Biotechnol Lett (2022) 44:535–545
https://doi.org/10.1007/s10529-022-03238-x

ORIGINAL RESEARCH PAPER

Antigenic and conserved peptides from diverse *Helicobacter pylori* antigens

Cecília R. C. Calado

Received: 10 September 2021 / Accepted: 14 February 2022 / Published online: 11 March 2022

Abstract Since the revolutionary finding of *Helicobacter pylori* as a common bacterial infection, that a high research effort for its eradication has been conducted. Epitope based-vaccine presents advantages over protein-based, as they can be designed to contain epitopes from diverse proteins, therefore, more easily representing the immune-variability of the bacterial population, while minimizing the toxicity associated to some whole proteins. In the present work, an iterative method, to design antigenic and conserved B-epitopes from diverse virulent factors of *H. pylori,* was established. The method considered the trade-off between epitopes antigenicity and conservation among the bacterial population. For the method validation, five virulent factors from *H. pylori* were selected. From each virulent factor, two epitopes were predicted, each with twelve residues of aminoacids. The corresponding ten peptides were synthesised and evaluated by enzyme-linked immunosorbent assay using polyclonal antibodies raised against a specific *H. pylori* strain. All ten peptides were recognised by the antibodies and were consequently antigenic and conserved. This result could strongly contribute to the design of a multivalent epitope-based vaccine, representing the immunogenetic variability within the bacterial population, leading to a sustained and effective immunogenic protection.

Keywords *Helicobacter pylori* · Vaccine · Epitopes · Multi-antigens · Peptide

C. R. C. Calado (✉)
CIMOSM - Centro de Investigação em Modelação e Otimização de Sistemas Multifuncionais, ISEL - Instituto Superior de Engenharia de Lisboa, Instituto Politécnico de Lisboa, R. Conselheiro Emídio Navarro 1, 1959-007 Lisboa, Portugal
e-mail: cecilia.calado@isel.pt

Introduction

The microaerophilic gram-negative bacterium *Helicobacter pylori* can be considered one of the most prevalent human pathogens, being present in over half the human population (Li and Perez-Perez 2018). It is believed that *H. pylori* colonization of the gastric mucosa is acquired early in life, usually before the age of 10 years, and mostly from an oral-oral or oral-faecal route, and mainly from mother to sons and among siblings (Goh et al. 2011). After infection and in the absence of antibiotic therapy, it generally persists for life, leading to a gastric chronic inflammation (Kusters et al. 2006; Cover et al. 2009). Depending on multiple factors (such as the bacteria characteristics, the host, and the environment), the inflammatory process can result in severe gastric pathologies such as ulcers, gastric cancer, or MALT lymphoma (Malfertheimer et al. 2017).

Current therapeutic guidelines recommend a first-line treatment comprising of a proton-pump inhibitor and two macrolide antibiotics. Treatment of infection can resolve ulcer disease in 90%

of patients and can result in 75% of cases to complete remission of early MALT lymphomas (Malfertheimer et al. 2017). However, relevant *H. pylori* resistance has been detected with macrolides (e.g., clarithromycin) and fluoroquinolones (e.g., levofloxacin) (Malfertheiner et al. 2011). Furthermore, even in cases of bacteria eradication, the re-infection may occur. Therefore, it is imperative to focus on vaccine development. However, and until now, an efficient vaccine is still to be developed.

One of the main obstacles to achieve an efficient immunization against *H. pylori*, is the bacteria high genetic diversity. This diversity results from high mutation and recombinant events associated to an inefficient DNA repair mechanism and high competence for transformation (Ailloud et al. 2019). Consequently, the bacteria population in each host can evolve, presenting niche specific adaptations and selective sweeps according to alterations on the environment as e.g., resulting from antibiotic treatments (Ailloud et al. 2019). Therefore, it is necessary to develop a vaccine covering this highly diverse and dynamic *H. pylori* population.

The target proteins usually selected for a *H. pylori* vaccine are immunogenic virulent factors. A high immunogenicity will result on a strong immune reaction, where the neutralization of virulent factors will impair the bacteria infection process. Examples of very appealing targets are the vacuolating cytotoxin A (VacA) (Moyat and Velin 2016; Liu et al. 2004) and the cytotoxin associated antigen A (CagA) (Takahashi-Kanemitsu et al. 2020). However, VacA and CagA presents a high genetic variability (Gressmann et al. 2005; Linz and Schuster 2007), besides CagA is not present in all bacteria strains (Gressmann et al. 2005). A vaccine based on diverse proteins, i.e., a multi-antigenic vaccine, covering the bacteria immunogenetic diversity, will minimize the probability of the bacteria escaping the immune system. According to this, diverse researchers have proposed vaccines based on multi-proteins as *e.g.*, based on VacA, CagA, and urease sub-unit B (UreB) (Liu et al. 2011), or by CagA, VacA and neutrophil-activating protein (NAP) (Malfertheiner et al. 2008), or by UreB, NAP, adhesin A (HpaA) and heat shock protein 60 (HSP60) (Guo et al. 2017).

On the present work, the following five target virulent factors were considered: VacA, CagA, Adhesin-binding fucosylated histo-blood group antigen (BabA), UreB and Flagellin A (FlaA).

VacA is an exotoxin secreted by *H. pylori* that causes the formation of vacuoles in the host cells and has been considered a good vaccine target (Moyat and Velin 2016; Liu et al. 2004). CagA is present between 50 and 70% of *H. pylori* strains and is delivered into gastric epithelial cells via bacterial type IV secretion, having been considered an oncogenic factor (Takahashi-Kanemitsu et al. 2020), and a target for vaccine construction (Stein et al. 2013). Both, VacA and CagA are therefore, secreted or injected by the bacteria, respectively, leading to modulation of diverse host pathways including the host immune response (Jones et al. 2010). BabA is an adherence factor that mediates the bacteria adhesion to the blood group antigen Lewis B on human epithelial cells (Ansari and Yamaoka 2017), and is therefore relevant in the first steps of the bacterial fixation and colonization, and consequently has been explored as a candidate for vaccine design (Prinz et al. 2003; Bai et al. 2004). The prevalence of BabA depends on the world region, were *e.g.* almost all strains from East Asia expressed it, and 10% of the Western strains lacks this gene (Sterbenc et al. 2019). BabA is an important virulent factor in gastric cancer development. Indeed, it has been shown that bacteria expressing BabA, VacA and CagA, are statistically associated to peptide ulcer and gastric cancer in relation to bacteria not expressing BAbA, but expressing VacA and CagA (Sterbenc et al. 2019). UreB, as with VacA and CagA, has been one of the most explored proteins for vaccine design. UreB is highly relevant for local buffering of gastric mucosa and consequently for bacteria gastric colonization (Voland et al. 2003), and for that several groups have already used UreB to induce immunization in animal models (Gomez-Duarte et al. 1998; Fujii et al. 2004; Yang et al. 2005; Morihara et al. 2007). UreB has been detected in the bacteria cytoplasm and periplasm (Marcus et al. 2013), but also on the bacteria surface (Haas et al. 2000). Flagellin A (FlaA) is a motility factor, involved in the colonization of the host and on the inflammatory processes (Gu et al. 2017) and has been suggested as a good vaccine target (Yan et al. 2003; Hamzehloo et al. 2019).

A limitation in developing multi-protein-based vaccines are the complexity of the production process, time, and costs. A more cost-effective

alternative to a mixture of proteins is an epitope-based vaccine, with epitopes from diverse proteins. These epitopes-based vaccines can be present as a set of peptides, or in alternative, the epitopes can be included in a polypeptide (Purcell et al. 2007; Dudek et al. 2010; Yang et al. 2015; Skwarczynski and Toth 2016; Nandy et al. 2018). These epitopes can cover a higher number of protein targets, and consequently the bacteria immunogenetic diversity. By defining specific regions from each protein, the toxicity associated to whole proteins, as with the toxins VacA and CagA, are also minimised or even avoided. By using specific sequences from each protein, it is also possible to reduce exacerbated inflammatory processes as associated, for example, with the NAP and minimise possible auto-immune reactions by avoiding regions common to human proteins (Roggen 2006; Skwarczynski and Toth 2016).

In the present work, an iterative method, enabling to determine epitopes from diverse target proteins, and that are antigenic and conserved among *H. pylori* strains was implemented. The method preliminary validation was based on the design of ten small peptides, two from each target protein, and evaluating its recognition by polyclonal antibodies raised against a specific *H. pylori* strain. Hopefully, this methodology could potentiate the design of a cost-effective vaccine against *H. pylori* representing conserved and antigenic regions from diverse immunogenic proteins.

Material and methods

Antigenicity analysis

The sequences of the following target proteins of *H. pylori* 26,695 strain were obtained from the ExPASy Proteomics (http://www.expasy.ch/): VacA, CagA, Adhesin-binding fucosylated histo-blood group antigen, UreB and Flagellin A (FlaA).

The protein antigenicity was predicted by the Antigenic Index (AI) as defined by Jameson and Wolf (1988), using the package Protean of the Lasergene (DNASTAR) software. The total antigenic index (TAI) was defined as the sum of the AI for each aminoacid residue on the peptide sequence. For each peptide associated to a specific target protein, the relative T.A.I. was determined considering a maximum of 100.

Conservation analysis

The protein sequences of different *H. pylori* strains were obtained by the SIB BLAST (Basic Local Alignment Search Tool) Network Service tool. Multiple alignments were made using the ClustalW algorithm (Thompson et al. 1994). In average, a total of 30 protein sequences for each target protein was considered. It was considered sequences from *H. pylori* isolated world-wide.

Selection of epitopes along the polypeptide

The selection of epitopes from each target protein was based on the following iterative algorithm (Fig. 1). A.I. was determined along the consensus sequence of the polypeptide; Linear epitopes with 12 aminoacid residues were considered. From each polypeptide, all combinations of linear epitopes were considered. For example, for a polypeptide with 1000 residues long, there will be 1000–12 possible different epitopes, since each epitope may start in any of the 1000 residues. The epitopes considered on the following steps should fulfil the following two conditions: (1) It must include an aminoacid residue with an A.I. > 12; (2) The central aminoacid residue, i.e., at the 6th or 7th position, should present the highest A.I from all residues. A minimum A.I. of 12 was empirically defined to obtain a reasonable number of epitopes from each polypeptide, resulting in the selection, in average, of one epitope in peptides with 40 to 90 aminoacids residues. The epitope conservation was determined only for epitopes that meet the two requirements pointed above. The epitopes that fulfil the two conditions and were present in at least 50% of all sequences, were then tested against the human genome. This last condition aims to minimise auto-immune reactions. For that, a BLAST search was conducted against the *Homo sapiens* taxonomic group.

The dimension of the epitopes, based on 12 residues, was supported on the following: it was intended to predict linear epitopes recognised by B-lymphocytes to promote an immune response against the proteins that can be at the bacteria surface or have been secreted to the extracellular space. B-lymphocytes utilizes the major histocompatibility complexes (MHC) class II (MHCII), to process the antigens bounded to the B-cell receptors and internalized antigens, and subsequently to present the epitopes

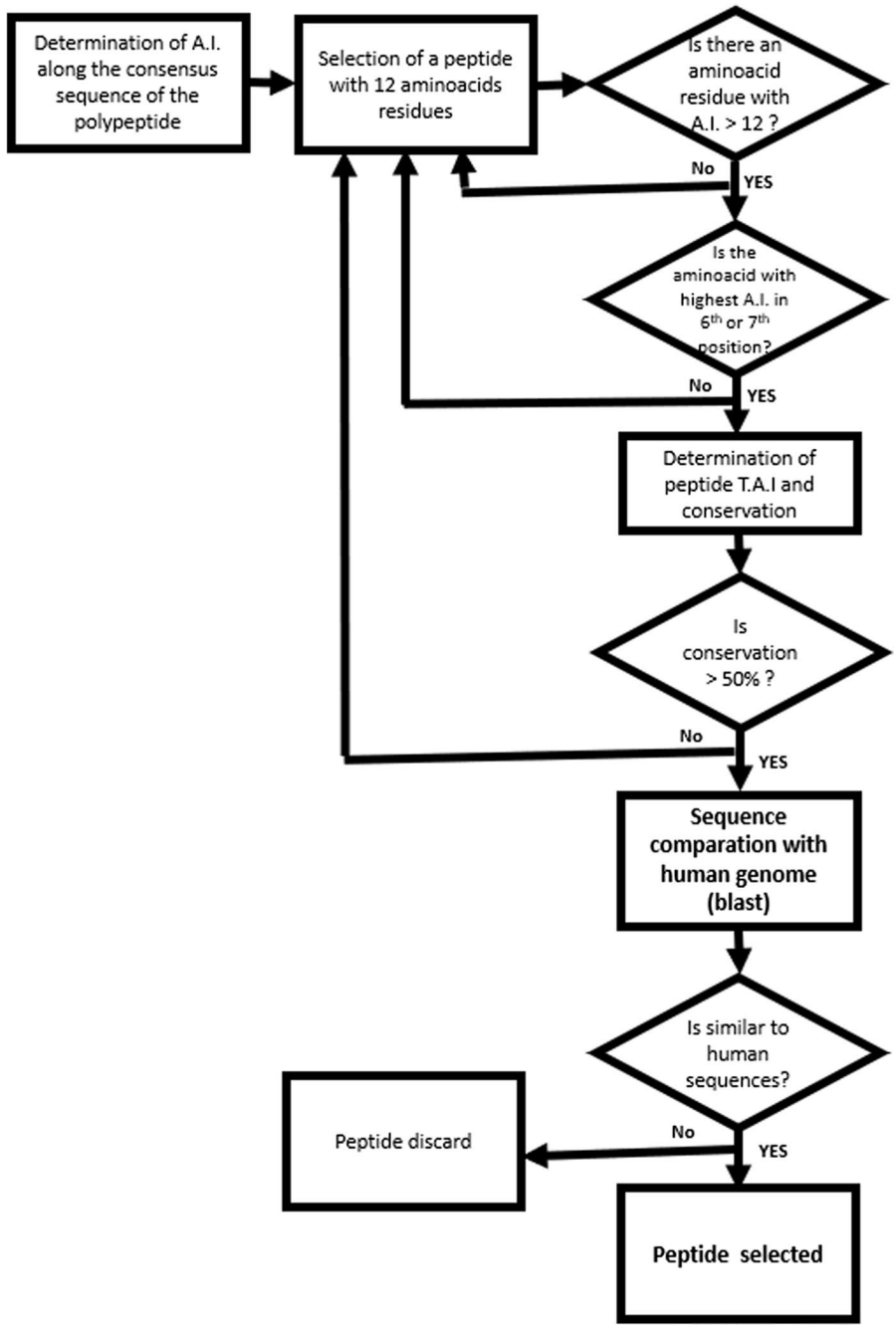

Fig. 1 Iterative method to determine epitopes along the polypeptide that simultaneously maximize the total of antigenic index (TAI) and its conservation among bacteria

to CD4 + T lymphocytes (Adler et al. 2017). The MHC class I (MHCI) usually presents epitopes with dimensions between 8 and 10 residues (Janeway et al. 2001), whereas most MHCI peptide ligands have 9 residues and, consequently, it is generally preferable to predict peptides with that size (Sanchez-Trincado et al. 2017). The MHCII usually presents epitopes with higher dimensions (Janeway et al. 2001), although the peptide binding core to this complex is also based only on 9 residues (Sanchez-Trincado et al. 2017). For that, there are authors that design small linear B-epitopes, as low as 5 to 6 residues (Amela et al. 2007). This low dimension, however, most probably will result in epitopes with low specificity.

Considering the above mentioned, and since CD4 + T lymphocytes usually recognise peptides between 12 and 16 residues in the context of MHCII (Hemmer et al. 2000), epitopes based on 12 residues were defined.

For validation of the iterative method, the following two epitopes from each target protein were selected: One peptide with the highest T.A.I. and conserved at least in 50% of the strains; a second peptide with the highest conservation, independently of the T.A.I.

Peptide synthesis

Peptides with ten aminoacids residues, each designed in the previous step, were chemically synthetized by Thermo Biopolymers (Germany).

Antigenic analysis of the designed peptides

Indirect Enzyme-Linked Immunosorbent Assay (ELISA) were performed in 96-well plates (Nunc, Immobilizer™—Amino Plates). The plates were coated with the synthetic peptides, dissolved in a carbonate buffer 0.1 M pH 9.6 (100 µl per well, of a 0.1 mg/ml solution), as instructed by the manufacturer. After overnight incubation at 4 °C, the plates were washed four times with 300 µl of phosphate-buffered saline, with 0.1% of tween 20 (PBST). The wells were blocked with 200 µl of blocking buffer (PBST with 5% of low-fat dry milk), for 1 h at 37 °C, with gentle agitation. The plates were washed with PBST (3 times, 250 µl each time), and then incubated with, the primary antibody (Rabbit antibody to *H. pylori*, Biodesign, Meridian Life Science) for 1 h, at 37 °C, with gentle agitation. The antibody was previously prepared (dilution 1:1500) in the blocking buffer (100 µl per well). The plates were washed with PBST (4 times, 250 µl each time), and then incubated with the secondary antibody (Anti-rabbit IgG, peroxidase-linked species-specific whole antibody, NA934) for 1 h, at 37 °C, with gentle agitation. This antibody was previously diluted (1:1500) in the blocking buffer (100 µl per well). The plates were washed with PBST (6 times, 200 µl each time). HRP activity was detected by chemiluminescence: a working solution of SuperSignal ELISA Pico Chemiluminescent Substrate (from Pierce) was prepared as instructed by the supplier and 100 µl was added to each well. For controls, in some cases, no peptides were added, while in other cases, the peptides were added but no primary antibody was applied. The chemiluminescence emission was measured using a microplate reader Synergy™ 2 (Biotek, USA). Analysis of variance (ANOVA) was conducted in Microsoft Excel.

Results

By application of the Jameson-Wolf method, it was predicted in all five target proteins several antigenic rich regions (Fig. 2), that are in accordance with these proteins being immunogenic, i.e., that leads to the production of antibodies in infected patients (Kimmel et al. 2000; Hass et al. 2002; Mini et al. 2006; Vitoriano et al. 2011). Since in all these proteins, the predicted antigenic regions were intercalated with regions with low antigenicity (Fig. 2), it is critical to choose the correct regions.

The iterative method described in Fig. 1, enabled the prediction of linear and sequential epitopes that are antigenic, conserved, and specific to the pathogen, i.e., are not present in the human genome. From the VacA, CagA, UreB, FlaA and BabA, it was possible to define 14, 27, 7, 7 and 21 epitopes along the polypeptide, respectively, presenting at least one residue with A.I. > 12 and with the highest A.I. at the centre of the epitope (Fig. 3). From all these epitopes, 7, 15, 7, 6 and 6 from VacA, CagA, UreB, FlaA and BabA, were present in at least 50% of the bacteria strains, respectively (Fig. 3).

From the multiple alignments, an almost inverse relationship was observed for VacA, CagA and BabA between the peptides TAI and its conservation among diverse strains (Fig. 3). This is according to the fact that most variable regions are usually the ones most recognized by the host immune cells, i.e., more antigenic, since the genetic variations enables the bacteria to escape from the immune system. In contrary, the most conserved regions, are, or critical for the protein function and consequently for the bacteria surviving, or not so accessible to the host immune system.

To validate the method, two epitopes for each target protein were selected, and its corresponding peptides synthesised and tested with an ELISA (Table 1). All ten peptides were recognized by polyclonal antibodies raised against a specific *H. pylori* strain (ATCC 43,504) (Fig. 4). Therefore, the predicted

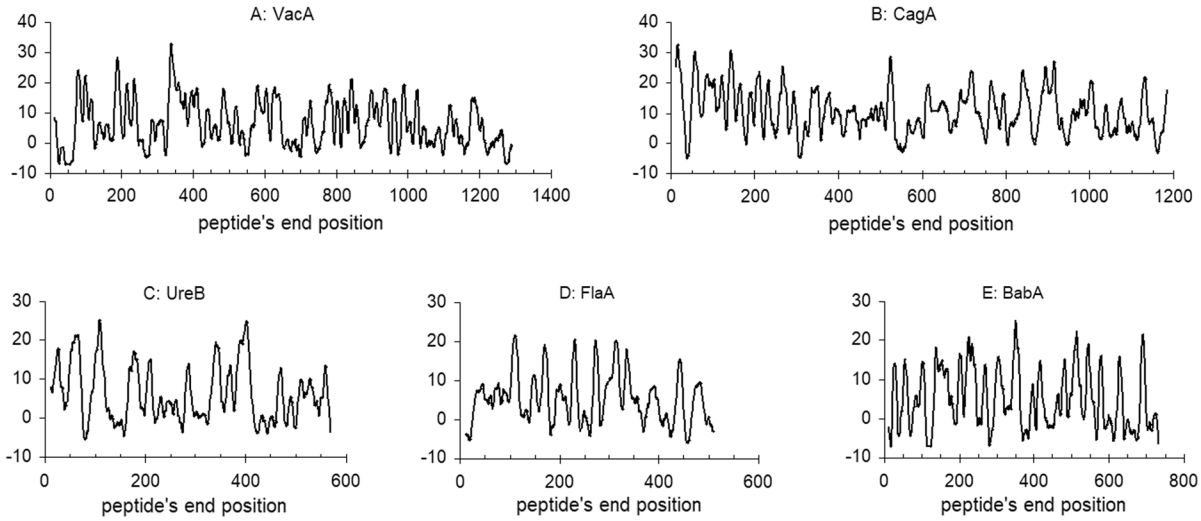

Fig. 2 **A**ntigenic index along the polypeptide of the following five target proteins. VacA; CagA; UreB; FlaA; and BabA

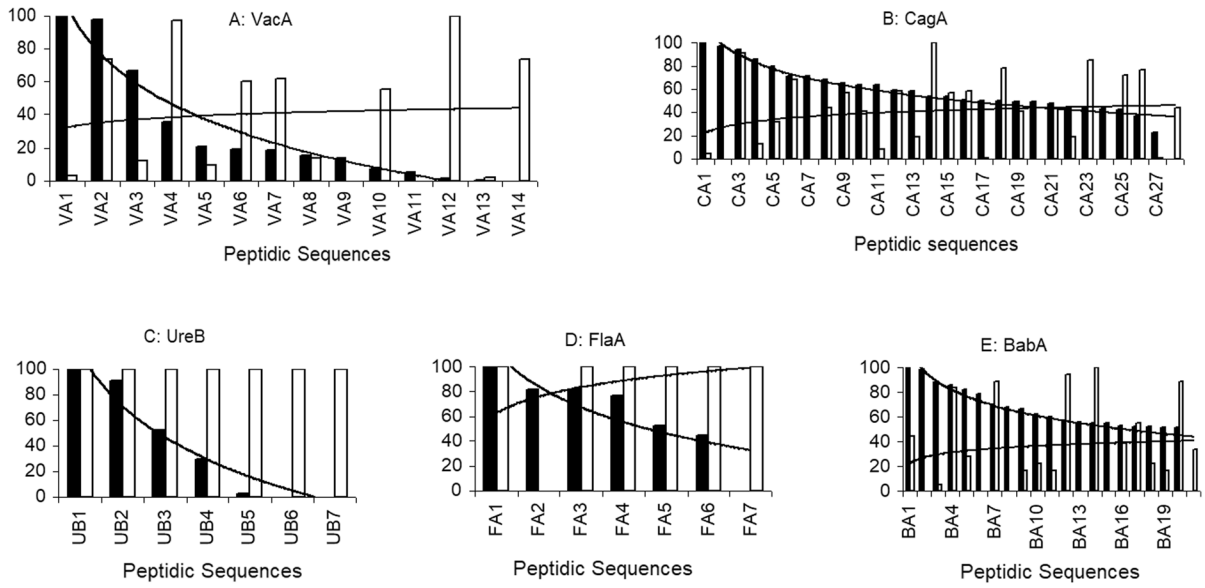

Fig. 3 Conservation (white bars) vs antigenicity (black bars) of peptide (with 12 aminoacids residues) of the following proteins: VacA; CagA; UreB; FlaA; and BabA. The lines are trendlines of the antigenicity and conservation. Data were normalised to the first peptide

peptides were all antigenic and conserved since were recognised by antibodies raised against a specific strain.

It is worth to point that the *H. pylori* strain (ATCC 43,504) used to generate the polyclonal antibodies produces all the five target proteins (*i.e.*, VacA, CAgA, UreB, FlaA and BabA) (Vitoriano et al. 2011; Kinoshita-Daitoku et al. 2020). For example, the same polyclonal antibodies used in the present manuscript also recognised in an 2D-western blotting, the whole proteins CagA, UreB, and FlaA obtained from diverse *H. pylori* strains (Vitoriano et al. 2011).

ANOVA of ELISA signals (R.L.U) pointed that, for each target protein, there was no significant

Table 1 Sequences of 10 peptides from diverse virulent factors of *H. pylori*

PROTEIN	PEPTIDE	AMINOACID SEQUENCE
BabA	BA4	SKKNNPYSPQGI
	BA14	KATNFLGKNNKL
CagA	CA3	DPDQKPIVDKND
	CA14	FNQKNMKDYSDS
FlaA	FA1	AQDGQTTESRKA
	FA3	ADIKKNDSDGRL
UreB	UB1	KGGNKDMQDGVK
	UB2	FGRLKEEKGDND
VacA	VA12	NRVGSGAGRKAS
	VA2	PEGGYKDKPNNT

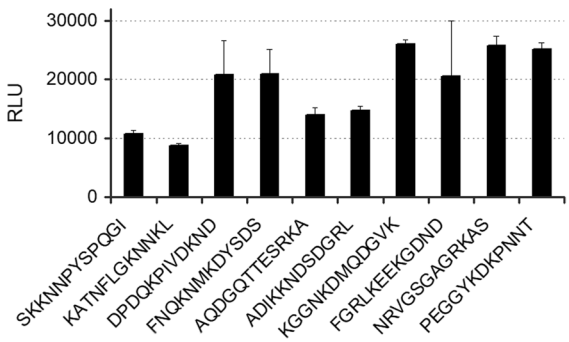

Fig. 4 ELISA chemiluminescence for the ten peptides described in Table 1, obtained from BabA, CagA, FlaA, UreB and VacA, respectively. Each result represents triplicates. (RLU: Relative Luminescence Units)

difference between the recognition of the two peptides ($p > 0.1$), with exception for the BabA peptides ($p < 0.001$) (Fig. 4). This last observation is in accordance with the fact that the second peptide of BabA presented 55% of the A.I. of the first peptide (Fig. 3). It was also observed that peptides from CagA, VacA and UreB presented significant higher R.L.U. than the peptides from FlaA and BabA ($p < 0.001$) (Fig. 4).

Discussion

An epitope-based vaccine, representing diverse target proteins of *H. pylori* can theoretically represent a cost-effective vaccine covering the immune-variability of the *H. pylori* population. These types of vaccines will also minimise possible toxicities associated to some proteins, as with the toxins CagA and VacA, minimises auto-immune reactions and exacerbate inflammatory processes, while leading to more rapid and economic production processes. However, to achieve all of this, it is critical to select epitopes from each target protein that are both antigenic and conserved.

The prediction of the proteins antigenicity was based on the Jameson-Wolf Antigenic Index algorithm, since it predicts linear antigenic determinants based on structural protein properties, as surface probability, backbone flexibility and secondary structure prediction (Jameson and Wolf 1988). The algorithm predicts the surface contour profile of the protein, *i.e.*, predicts which regions are at the protein surface. The algorithm application to a large number of proteins resulted in a high accuracy to predict antigenic determinants (Kolaskar and Tongaonkar 1990), and consequently it has been applied to design diverse vaccines as pointed e.g., in Amela et al. (2007), Frikha-Gargouri et al. (2008), Da Silva et al. (2009), Zhao et al. (2011), Li et al. (2014), and Hossain et al. (2017).

The algorithm is particularly relevant to predict linear and sequential B-lymphocytes epitopes (Roggen et al. 2006), as intended for *H. pylori,* to promote a humoral immune response mediated by B-lymphocytes. The subsequent antibodies produced by the B-cells will recognise the target proteins present at the bacteria surface or that were secreted towards the extracellular space. The identification of the bacteria ultimately leads to its neutralization. In a similar mode, the protein recognition will also lead to the proteins activity neutralization, and consequently on the infection process impairment. For example, UreB and BabA are proteins that can potentiate the bacteria recognition by the immune system, where all five targets are virulent factors, and whose neutralization may impair the infection process. Three of the target proteins are critical for the bacteria survival in the stomach hostile environment due to their function as medium buffer (UreB), bacteria adhesion (BabA) and movement (FlaA). The remaining two target proteins are critical to regulate diverse pathways of the human host towards the bacteria maintenance, including the immune host response (VacA and CagA). It is also worth mentioning that the prediction of linear epitopes, instead of conformational ones, is crucial when working with small dimension epitopes.

In the present work, it was evaluated the antigenicity and conservation for VacA, CagA, UreB, FlaA and BabA (Fig. 2). The high conservation observed among FlaA and UreB was according to other authors. For example, FlaA was one of the most conserved immunoproteins in a study including the proteome of ten different strains (Vitoriano et al. 2011). In that study, proteins were first separated in a 2D-electrophoresis gel, and subsequently recognized in a Western blotting by polyclonal antibodies. The antibodies detected FlaA in eight out of the ten *H. pylori* strains proteome. This pointed FlaA as the most conserved protein among the evaluated strains. Similar results were obtained by other immunoproteomics works, where FlaA, FlaB and GroEL, obtained from 3 different strains, were the only proteins that reacted with serum from all 5 infected patients (Mini et al. 2006). In another study, FlaA was the only protein that reacted with all the 16 human sera from infected patients (Kimmel et al. 2000). In this last work, the other more immune-conserved proteins were GroEl, that was recognised in 14 sera, and UreB and NAP that were recognized by 11 out of 16 sera. Hass et al. (2002) also observed that, from the target proteins evaluated in the present work, the most immune-conserved was UreB, which was detected in 14 out of 24 human sera from infected patients. There are some works pointing the genetic diversity of the *flaA* gene (Forbes et al. 1995), however, most of the diversity were due from nucleotide synonymous substitutions, i.e., nucleotide substitutions, which do not result in aminoacidic modifications. Indeed, it has been detected a ratio between synonymous and non-synonymous substitutions for the *flaA* gene between 52 and 58 (Suerbaum 1998), much higher than the observed for the *cagA* gene, between 2 and 5 (van der Ende et al. 1998; Achtman and Suerbaum 2000), or for the *vacA* gene, between 1.7 and 4 (Atherton et al. 1999), or for the *babA* gene, between 2.3 and 5.2 (Pride et al. 2001; Ghose et al. 2002). For VacA, CagA and BabA it was observed an inverse relationship between T.A.I. and conservation (Fig. 3), according to these targets' high genetic variability (Suerbaum et al. 1998; van der Ende et al. 1998; Athtman and Suerbaum 2000; Pride et. al, 2001).

The algorithm presented here, promoted the selection of conserved and antigenic epitopes from all five targets. The predicted two peptides from each target protein were all commonly recognised by the polyclonal antibodies raised against a specific *H. pylori* strain. Therefore, the ten designed peptides are antigenic and conserved. Interestingly, the second peptide from BabA was less antigenic than the first one, i.e., presented a lower ELISA signal in relation to the first peptide. This is according to the prediction for the second peptide of half of the theoretically antigenicity in relation to the first peptide. The ELISA also pointed as more antigenic peptides, the ones obtained from UreB, VacA and CAgA in relation to peptides from FlaA and BabA. This is according to UreB being one of the most abundant proteins of the bacteria, that may represent 15% of the total protein (Hu et al. 1990). The high quantity of the protein will potentiate the immune host recognition. The VacA and CagA high antigenicity are also according to these proteins being major toxins secreted or injected by the bacteria, respectively, which translates into a strong interaction with human cells, including immune cells (Sterbenc et al. 2019). All these observations corroborate the epitopes selection process.

Briefly, the iterative method presented, enabled to predict, from a defined target protein, epitopes that are antigenic and conserved. The method described could therefore potentiate the design of multivalent vaccines, based on epitopes from multi-target proteins, towards the representation of the bacteria immune–variability. The use of multiple epitopes from diverse target proteins instead of the whole protein presents the advantage of enabling covering a higher dimension of antigenic proteins as targets, while minimizing the risk of adverse immune reactions as exacerbated inflammatory processes, autoimmune reactions, and even toxicity associated to whole proteins. Hopefully, this methodology could potentiate the design of a cost-effective vaccines against *H. pylori*.

Acknowledgements *Fundação para a Ciência e Tecnologia*, Portugal, with grant number PTDC/BIO/69242/2006. The present work was partially conducted in the Engineering & Health Laboratory, that resulted from a collaboration protocol established between *Universidade Católica Portuguesa* and IPL.

Funding The authors have not disclosed any funding.

Declarations

Conflict of interest The author declares that they have no conflict of interest.

References

Achtman M, Suerbaum S (2000) Sequence variation in *Helicobacter pylori*. Trends Microbiol 8:57–58. https://doi.org/10.1016/s0966-842x(99)01676-5

Adler LN, Jiang W, Bhamidipati K, Millican M, Macaubas C, Hung S-C, Mellins ED (2017) The other function: class II-restricted antigen presentation by B cells. Front Immunol 8:319. https://doi.org/10.3389/fimmu.2017.00319

Ailloud F, Didelot X, Woltemate S, Pfaffinger G, Overmann J, Bader RC, Schulz C, Malfertheiner M, P, Suerbaum S, (2019) Within-host evolution of *Helicobacter pylori* shaped by niche-specific adaptation, intragastric migrations and selective sweeps. Nat Commun 10:2273. https://doi.org/10.1038/s41467-019-10050-1

Amela I, Cedano J, Querol E (2007) Pathogen proteins eliciting antibodies do not share epitopes with host proteins: a bioinformatics approach. PLoS ONE 2:e512. https://doi.org/10.1371/journal.pone.0000512

Ansari S, Yamaoka Y (2017) *Helicobacter pylori* BabA in adaptation for gastric colonization. World J Gastroenterol 23(23):4158–4169. https://doi.org/10.3748/wjg.v23.i23.4158

Atherton JC, Sharp PM, Cover TL, Gonzalez-Valencia G, Peek RM, Thompson SA, Hawkey CJ, Blaser M (1999) Vacuolating Cytotoxin (vacA) alleles of *helicobacter pylori* comprise two geographically widespread types, m1 and m2, and have evolved through limited recombination. Curr Microbiol 39(4):211–218. https://doi.org/10.1007/s002849900447

Bai Y, Zhang Y-L, Chen Y, Jin J-F, Zhang Z-S, Zhou DY (2004) Cloning and expression and immunogenicity of *Helicobacter pylori* BabA2 gene. World J Gastroenterol 10(17):2560–2562. https://doi.org/10.3748/wjg.v10.i17.2560

Cover TL, Blaser MJ (2009) *Helicobacter pylori* in health and disease. Gastroenterology 136(6):1863–1873. https://doi.org/10.1053/j.gastro.2009.01.073

Da Silva AN, Nascimento EJ, Cordeiro MT, Gil LH, Abath FG, Montenegro SM, Marques ET (2009) Identification of continuous human B-cell epitopes in the envelope glycoprotein of dengue vírus type 3 (DENV-3). PLoS ONE 4:e7425. https://doi.org/10.1371/journal.pone.0007425

Dudek NL, Perlmutter P, Aguilar M-I, Croft NP, Purcell AW (2010) Epitope discovery and their use in peptide based vaccines. Curr Pharm Design 16(28):3149–3157. https://doi.org/10.2174/138161210793292447

Forber KJ, Fang Z, Pennington TH (1995) Allelic variation in the *Helicobacter pylori* flagellin genes flaA and flaB: its consequences for strain typing schemes and population structure. Epidemiol Infect 114(2):257–266. https://doi.org/10.1017/s0950268800057927

Frikha-Gargouri O, Gdoura R, Znazen A, Gargouri B, Gargouri J, Rebai A, Hammami A (2008) Evaluation of an in silico predicted specific and immunogenic antigen from the OmcB protein for the serodiagnosis of Chlamydia trachomatis infections. Bmcmicrobiol 8:217. https://doi.org/10.1186/1471-2180-8-217

Fujii R, Morihara F, Fukushima K, Oku T, Hifumi E, Uda T (2004) Recombinant antigen from *Helicobacter pylori* urease as vaccine against *H. pylori*-associated disease. Biotechnol Bioeng 86(7):737–746. https://doi.org/10.1002/bit.20047

Ghose C, Perez-Perez G, Dominguez-Bello MG (2002) East Asian genotypes of *Helicobacter pylori* strains in Amerindians provide evidence of its ancient human carriage. Proc Nat Acad Sci USA 99(23):15107–15111. https://doi.org/10.1073/pnas.242574599

Goh KL, Chan WK, Shiota S, Yamaoka Y (2011). Epidemiology of *Helicobacter pylori* infection and public health implications. Helicobacter 16 Suppl 1(01):1–9. https://doi.org/10.1111/j.1523-5378.2011.00874.x.

Gomez-Duarte OG, Lucas B, Yan ZX, Panthel K, Haas R, Meyer TF (1998) Protection of mice against gastric colonization by *Helicobacter pylori* by single oral dose immunization with attenuated Salmonella typhimurium producing urease subunits A and B. Vaccine 16(5):460–471. https://doi.org/10.1016/s0264-410x(97)00247-8

Gressmann H, Linz B, Ghai R, Pleissner K-P, Schlapbach R, Yamaoka Y, Kraft C, Suerbaum S, Meyer TF, Achtman M (2005) Gain and loss of multiple genes during the evolution of *Helicobacter pylori*. PLoS Genet 1(4):e43. https://doi.org/10.1371/journal.pgen.0010043

Gu H (2017) Role of flagella in the pathogenesis of *Helicobacter pylori*. Curr Microbiol 74(7):863–869. https://doi.org/10.1007/s00284-017-1256-4

Guo L, Yang H, Tang F, Yin R, Liu H, Gong X, Wei J, Zhang Y, Xu G, Liu K (2017) Oral immunization with a multivalent epitope-based vaccine, based on NAP, urease, HSP60, and HpaA, provides therapeutic effect on *H. pylori* infection in mongolian gerbils. Front Cell Infect Microbiol 7:349. https://doi.org/10.3389/fcimb.2017.00349

Haas G, Karaali G, Ebermayer K, Metzger WG, Lamer S, Zimny-Arndt U et al (2002) Immunoproteomics of *Helicobacter pylori* infection and relation to gastric disease. Proteomics 2(3):313–324. https://doi.org/10.1002/1615-9861(200203)2:3%3c313::aid-prot313%3e3.0.co;2-7

Hamzehloo Z, Mosayebi G, Khansarinejad B, Zolfaghari M, Abtahi H (2019) Antigenicity identification of a novel recombinant multi-epitope antigen based on FlaA and UreB antigens of *Helicobacter pylori*. Jundishapur J Microbiol 2(5):e66502. https://doi.org/10.5812/jjm.66502

Hemmer B, Kondo T, Gran B, Pinilla C, Cortese I, Pascal J, Tzou A, McFarland HF, Houghten R, Martin R (2000) Minimal peptide length requirements for $CD4^+$ T cell clones—implications for molecular mimicry and T cell survival. Int Immunol 12(3):375–383. https://doi.org/10.1093/intimm/12.3.375

Hossain G, Ueda K, Hossain G, Ueda K (2017) Investigation of a novel hepatitis B virus surface antigen (HBsAg) escape mutant affecting immunogenicity. PLoS ONE 12(1):e0167871. https://doi.org/10.1371/journal.pone.0167871

Hu LT, Mobley HL (1990) Purification and *N-terminal* analysis of urease from *Helicobacter pylori*. Infect Immun 58(4):992–998

Jameson BA, Wolf H (1988) The antigenic index: a novel algorithm for predicting antigenic determinants. Comput Appl Biosci 4(1):181–186. https://doi.org/10.1093/bioinformatics/4.1.181

Janeway CA, Travers P, Walport M, Shlomchik MJ (2001) Immunobiology: the immubne system in health and disease, 5th edn. Garlan Sciences, New York

Jones KR, Whitmire JM, Merrell DS (2010) A tale of two toxins: *Helicobacter pylori* CagA and VacA modulate host pathways that impact disease. Front Microbiol 1:115. https://doi.org/10.3389/fmicb.2010.00115

Li J, Perez-Perez GI (2018) *Helicobacter pylori* the latent pathogen or an ancestral commensal organism. Front Microbiol 9:609. https://doi.org/10.3389/fmicb.2018.00609

Li W-Z, Wang J, Long R, Su G-H, Bukhory D-K, Dai J, Jin N, Huang S-Y, Jia P, Li T, Fan C, Liu K, Wang Z (2014) Novel antibody against a glutamic acid-rich human fibrinogen-like protein 2-derived peptide near Ser91 inhibits hfgl2 prothrombinase activity. PLoS ONE 9(4):e94551. https://doi.org/10.1371/journal.pone.0094551

Linz B, Schuster SC (2007) Genomic diversity in Helicobacter and related organisms. Res Microbiol 158(10):737–744. https://doi.org/10.1016/j.resmic.2007.09.006

Liu K-Y, Shi Y, Luo P, Yu S, Chen L, Zhao Z, Mao X-H, Guo G, Wu C, Zou Q-M (2011) Therapeutic efficacy of oral immunization with attenuated *Salmonella typhimurium* expressing *Helicobacter pylori* CagA, VacA and UreB fusion proteins in mice model. Vaccine 29:6679–6685. https://doi.org/10.1016/j.vaccine.2011.06.099

Liu XL, Li S-Q, Liu C-J, Tao H-X, Zhang Z-S (2004) Antigen epitope of *Helicobacter pylori* vacuolating cytotoxin A. World J Gastroenterol 10(16):2340–2343. https://doi.org/10.3748/wjg.v10.i16.2340

Kimmel B, Bosserhoff A, Frank R, Gross R, GoebelBeier WD (2000) Identification of immunodominant antigens from *Helicobacter pylori* and evaluation of their reactivities with sera from patients with different gastroduodenal pathologies. Infeccion and Immunity 68:915–920. https://doi.org/10.1128/IAI.68.2.915-920.2000

Kinoshita-Daitoku R, Ogura Y, Kiga K, Maruyama F, Kondo T, Nakagawa I, Hayashi T, Mimuro H (2020) Complete genome sequence of *Helicobacter pylori* strain ATCC 43504, a type strain that can infect gerbils. Microbiol Resour Announc 9(18):e00105-e120. https://doi.org/10.1128/MRA.00105-20

Kolaskar AS, Tongaonkar PC (1990) A semi-empirical method for prediction of antigenic determinants on protein antigens. FEBS Lett 276(1–2):172–174. https://doi.org/10.1016/0014-5793(90)80535-q

Kusters JG, van Vliet AH, Kuipers EJ (2006) Pathogenesis of *Helicobacter pylori* infection. Clin Microbiol Rev 19(3):449–490. https://doi.org/10.1128/CMR.00054-05

Malfertheiner P, Bazzoli F, Delchier JC, Celiński K, Giguère M, Rivière M, Mégraud F (2011) *Helicobacter pylori* eradication with a capsule containing bismuth subcitrate potassium, metronidazole, and tetracycline given with omeprazole versus clarithromycin-based triple therapy: a randomized, open-label, non-inferiority, phase 3 trial. Lancet 377(9769):905–913. https://doi.org/10.1016/S0140-6736(11)60020-2

Malfertheiner P, Megraud F, O'Morain CA, Gisbert JP, Kuipers EJ, Axon AT, Bazzoli F, Gasbarrini A, Atherton J, Graham DY, Hunt R, Moayyedi P, Rokkas T, Rugge M, Selgrad M, Suerbaum S, Sugano K, El-Omar EM (2017) Management of *Helicobacter pylori* infection—the maastricht v/florence consensus report. Gut 66(1):6–30

Malfertheiner P, Schultze V, Rosenkranz B, Kaufmann SH, Ulrichs T, Novicki D, Norelli F, Conttorni M, Peppoloni S, Berti D, Tornese D, Ganju J, Palla E, Rappuoli R, Scharschmidt BF, Del Giudice G (2008) Safety and immunogenicity of an intramuscular *Helicobacter pylori* vaccine in noninfected volunteers: a phase I study. Gastroenterology 135:787–795. https://doi.org/10.1053/j.gastro.2008.05.054

Marcus EA, Sachs G, Scott DR (2013) The role of ExbD in periplasmic pH homeostasis in *Helicobacter pylori*. Helicobacter 18(5):363–372. https://doi.org/10.1111/hel.12055

Mini R, Bernardini G, Salzano MA, Renzone G, Scaloni A, Figura N, Santucci A (2006) Comparative proteomics and immunoproteomics of *Helicobacter pylori* related to different gastric pathologies. J Chromatogr B 833:63–79. https://doi.org/10.1016/j.jchromb.2005.12.052

Morihara F, Fujii R, Hifumi E, Nishizono A, Uda T (2007) Effects of vaccination by a recombinant antigen ureB138 (a segment of the beta-subunit of urease) against *Helicobacter pylori* infection. J Med Microbiol 56(Pt 6):847–853. https://doi.org/10.1099/jmm.0.47061-0

Moyat M, Velin D. (2016). Use of VacA as a Vaccine Antigen. Toxins 8(6): 181. doi: 10.3390/toxins8060181.

Nandy A, Dey S, Roy P, Basak SC (2018) Epidemics and peptide vaccine response: a brief review. Curr Top Med Chem 18(26):2202–2208. https://doi.org/10.2174/1568026618666181112144745

Pride DT, Meinersmann RJ, Blaser MJ (2001) Allelic variation within *Helicobacter pylori* babA and babB. Infect Immun 69(2):1160–1171. https://doi.org/10.1128/IAI.69.2.1160-1171.2001

Prinz C, Hafsi N, Voland P (2003) *Helicobacter pylori* virulence factors and the host immune response: implications for therapeutic vaccination. Trends Microbiology 11(3):134–138. https://doi.org/10.1016/S0966-842X(03)00024-6

Purcell AW, McCluskey J, Rossjohn J (2007) More than one reason to rethink the use of peptides in vaccine design. Nat Rev Drug Discov 6(5):404–414. https://doi.org/10.1038/nrd2224

Roggen EL (2006) Recent developments with B-cell epitope identification for predictive studies. J Immunotox 3(3):137–149. https://doi.org/10.1080/15476910600845690

Sanchez-Trincado JL, Gomez-Perosanz M, Reche PE (2017) Fundamentals and methods for T- and B-cell epitope prediction. J Immunol 2017:2680160. https://doi.org/10.1155/2017/2680160

Skwarczynski M, Toth I (2016) Peptide-based synthetic vaccines. Chem Sci 7(2):842–854. https://doi.org/10.1039/c5sc03892h

Stein M, Ruggiero P, Rappuoli R, Bagnoli F (2013) *Helicobacter pylori* CagA: from pathogenic mechanisms to its use as an anti-cancer vaccine. Front Immunol 4:328. https://doi.org/10.3389/fimmu.2013.00328

Sterbenc A, Jarc E, Poliak M, Homan M (2019) *Helicobacter pylori* virulence genes. World J Gastroenterol

25(33):4870–4884. https://doi.org/10.3748/wjg.v25.i33.4870

Suerbaum S, Smith JM, Bapumia K, Morelli G, Smith NH, Kunstmann E, Dyrek I, Achtman M (1998) Free recombination within *Helicobacter pylori*. Proc Natl Acad Sci USA 95:12619–12624. https://doi.org/10.1073/pnas.95.21.12619

Takahashi-Kanemitsu A, Knight CT, Hatakeyama M (2020) Molecular anatomy and pathogenic actions of Helicobacter pylori CagA that underpin gastric carcinogenesis. Cell Mol Immunol 17(1):50–63. https://doi.org/10.1038/s41423-019-0339-5

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22(22):4673–4680. https://doi.org/10.1093/nar/22.22.4673

van der Ende A, Pan ZJ, Bart A, van der Hulst RW, Feller M, Xiao SD, Tytgat GNJ, Dankert J (1998) cagA-positive *Helicobacter pylori* populations in China and The Netherlands are distinct. Infect Immun 66:1822–1826

Vitoriano I, Rocha-Gonçalves A, Carvalho T, Oleastro M, Calado CRC, Roxo-Rosa M (2011) Antigenic diversity among Portuguese clinical isolates of *Helicobacter pylori*. Helicobacter 16(2):153–168. https://doi.org/10.1111/j.1523-5378.2011.00825.x

Voland P, Weeks DL, Marcus EA, Prinz C, Sachs G, Scott D (2003) Interactions among the seven *Helicobacter pylori* proteins encoded by the urease gene cluster. Am J Physiol Gastrointest Liver Physiol 284(1):G96–G106. https://doi.org/10.1152/ajpgi.00160.2002

Yan J, Liang S-H, Mao Y-F, Li L-W, Li S-PP (2003) Construction of expression systems for flaA and flaB genes of *Helicobacter pylori* and determination of immunoreactivity and antigenicity of recombinant proteins. World J Gastroenterol 9(10):2240–2250

Yang J, Dai L-x, Pan X, Wang H, Li B, Zhu J, Li M-y, Shi X-L, Wang B-n (2015) Protection against *Helicobacter pylori* infection in BALB/c mice by oral administration of multi-epitope vaccine of CTB-UreI-UreB. Pathog Dis 73(5):ftv026

Yang XL, Liu W-C, Yang W-W, Zhong D, Liu Y-H, Zhang J-D, Jiang J-H, Li S-S (2005) Oral immunization of mice with vaccine of attenuated *Salmonella typhimurium* expressing *Helicobacter pylori* urease B subunit. Biomed Environ Sci 18(6):411–418

Zhao J-W, Zhang S, Yang H-G, Zhang C-Z, Yu X-L, Wen Z-L, Gao Y-F, Guo X-K, Qi Y-M, Zhang S-L (2011) Cloning, expression and immunological evaluation of a short fragment from Rv3391 of Mycobacterium tuberculosis. Annals Microb 61(2):345–353. https://doi.org/10.1007/s13213-010-0148-7