* 172: 5606–5615.
7. Rodriguez-Ortega MJ, Norais N, Bensi G, Liberatori S, Capo S, et al. (2006) Characterization and identification of vaccine candidate proteins through analysis of the group A *Streptococcus* surface proteome. *Nat Biotechnol* 24: 191–197.
8. Poolman JT (1995) Development of a meningococcal vaccine. *Infect Agents Dis* 4: 13–28.
9. van den Elsen J, Vandepitte-Rutten L, Kroon J, Gros P (1999) Bactericidal antibody recognition of meningococcal PorA by induced fit. Comparison of liganded and unliganded Fab structures. *J Biol Chem* 274: 1495–1501.
10. Benoist C, Mathis D (2001) Autoimmunity provoked by infection: how good is the case for T cell epitope mimicry? *Nat Immunol* 2: 797–801.
11. Matzinger P (2002) The danger model: a renewed sense of self. *Science* 296: 301–305.
12. Chatenoud L (2000) [Immunotolerance and autoimmunity]. *Rev Prat* 50: 1497–1505.
13. Gattiker A, Michoud K, Rivoire C, Auchincloss AH, Coudert E, et al. (2003) Automated annotation of microbial proteomes in SWISS-PROT. *Comput Biol Chem* 27: 49–58.
14. Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, et al. (2003) ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res* 31: 3784–3788.
15. Saha S, Bhasin M, Raghava GP (2005) Bcipep: a database of B-cell epitopes. *BMC Genomics* 6: 79.
16. Peters B, Sidney J, Bourne P, Bui HH, Buus S, et al. (2005) The immune epitope database and analysis resource: from vision to blueprint. *PLoS Biol* 3: e91.
17. Toseland CP, Clayton DJ, McSparron H, Hemsley SL, Blythe MJ, et al. (2005) Antigen: a quantitative immunology database integrating functional, thermodynamic, kinetic, biophysical, and cellular data. *Immunity Res* 1: 4.
18. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215: 403–410.
19. Aloy P, Cedano J, Oliva B, Aviles FX, Querol E (1997) 'TransMem': a neural network implemented in Excel spreadsheets for predicting transmembrane domains of proteins. *Comput Appl Biosci* 13: 231–234.
20. Bendtsen JD, Nielsen H, von Heijne G, Brunak S (2004) Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol* 340: 783–795.
21. Kolaskar AS, Tongaonkar PC (1990) A semi-empirical method for prediction of antigenic determinants on protein antigens. *FEBS Lett* 276: 172–174.
22. Saha S, Raghava GP (2006) Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. *Proteins* 65: 40–48.
23. Saha S, Raghava GP (2004) BcPred: Prediction of Continuous B-Cell Epitopes in Antigenic Sequences Using Physico-chemical Properties. 197–204 p.
24. Fitzpatrick DA, Greevey CJ, McInerney JO (2005) Evidence of positive Darwinian selection in putative meningococcal vaccine antigens. *J Mol Evol* 61: 90–98.
25. Vermont C, van den Doolbasteen G (2002) *Neisseria meningitidis* serogroup B: laboratory correlates of protection. *FEMS Immunol Med Microbiol* 34: 89–96.
26. Ala'Aldeen DA, Bariello SP (1996) The meningococcal transferrin-binding proteins 1 and 2 are both surface exposed and generate bactericidal antibodies capable of killing homologous and heterologous strains. *Vaccine* 14: 49–53.
27. Ala'Aldeen DA, Stevenson P, Griffiths E, Gorringe AR, Irons LI, et al. (1994) Immune responses in humans and animals to meningococcal transferrin-binding proteins: implications for vaccine design. *Infect Immun* 62: 2984–2990.
28. Turner PC, Thomas CE, Stojiljkovic I, Elkins C, Kizel G, et al. (2001) Neisserial TonB-dependent outer-membrane proteins: detection, regulation and distribution of three putative candidates identified from the genome sequences. *Microbiology* 147: 1277–1290.
29. Cazalet C, Rusniok C, Bruggemann H, Zidane N, Magnier A, et al. (2004) Evidence in the *Legionella pneumophila* genome for exploitation of host cell functions and high genome plasticity. *Nat Genet* 36: 1165–1173.
30. Weeratna R, Stabler DA, Edelstein PH, Ripley M, Marrie T, et al. (1994) Human and guinea pig immune responses to *Legionella pneumophila* protein antigens OmpS and Hsp60. *Infect Immun* 62: 3454–3462.
31. Chang B, Kura F, Amemura-Maekawa J, Koizumi N, Watanabe H (2005) Identification of a novel adhesion molecule involved in the virulence of *Legionella pneumophila*. *Infect Immun* 73: 4272–4280.
32. Beghetto E, Gargano N, Ricci S, Garufi G, Peppoloni S, et al. (2006) Discovery of novel *Streptococcus pneumoniae* antigens by screening a whole-genome lambda-display library. *FEMS Microbiol Lett* 262: 14–21.
33. Maiorca D, Margarit I, Rinaudo CD, Masignani V, Mora M, et al. (2005) Identification of a universal Group B streptococcus vaccine by multiple genome screen. *Science* 309: 148–150.
34. Adamou JE, Heinrichs JH, Erwin AL, Walsh W, Gayle T, et al. (2001) Identification and characterization of a novel family of pneumococcal proteins that are protective against sepsis. *Infect Immun* 69: 949–958.
35. Wizemann TM, Heinrichs JH, Adamou JE, Erwin AL, Kunsch C, et al. (2001) Use of a whole genome approach to identify vaccine molecules affording protection against *Streptococcus pneumoniae* infection. *Infect Immun* 69: 1593–1598.
36. Chiavolini D, Memmi G, Maggi T, Iannelli F, Pozzi G, et al. (2003) The three extra-cellular zinc metalloproteinases of *Streptococcus pneumoniae* have a different impact on virulence in mice. *BMC Microbiol* 3: 14.
37. Liu DF, Mason KW, Mastri M, Pazirandeh M, Cutter D, et al. (2004) The C-terminal fragment of the internal 110-kilodalton passenger domain of the Hap protein of nontypeable *Haemophilus influenzae* is a potential vaccine candidate. *Infect Immun* 72: 6961–6968.
38. El-Adhami W, Kyd JM, Bastin DA, Cripps AW (1999) Characterization of the gene encoding a 26-kilodalton protein (OMP26) from nontypeable *Haemophilus influenzae* and immune responses to the recombinant protein. *Infect Immun* 67: 1935–1942.
39. Murphy TF, Kirkham C, Lessig AJ (2006) Construction of a mutant and characterization of the role of the vaccine antigen P6 in outer membrane integrity of nontypeable *Haemophilus influenzae*. *Infect Immun* 74: 5169–5176.
40. Janson H, Ruan M, Forsgren A (1993) Limited diversity of the protein D gene (hpd) among encapsulated and nonencapsulated *Haemophilus influenzae* strains. *Infect Immun* 61: 4546–4552.
41. Zagursky RJ, Ooi P, Jones KF, Fiske MJ, Smith RP, et al. (2000) Identification of a *Haemophilus influenzae* 5'-nucleotidase protein: cloning of the nucA gene and immunogenicity and characterization of the NucA protein. *Infect Immun* 68: 2525–2534.
42. Corech R, Rao A, Laxova A, Moss J, Rock MJ, et al. (2005) Early immune response to the components of the type III system of *Pseudomonas aeruginosa* in children with cystic fibrosis. *J Clin Microbiol* 43: 3956–3962.
43. Goure J, Broz P, Attree O, Cornelis GR, Attree I (2005) Protective anti-V antibodies inhibit *Pseudomonas* and *Yersinia* translocon assembly within host membranes. *J Infect Dis* 192: 218–225.
44. Moss J, Ehrmantraut ME, Banwart BD, Frank DW, Barbieri JT (2001) Sera from adult patients with cystic fibrosis contain antibodies to *Pseudomonas aeruginosa* type III apparatus. *Infect Immun* 69: 1185–1188.
45. Price BM, Galloway DR, Baker NR, Gilleland LB, Staczek J, et al. (2001) Protection against *Pseudomonas aeruginosa* chronic lung infection in mice by genetic immunization against outer membrane protein F (OprF) of *P. aeruginosa*. *Infect Immun* 69: 3510–3515.
46. Thomas LD, Kyd JM, Bastin DA, Dunkley ML, Cripps AW (2003) Immunisation with non-integral OMPs promotes pulmonary clearance of *Pseudomonas aeruginosa*. *FEMS Immunol Med Microbiol* 37: 155–160.
47. Umelo-Njaka E, Nomellini JF, Bingle WH, Glasier LG, Irvin RT, et al. (2001) Expression and testing of *Pseudomonas aeruginosa* vaccine candidate proteins prepared with the Caulobacter crescentus S-layer protein expression system. *Vaccine* 19: 1406–1415.
48. Duchene M, Barron C, Schweizer A, von Specht BU, Domdey H (1989) *Pseudomonas aeruginosa* outer membrane lipoprotein I gene: molecular cloning, sequence, and expression in *Escherichia coli*. *J Bacteriol* 171: 4130–4137.
49. Worgall S, Krause A, Rivara M, Hee KK, Vintayen EV, et al. (2005) Protection against *P. aeruginosa* with an adenovirus vector containing an OprF epitope in the capsid. *J Clin Invest* 115: 1281–1289.
50. Abdissa A, Asrat D, Kronvall G, Shittu B, Achiko D, et al. (2006) High diversity of group A streptococcal emm types among healthy schoolchildren in Ethiopia. *Clin Infect Dis* 42: 1362–1367.
51. Okamoto S, Tamura Y, Terao Y, Hamada S, Kawabata S (2005) Systemic immunization with streptococcal immunoglobulin-binding protein Sib 35 induces protective immunity against group A: *Streptococcus* challenge in mice. *Vaccine* 23: 4852–4859.
52. Vohra H, Dey N, Gupta S, Sharma AK, Kumar R, et al. (2005) M protein conserved region antibodies opsonise multiple strains of *Streptococcus pyogenes* with sequence variations in C-repeats. *Res Microbiol* 156: 575–582.
53. McMillan DJ, Davies MR, Good MF, Srivastava KS (2004) Immune response to superoxide dismutase in group A streptococcal infection. *FEMS Immunol Med Microbiol* 40: 249–256.
54. Guzman CA, Talay SR, Molinari G, Medina E, Chhatwal GS (1999) Protective immune response against *Streptococcus pyogenes* in mice after intranasal vaccination with the fibronectin-binding protein SfbI. *J Infect Dis* 179: 901–906.
55. Seepersaud R, Hannify SB, Mayne P, Sizer P, Le Page R, et al. (2005) Characterization of a novel leucine-rich repeat protein antigen from group B streptococci that elicits protective immunity. *Infect Immun* 73: 1671–1683.
56. Banks DJ, Porcella SF, Barbier KD, Beres SB, Philips LE, et al. (2004) Progress toward characterization of the group A *Streptococcus* metagenome: complete genome sequence of a macrolide-resistant serotype M6 strain. *J Infect Dis* 190: 727–738.
57. McArthur JD, Walker MJ (2006) Domains of group A streptococcal M protein that confer resistance to phagocytosis, opsonization and protection: implications for vaccine development. *Mol Microbiol* 59: 1–4.
58. Moyle PM, Olive C, Ho MF, Burgess M, Karpali L, et al. (2006) Method for the synthesis of multi-epitopic *Streptococcus pyogenes* lipopeptide vaccines using native chemical ligation. *J Org Chem* 71: 6846–6850.
59. Li B, Jiang L, Song Q, Yang J, Chen Z, et al. (2005) Protein microarray for profiling antibody responses to *Yersinia pestis* live vaccine. *Infect Immun* 73: 3734–3739.
60. Philipovskiy AV, Cowan C, Wulff-Strobel CR, Burnett SH, Kerschen EJ, et al. (2005) Antibody against V antigen prevents Yop-dependent growth of *Yersinia pestis*. *Infect Immun* 73: 1532–1542.
61. Benner GE, Andrews GP, Byrne WR, Strachan SD, Sample AK, et al. (1999) Immune response to *Yersinia* outer proteins and other *Yersinia pestis* antigens after experimental plague infection in mice. *Infect Immun* 67: 1922–1928.

62. Mattoo S, Miller JF, Cotter PA (2000) Role of *Bordetella bronchiseptica* fimbriae in tracheal colonization and development of a humoral immune response. *Infect Immun* 68: 2024–2033.

63. Keil DJ, Burns EH, Kisker WR, Bemis D, Fenwick B (1999) Cloning and immunologic characterization of a truncated *Bordetella bronchiseptica* filamentous hemagglutinin fusion protein. *Vaccine* 18: 860–867.

64. Hausman SZ, Burns DL (2000) Use of pertussis toxin encoded by ptx genes from *Bordetella bronchiseptica* to model the effects of antigenic drift of pertussis toxin on antibody neutralization. *Infect Immun* 68: 3763–3767.

65. Vianu M, Zouali M (2005) Effect of the B cell superantigen protein A from *S. aureus* on the early lupus disease of (NZBxNZW) F1 mice. *Mol Immunol* 42: 849–855.

66. Weichhart T, Horky M, Sollner J, Gangl S, Henics T, et al. (2003) Functional selection of vaccine candidate peptides from *Staphylococcus aureus* whole-genome expression libraries in vitro. *Infect Immun* 71: 4633–4641.

67. Vytytska O, Nagy E, Blugel M, Meyer HE, Kurzbauer R, et al. (2002) Identification of vaccine candidate antigens of *Staphylococcus aureus* by serological proteome analysis. *Proteomics* 2: 580–590.

68. Lugtenberg B, van Boxtel R, Evenberg D, de Jong M, Storm P, et al. (1986) Biochemical and immunological characterization of cell surface proteins of *Pasteurella multocida* strains causing atrophic rhinitis in swine. *Infect Immun* 52: 175–182.

69. Mascarell L, Fayolle C, Bauche C, Ladant D, Leclerc C (2005) Induction of neutralizing antibodies and Th1-polarized and CD4-independent CD8+ T-cell responses following delivery of human immunodeficiency virus type 1 Tat protein by recombinant adenylate cyclase of *Bordetella pertussis*. *J Virol* 79: 9872–9884.

70. Dias WO, Horton DS, Takahashi CM, Raw I (1994) A *Bordetella pertussis* acellular vaccine candidate: antigenic characterization and antibody induction. *Braz J Med Biol Res* 27: 2607–2611.

71. Novotny P, Chubb AP, Cownley K, Charles IG (1991) Biologic and protective properties of the 69-kDa outer membrane protein of *Bordetella pertussis*: a novel formulation for an acellular pertussis vaccine. *J Infect Dis* 164: 114–122.

72. Askelof P, Rodmalm K, Wrangsell G, Larsson U, Svenson SB, et al. (1990) Protective immunogenicity of two synthetic peptides selected from the amino acid sequence of *Bordetella pertussis* toxin subunit S1. *Proc Natl Acad Sci U S A* 87: 1347–1351.

73. Novotny P, Chubb AP, Cownley K, Charles IG (1991) A novel bivalent acellular pertussis vaccine based on the 69 kDa protein and FHA. *Dev Biol Stand* 73: 243–249.

74. Erkkila L, Laitinen K, Haasio K, Tiirila T, Jauhainen M, et al. (2004) Heat shock protein 60 autoimmunity and early lipid lesions in cholesterol-fed C57BL/6JBom mice during *Chlamydia pneumoniae* infection. *Atherosclerosis* 177: 321–328.

75. Sambri V, Donati M, Storni E, Di Leo K, Agnusdei M, et al. (2004) Experimental infection by *Chlamydia pneumoniae* in the hamster. *Vaccine* 22: 1131–1137.

76. Montigiani S, Falugi F, Scarselli M, Finco O, Petracca R, et al. (2002) Genomic approach for analysis of surface proteins in *Chlamydia pneumoniae*. *Infect Immun* 70: 368–379.

77. Finco O, Bonci A, Agnusdei M, Scarselli M, Petracca R, et al. (2005) Identification of new potential vaccine candidates against *Chlamydia pneumoniae* by multiple screenings. *Vaccine* 23: 1178–1188.

78. Ring GO (1927) Ocular Disease, Notably of the Uveal Tract, Induced by *Streptococcus Viridans* Infection. *Trans Am Ophthalmol Soc* 25: 93–104.

79. Nguyen L, Pieters J (2005) The Trojan horse: survival tactics of pathogenic mycobacteria in macrophages. *Trends Cell Biol* 15: 269–276.

80. Birkeland HC, Stenmark H (2004) Protein targeting to endosomes and phagosomes via FYVE and PX domains. *Curr Top Microbiol Immunol* 282: 89–115.

81. Vieira OV, Botelho RJ, Rameh L, Brachmann SM, Matsuo T, et al. (2001) Distinct roles of class I and class III phosphatidylinositol 3-kinases in phagosome formation and maturation. *J Cell Biol* 155: 19–25.

82. Lindmo K, Stenmark H (2006) Regulation of membrane traffic by phosphoinositide 3-kinases. *J Cell Sci* 119: 605–614.

83. Bannantine JP, Griffith RS, Viratyosin W, Brown WJ, Rockey DD (2000) A secondary structure motif predictive of protein localization to the chlamydial inclusion membrane. *Cell Microbiol* 2: 35–47.

84. Miyamoto Y, Yamauchi J, Sanbe A, Tanoue A (2007) Dock6, a Dock-C subfamily guanine nucleotide exchanger, has the dual specificity for Rac1 and Cdc42 and regulates neurite outgrowth. *Exp Cell Res* 313: 791–804.