



# A Novel Design of Multi-epitope Vaccine Against *Helicobacter pylori* by Immunoinformatics Approach

Junfei Ma<sup>1</sup> · Jingxuan Qiu<sup>1</sup> · Shuying Wang<sup>1</sup> · Qianyu Ji<sup>1</sup> · Dongpo Xu<sup>1</sup> · Haiwang Wang<sup>3</sup> · Zhiguang Wu<sup>3</sup> · Qing Liu<sup>1,2</sup>

Accepted: 25 November 2020 / Published online: 2 January 2021

© The Author(s), under exclusive licence to Springer Nature B.V. part of Springer Nature 2021

## Abstract

*Helicobacter pylori* (*H. pylori*) is a gram-negative spiral bacterium that caused infections in half of the world's population and had been identified as type I carcinogen by the World Health Organization. Compared with antibiotic treatment which could result in drug resistance, the vaccine therapy is becoming a promising immunotherapy option against *H. pylori*. Further, the multi-epitope vaccine could provoke a wider immune protection to control *H. pylori* infection. In this study, the in-silico immunogenicity calculations on 381 protein sequences of *H. pylori* were performed, and the immunogenicity of selected proteins with top-ranked score were tested. The B cell epitopes and T cell epitopes from three well performed proteins UreB, PLA1, and Omp6 were assembled into six constructs of multi-epitope vaccines with random orders. In order to select the optimal constructs, the stability of the vaccine structure and the exposure of B cell epitopes on the vaccine surface were evaluated based on structure prediction and solvent accessible surface area analysis. Finally Construct S1 was selected and molecular docking showed that it had the potential of binding TLR2, TLR4, and TLR9 to stimulate strong immune response. In particular, this study provides good suggestions for epitope assembly in the construction of multi-epitope vaccines and it may be helpful to control *H. pylori* infection in the future.

**Keywords** *Helicobacter pylori* · Multi-epitope vaccine · Reverse vaccinology · Structural vaccinology · Immunoinformatics

## Introduction

*Helicobacter pylori* infection has been shown to be associated with multiple gastric diseases such as chronic gastritis, peptic ulcers, gastric malignancies (Peek and Blaser

2002; Sigal et al. 2015; Suerbaum and Michetti 2002), and non-gastric diseases including iron deficiency anemia, idiopathic thrombocytopenic purpura, non-alcoholic fatty liver (Barabino 2002; Emilia et al. 2004; Okushin et al. 2015). Currently the *H. pylori* infection was usually treated by an antibiotic-based triple regimen although it causes a series of problems such as increased antimicrobial resistance and intestinal flora disturbance (Suzuki et al. 2019; Wang and Huang 2014). Therefore, therapeutic vaccination against *H. pylori* infection could be a more effective and safer immunotherapy avoiding the drawbacks of antibiotics.

Screening of antigen targets is key to vaccine development. In the previous studies, many antigens such as urease, neutrophil activating protein (NAP), and superoxide dismutase (SOD) have been proved to be the excellent candidates for their effective immune response against *H. pylori* (Corthesy-Theulaz et al. 1998; Every et al. 2011; Peng et al. 2018). In most instances, we must face that verified antigens have limited effect on clearance of *H. pylori* in the stomach (Sutton and Chionh 2013). So it is necessary to predict new candidate antigens through reverse vaccination for the ability of high-throughput screening. The strategies for predicting

---

Junfei Ma and Jingxuan Qiu have contributed to the work equally and should be regarded as co-first authors.

---

**Supplementary Information** The online version of this article (<https://doi.org/10.1007/s10989-020-10148-x>) contains supplementary material, which is available to authorized users.

---

✉ Qing Liu  
liuq@usst.edu.cn

<sup>1</sup> School of Medical Instrument and Food Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China

<sup>2</sup> Laboratory for Marine Fisheries Science and Food Production Processes, Qingdao National Laboratory for Marine Science and Technology, Qingdao 266071, China

<sup>3</sup> Beijing DQY Agricultural Technology Co., Ltd, Beijing 100094, China

antigens through reverse vaccination based on bioinformatics have been successfully applied to multiple pathogens including serogroup *B meningococcus*, *Klebsiella pneumoniae*, *Mycobacterium abscessus* (Giuliani et al. 2006; Lundberg et al. 2013; Shanmugham and Pan 2013). The reverse vaccination approach, which has become an alternative attractive way to identify all potential antigen targets, could accelerate the antigen discovery process and reduce antigen failure rate compared to traditional methods.

In addition to this, an emphasis on vaccine design and development has moved to the generation of recombinant multi-epitope vaccines. A multi-epitope vaccine consists of B and T epitopes from several different antigens in a certain order, with reasonable linkers which could avoid generation of junctional epitopes and promote the antigen presentation process (Livingston et al. 2002; Nezafat et al. 2017). As a new strategy relative to the single intact antigen vaccines, it has been shown to have benefits that the construct of multiple epitopes could activate a broader protective immune response and it could lead to an effective immunoreaction blocking multiple pathogenic channels for the control of *H. pylori* (Meza et al. 2017). Recent studies have reported that a variety of multi-epitope vaccines against *H. pylori* could induce comparatively high levels of specific antibodies against multiple antigens which are the sources of epitopes (Guo et al. 2017a, b; Pan et al. 2018).

To evaluate the immunogenicity of target protein, the property of epitopes which could be presented by immune cell to trigger immune response should be estimated. Epitopes could be divided into T cell epitopes and B cell epitopes, T cell epitopes were presented by MHC molecule and be recognized by T cell, while B cell epitopes were recognized by antibodies (Gold and Reth 2019; Neefjes et al. 2011). Based on the structure features, conformational epitopes were defined as the epitopes composed of segments, which were discontinuous in sequence but compact in spatial. And linear epitopes refers to the effective peptides in protein sequence (Van Regenmortel 2009; Zhang et al. 2014). In order to construct multi-epitope vaccines, linear epitopes including MHC-binding peptide and B cell linear epitopes were taken into consideration in this study.

Due to the string-of-beads structures of multi-epitope vaccines, diverse constructs may be made by changing the order of the epitope segments. The constructs with different orders make a difference in stability of the recombinant antigen, the exposure of the critical domains, and the docking with pathogen-associated molecular patterns (PAMP). Proper presentation of antigens, which can efficiently induce the immune system, is strongly dependent on the optimal structural stability of the vaccine construct (Scheibelhofer et al. 2017). In this approach, computer-aided design could invoke available structural information to engineering and designing immunogens (Dormitzer et al. 2012; Kulp and

Schief 2013). Being evaluated by predictive structure and other physicochemical features including hydrophilicity-hydrophobicity, solubility, and pI, the construct with optimal order will be selected and considered as the most promising candidate vaccine for further experimental verification (Negahdaripour et al. 2018).

In this paper, candidate antigens PLA1 and Omp6 were screened based on in-silico immunogenicity calculations and the experiment of antigen verification. Then multi-epitope vaccines were constructed using predicted epitopes of PLA1 and Omp6; the queried epitopes of UreB, which is the most successful antigen of *H. pylori* were also considered into the constructs. Subsequently, six different constructs of multi-epitope vaccines were evaluated based on their predicted structures. To select the best multi-epitope antigen construct, the stability of the construct and the exposure of the B cell epitopes on the vaccine surface were considered. Structure validation servers were applied for evaluating the stability of the construct and assessed the exposure of the B cell epitopes on the vaccine surface by solvent accessible surface area analysis. Construct S1 that performed well in above two aspects was selected. Molecular docking with TLR2, TLR4, and TLR9 showed that Construct S1 had the potential to bind TLRs. Finally, codon optimization and in-silico cloning of the construct sequence were used to increase its expression level in *E. coli* K12. In short, we propose a novel strategy for designing multi-epitope vaccine against *H. pylori* based on reverse vaccinology and structural vaccinology.

## Materials and Methods

### Bacteria and Culture

The mouse-adapted *H. pylori* strain SS1 and other 12 clinical isolates strains *H. pylori* A-L were obtained from Shanghai Institute of Digestive Disease and then preserved in our laboratory. *H. pylori* was cultured on Columbia plates containing 7% newborn bovine serum and *H. pylori* selective supplement SR0147E (Oxoid, England) under microaerophilic conditions at 37 °C for 3–4 days.

### Dataset

To construct the protein dataset of *H. pylori* for antigen screening, available sequences were collected from the following public protein databases. (1) Proteogen database (Yang et al. 2011) for protective antigens. (2) AntigenDB database (Ansari et al. 2010) for antigen proteins, which have been validated experimentally. (3) OMPdb database (Tsirigos et al. 2011) for  $\beta$ -barrel outer membrane proteins from Gram-negative bacteria. (4) PSORTdb database (Rey et al. 2005) for secreted proteins and outer membrane

proteins of *H. pylori*. There is a total number of 14,702 protein sequences for *H. pylori*. And the sequence redundancy was reduced using CD-HIT (Fu et al. 2012) by selecting the representative sequence for sequences with identity above 90%. In the end, a non-redundant protein dataset of 381 sequences were constructed for *H. pylori*.

## Epitope Prediction

### MHC-I Binding Peptide Prediction

To predict the peptides which could be presented by MHC-I molecule, NetMHCpan4.0 server (Jurtz et al. 2017) were used to screen each of 381 protein sequences. A total of 8 MHC subtypes were selected for mice, including H-2-Db, H-2-Dd, H-2-Kb, H-2-Kd, H-2-Kk, H-2-Ld, H-2-Qa1, H-2-Qa2. All of the 9-mer peptides in the input protein were screened by a sliding window, and the binding affinity between MHC molecule and peptides was ranked. The strong binding peptides were selected as the peptides with binding affinity ranked within top 0.5% and weak binding peptides refer to the peptides with binding affinity ranked within top 2%. In this model, a peptide was defined as the significant peptides  $S_j$ , if it could bind to more than half of H-2 subtypes.

For each input protein  $P$  in dataset, the MHC-I binding score  $M_I(P)$  was defined as formula (1)

$$M_I(P) = \frac{N(S_j)}{\text{len}(P)} \quad (1)$$

In which,  $N(S_j)$  refers to the number of significant peptides  $S_j$  in protein  $P$ . And  $\text{len}(P)$  represents the length of protein  $P$ .

### MHC-II Binding Peptide Prediction

For each of 381 protein sequences, the sequences of proteins were input to NetMHCIIpan 3.2 server (Jensen et al. 2018) to predict the peptides, which could bind to MHC-II molecule. The length of screened peptides was set as 15-mer. The peptides with predicted binding affinity with MHC-II molecule ranked in top 2% were defined as strong binding peptides. The weak binding peptides referred to the peptides with top 10% ranked binding affinity. All of the available peptides in each protein were screened and the binding ability between peptides and the MHC-II subtypes including H-2-IAb, H-2-IEb, H-2-IAd, and H-2-IEd were predicted. In this model, the significant peptides  $S_{II}$  was defined, if the peptides could bind to more than half of selected MHC-II subtypes.

For protein  $P$ , MHC-II binding peptide score  $M_{II}(P)$  was defined as formula (2).

$$M_{II}(P) = \frac{N(S_{II})}{\text{len}(P)} \quad (2)$$

where  $S_{II}$  represents significant MHC-II binding peptides included in protein  $P$ .  $N(S_{II})$  refers to the number of  $S_{II}$ . And  $\text{len}(P)$  represents the length of protein  $P$ .

### B Cell Linear Epitope Prediction

Apart from the above T cell epitope prediction, B cell linear epitopes peptides or epitope residues were also predicted by two tools including LBtope server (Singh et al. 2013) and BepiPred2.0 server (Jespersen et al. 2017). In LBtope server, all the available k-mer peptides split from the input protein by sliding window were screened. The probability score of each peptide to serve as a linear epitope was calculated. In this model, the k-mer was defined as the defaulting setting of 15-mer. The possible linear epitope peptides were defined with the cutoff of percent probability for correct prediction setting as 60%. For protein  $P$ , the LBtope score  $B_1(P)$  was calculated as:

$$B_1(P) = \frac{N(\text{epi\_pep})}{\text{len}(P)} \quad (3)$$

where the  $\text{epi\_pep}$  refers to the predicted linear peptide by LBtope for protein  $P$ .  $N(\text{epi\_pep})$  represents the number of predicted peptides. And  $\text{len}(P)$  represents the number of amino acids in protein  $P$ .

In the calculation of BepiPred1.0, the results were predicted whether each residue in the input protein could serve as epitope residues. Thus, for protein  $P$ , the score for BepiPred  $B_2(P)$  was defined as formula (4).

$$B_2(P) = \frac{N(\text{epi\_rsd})}{\text{len}(P)} \quad (4)$$

In which the  $N(\text{epi\_rsd})$  represented the number of epitope residues which was predicted by BepiPred for protein  $P$ . And  $\text{len}(P)$  represented the length of protein  $P$ .

### Epitope Mapping

To give a comprehensive evaluation of immunogenicity for each protein, the epitope mapping score was also designed in this model to reflect whether the protein contained the experimentally validated epitopes. First, a total of 39,879 experimental linear epitope recognized by mice were first collected from the immune epitope database and analysis resource (IEDB) (Vita et al. 2019). Then protein sequence of each 381 proteins from *H. pylori* was compared to the above epitopes. And the epitope mapping score  $EM(P)$  for protein  $P$  was set as 1, if the protein from *H. pylori* contained any of known epitopes, otherwise the  $EM(P)$  was set as 0.

## Immunogenicity Score

Before the calculation of final immunogenicity score for each protein, each of above scores should be normalized first. To eliminate the bias caused by different range of score. Each score was normalized in the range of 0 to 1, based on the maximum and minimum values. For the protein  $P$ , the MHC-I binding score  $M_I(P)$  was normalized as  $NM_I(P)$  as following:

$$NM_I(P) = \frac{M_I(P) - \min(M_I)}{\max(M_I) - \min(M_I)} \quad (5)$$

where the  $\max(M_I)$  referred to the maximum value of  $M_I$  for all the 381 protein sequences, and the  $\min(M_I)$  referred to the minimum value of  $M_I$  for all 381 *H. Pylori* proteins.

The final immunogenicity score  $I(P)$  were defined as the sum of above normalized scores as formula (6).

$$I(P) = NM_I(P) + NM_{II}(P) + NB_1(P) + NB_2(P) + EM(P) \quad (6)$$

where the  $NM_{II}(P)$ ,  $NB_1(P)$ ,  $NB_2(P)$  refer to the normalized score of  $M_{II}(P)$ ,  $B_1(P)$  and  $B_2(P)$  respectively.

The immunogenicity score for all 381 protein sequences from *H. pylori* were calculated respectively, and the immunogenicity scores were ranked in descending order to select the potential antigens for *H. pylori*.

## Gene Conservation Analysis of Screened Antigens

The genomic DNA as PCR template from 13 *H. pylori* strains was extracted. Then 6 antigen genes we screening were amplified using primers shown in Table 1. Agarose gel electrophoresis was used to identify them. Then three widely distributed gene *ureB*, *HP0499*, and *HP0229* were cloned into T-vector and send to the Beijing Genomics Institute for sequencing after PCR and purification. According to the sequencing report, the phylogenetic tree was made based on gene sequences by MEGA6.

## Cloning of UreB, PLA1 and Omp6

According to the phylogenetic tree analysis, *ureB* (*H. pylori* G), *HP0499* (*H. pylori* L) and *HP0229* (*H. pylori* E) were chosen as the final sequence to be cloned into the plasmid pET30a vector. Super-fidelity PCR amplification was used to obtain accurate sequences with Phanta Max Super-Fidelity DNA Polymerase (Vazyme, China) and ClonExpress II One Step Cloning Kit (Vazyme, China) to clone the target genes into the plasmids. The recombinant plasmids were transformed into *E. coli* DH5 $\alpha$  competent cells and screened on LB plates with kanamycin (50  $\mu$ g/mL). Subsequently, the inserts were confirmed via PCR and sequencing.

## Expression and Purification of UreB, PLA1 and Omp6

*E. coli* BL21(DE3) competent cells were transformed with pET30a-UreB/PLA1/Omp6 and inoculated in LB broth with kanamycin (50  $\mu$ g/mL). All three expression was induced by addition of Isopropyl  $\beta$ -D-1-Thiogalactopyranoside (IPTG) after the optical density (OD600) had reached 0.6. The gradient of IPTG concentration and induced temperature were explored. After 6 h the whole proteins were extracted from the collected cells by One Step Bacteria Active Protein Extraction Kit (Sangon Biotech, China) and analyzed by Sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE). The soluble histidin-tagged proteins were purified using affinity chromatography (Ni-IDA-Sefinose Column, Sangon Biotech, China).

## Western Blot

Purified recombinant proteins were evaluated by western blot probed with a rabbit anti-*H. pylori* antibody (GeneTex, USA). Briefly, purified UreB, PLA1 and Omp6 were applied to 12% SDS-PAGE under denaturing conditions and electro transferred on to a nitrocellulose membrane at 100 V for 1.5 h. Nonspecific binding sites were blocked overnight at 4 °C in blocking buffer (5% non-fat milk in TBST, pH 7.4, with 0.05% Tween-20). Membranes were washed three times for 10 min with TBST and further incubated with a 1:5,000 dilution of rabbit polyclonal anti-*H. pylori* antibody at 37 °C for 2 h. After the same washing way, PBST, the membranes

**Table 1** Primer sequences of antigen gene amplification

Name	Forward primer (5' → 3')	Reverse primer (5' → 3')
<i>hopQ</i>	ATGAAAAAACGAAAAAAC	TTAATAAGCGAACACATAAT
<i>HP0175</i>	ATGAAAAAATATCTTAA	TTACTTGTGATAACAATTT
<i>HP1156</i>	TTGCAAACTTTGTTTTAG	TCAAAGCCTATGTTGTAGC
<i>ureB</i>	ATGAAAAGATTAGCAGAA	CTAGAAAATGCTAAAGAGTT
<i>HP0499</i>	ATGAAAAGCATTTTGCTCTT	TTAAGGGTTTAAGCGTATTC
<i>HP0229</i>	ATGAAAAAACGATTTTACT	TCAATAGCCAAAACATAAT



were incubated in a 1:10,000 dilution of Alexa Fluor Plus 800-conjugated goat anti-rabbit IgG (Invitrogen, USA) at 37 °C for 2 h. The protein bands were visualized with Odyssey Imager Dual-color infrared imaging (Li-Cor, USA).

### Epitopes Acquisition of UreB, PLA1 and Omp6

The fragments containing CD4<sup>+</sup> T cell epitopes and B cell epitopes in PLA1 and Omp6 were predicted by T and B cell epitope prediction tools. The B-cell epitopes in PLA1 and Omp6 were predicted by BepiPred 2.0 using the epitope threshold at 0.5. The T cell epitopes in PLA1 and Omp6 were predicted by NetMHCIIpan 3.2 Server. Mouse H-2 IAd and H-2 IEed were selected as the allele type due to the fact that the design was for BALB/C mice and other parameters were set as the default. All the real B and T cell epitopes in UreB was obtained from IEDB.

### Design and Structure Prediction of Multi-epitope Vaccines

Twenty constructs of tandem copies including 7 B cell epitopes, 6 T cell epitopes and the linkers “KK”, “GGGS” were determined randomly. The selected linkers could effectively connect the epitopes without changing the immunogenicity of designed epitopes and performed well in the improvements of protein stability, domain-domain orientation, and solubility (Kavoosi et al. 2007; Klement et al. 2015; Nezafat et al. 2017). After different attempts based on elementary evaluation of VaxiJen 2.0 (Doytchinova and Flower 2007), 6 constructs with higher VaxiJen scores numbered S1-S6 were obtained. Subsequently, the I-TASSER Server predicted the 3D structure of every construct sequence based on the degree of similarity between target protein and available template structure from PDB.

### Structure Validation of the Selected Multi-epitope Vaccine Construct

ProSA-web, RAMPAGE, and ERRAT sever were used for evaluating six multi-epitope vaccine constructs based on their predicted structures. The Z-score calculated by ProSA-web could assess the overall quality of predicted structure (Wiederstein and Sippl 2007). The main Ramachandran plot from RAMPAGE was used for calculating phi–psi torsion angles for each amino acid in the vaccine structure (Lovell et al. 2003). Finally, ERRAT sever was used for distinguishing between correctly and incorrectly determined structure regions based on the statistics of highly refined structures (Colovos and Yeates 1993).

### Solvent Accessible Surface Area (SASA) Analysis of B Cell Epitopes in Multi-epitope Constructs

Based on the 3D structures of multi-epitope constructs, the SASA of B cell epitopes in them was calculated using Tcl command of VMD software. Besides, the SASA of B cell epitope in every antigen was shown in their structures.

### Molecular Docking with Toll-Like Receptors (TLRs)

Based on the 3D structures of multi-epitope antigens, the molecular docking of the selected multi-epitope construct with TLR2, TLR4, and TLR9 was taken using ZDOCK server (Pierce et al. 2014) and the one with best docking of top 10 complexes was chosen as the next analysis object. Then the docking interface area and binding free energy ( $\Delta^iG$ ) was reported by PDBePISA server (Krissinel and Henrick 2007) after the complex structures had been submitted.

### Evaluation of In-Silico Cloning of Vaccine Constructs

In order to achieve maximum expression in *E. coli* system, Jcat (Grote et al. 2005) was used for reverse translation and optimization of codons. Jcat is a rapid and easy method that could ensure maximum expression by calculating the GC content and CAI value (Grote et al. 2005). Restriction sites of BamHI and KpnI were added to the reverse translated sequence. The optimized sequence was then cloned into pET30a plasmid using Snapgene software.

## Results

### Selection of Candidate Antigen Protein by Immunogenicity Calculation

In this study, an in-silico model was constructed to screen the antigen protein as vaccine target among all the available proteins of *H. pylori*. An effective antigen protein should meet the requirements including the following. (1) The candidate antigen protein should be able to trigger the immune response in hosts. (2) The vaccine targets should be accessible to the immune effectors. Based on the above hypothesis, the immunogenicity scoring model was constructed to evaluate the immunogenicity of selected proteins for *H. pylori*.

At first, 14,702 proteins including outer membrane proteins and secreted proteins of *H. Pylori* were collected from public databases, and a none-redundant dataset of 381 protein sequences was constructed for antigen screening. Then for each protein, the immunogenicity score was calculated which reflect the ability to serve as a potential antigen. The epitope prediction score including MHC-I binding

prediction, MHC-II binding prediction along with B cell linear epitope prediction was calculated, which refers to the ability of the protein to be recognized by MHC molecule or antibodies. Based on the descending order of immunogenicity score among 381 protein sequences, among the top 10% ranked proteins, six proteins were randomly selected for further validation, including protein HopQ (gene: *hopQ*), Omp25 (gene: *HP1156*), peptidylprolyl isomerase (gene: *HP0175*), Omp6 (gene: *HP0229*), phospholipase A1 (PLA1, gene: *HP0499*), and the widely used UreB (gene: *ureB*).

### Distribution of Six Antigen Genes in Thirteen *H. pylori* Strains

According to the results of agarose gel electrophoresis which were shown in Figure S1, three genes were identified in a small number of *H. pylori* strains, including *hopQ*, *HP1156*, *HP0175*, while the others were all identified in thirteen *H. pylori* strains, including *ureB*, *HP0499*, *HP0229*. Correspondingly, only three antigens UreB, PLA1, and Omp6 were evaluated next due to their wide distribution in all thirteen *H. pylori* strains.

### Phylogenetic Tree Analysis of Gene *ureB*, *HP0499* and *HP0229*

Figure 1 shows the phylogenetic relationships of gene *ureB* (Fig. 1a), *HP0499* (Fig. 1b) and *HP0229* (Fig. 1c) from *H. pylori* SS1 and 12 clinical isolates *H. pylori* A–L. For example, the *ureB* sequence of *H. pylori* G has the highest conservation in 13 *H. pylori* strains according to the phylogenetic tree analysis. Compared with *H. pylori* G, the *ureB* sequence similarity of the other 12 *H. pylori* strains could reach more than 96.8%. This indicates that *ureB* gene is highly conserved among 13 *H. pylori* strains and could be used as a broad-spectrum antigen. The same analysis as *ureB*, the *HP0499* sequence of *H. pylori* L and the *HP0229* sequence of *H. pylori* E are highly conserved. The sequence similarity can reach 96.2% (*HP0499*) and 90.7% (*HP0229*) respectively, which also proves that they could be regarded as broad-spectrum antigens.

### Expression, Purification, and Antigen Characteristics of UreB, PLA1 and Omp6

The UreB protein (about 66 kD) was abundantly expressed in inclusion body of *E. coli* BL21(DE3)/pET30a-UreB induced with 0.2 mM IPTG at 37 °C. After purification by Ni<sup>2+</sup>-NTA affinity chromatography, the purity of the UreB protein was shown as analyzed by SDS-PAGE (Fig. 2a; Lane 1). Similarly, the PLA1 (about 42 kD) and Omp6 (about 53 kD) proteins were also mainly expressed in inclusion body of *E. coli* BL21(DE3)/pET30a-PLA1 and pET30a-Omp6

induced with 0.5 mM IPTG at 30°C. The purified PLA1 and Omp6 proteins were shown in Fig. 2a (lanes 2 and 3). The immunoreactivity of UreB, PLA1, and Omp6 was identified by Western blotting. UreB, PLA1 and Omp6 proteins could be recognized by Rabbit anti-*H. pylori* polyclonal antibody (Fig. 2b), which elucidated that they could be used as candidate antigens against *H. pylori*.

### Epitopes Acquisition of UreB, PLA1 and Omp6

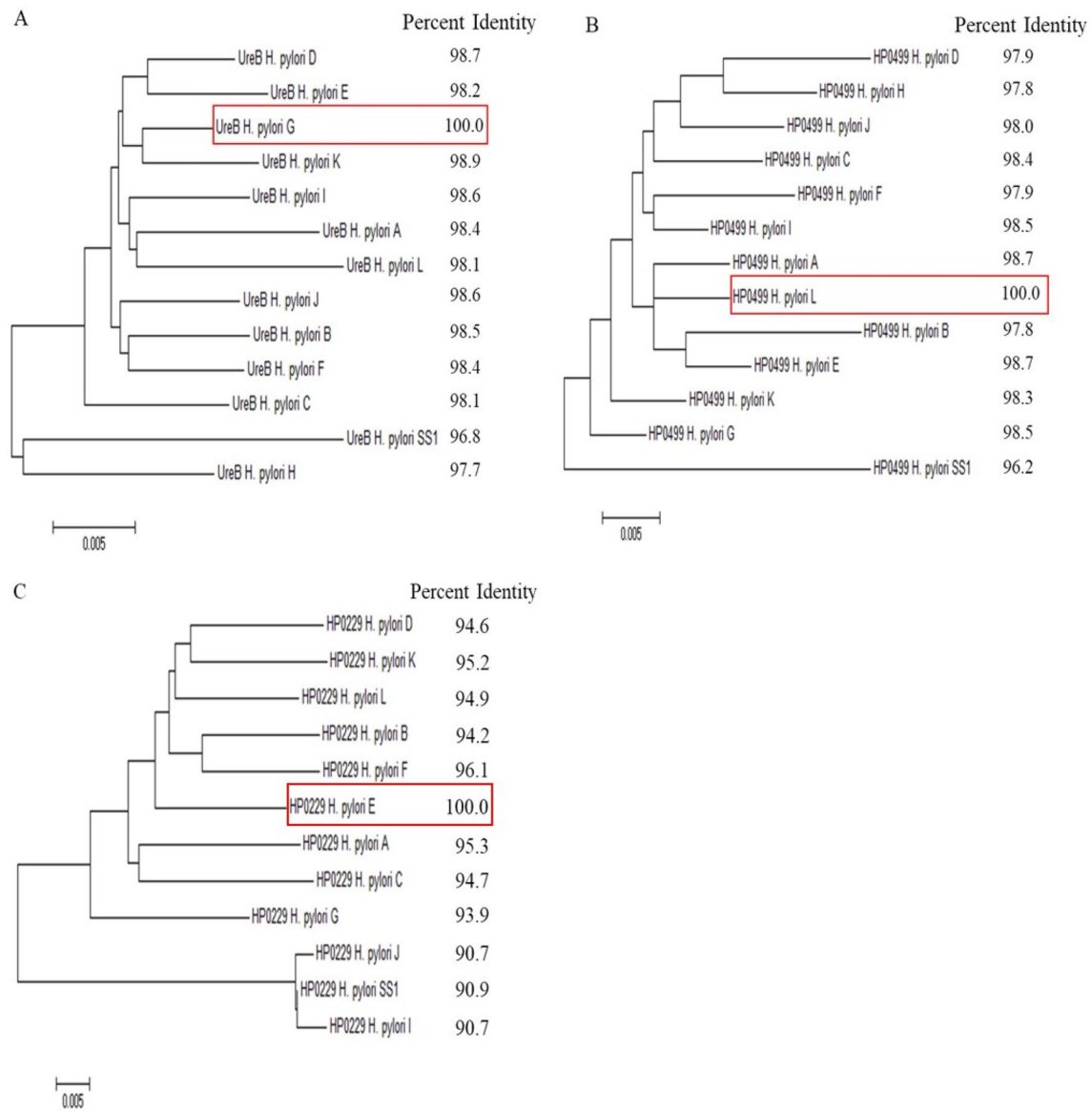
B cell epitope-inducing humoral immune response and CD4<sup>+</sup> T cell epitope-inducing cellular immune response must be considered in the treatment of *H. pylori* infection (Wilson and Crabtree 2007). The experimentally verified B and T epitopes of UreB could be queried via IEDB which is a freely available resource cataloging experimental data on antibody and T cell epitopes. Besides, the B and T cell epitopes with high scores were picked based on the predictions of BepiPred and NetMHCIIpan server. All the B cell epitopes were shown in Table 2 and all the T cell epitopes were shown in Table 3.

### Design and Structure Prediction of Multi-epitope Vaccines

Six constructs were produced by seven B cell epitopes and six T cell epitopes with random orders that showed in Fig. 3 and their sequences were shown in Table S1. Then their structures were predicted by I-TASSER Server which provides C-score for assessing model quality. C-score (Yang et al. 2015) in Table 4 is typically in the range of [− 5, 2], where a C-score of a higher value signifies a model with a higher confidence and vice-versa. All six structures of Constructs S1–S6 were shown in Fig. 4.

### Structure Validations of the Six Multi-epitope Vaccine Constructs

In order to further evaluate the structure reliability of six constructs, ProSA-web, RAMPAGE, and ERRAT servers were used for assessing the quality of the predicted structure. Taking Construct S1 as an example, the Z-score of the structure calculated by ProSA-web server was -2.39, which is in the range of native protein conformation scores (Fig. 5a). Ramachandran plot from RAMPAGE showed that 74.9% of residues were in favored regions, 15.8% of residues were in allowed regions, 9.3% of residues were in outlier regions (Fig. 5b). ERRAT results showed that the overall quality factor of Construct S1 was 86.250 (Fig. 5c). All these results indicated that the predicted structure of Construct S1 was reasonable and reliable. As shown in Table 4, the structure validation results of the other constructs were given (figures not shown). The structure validations of Constructs S1, S2,



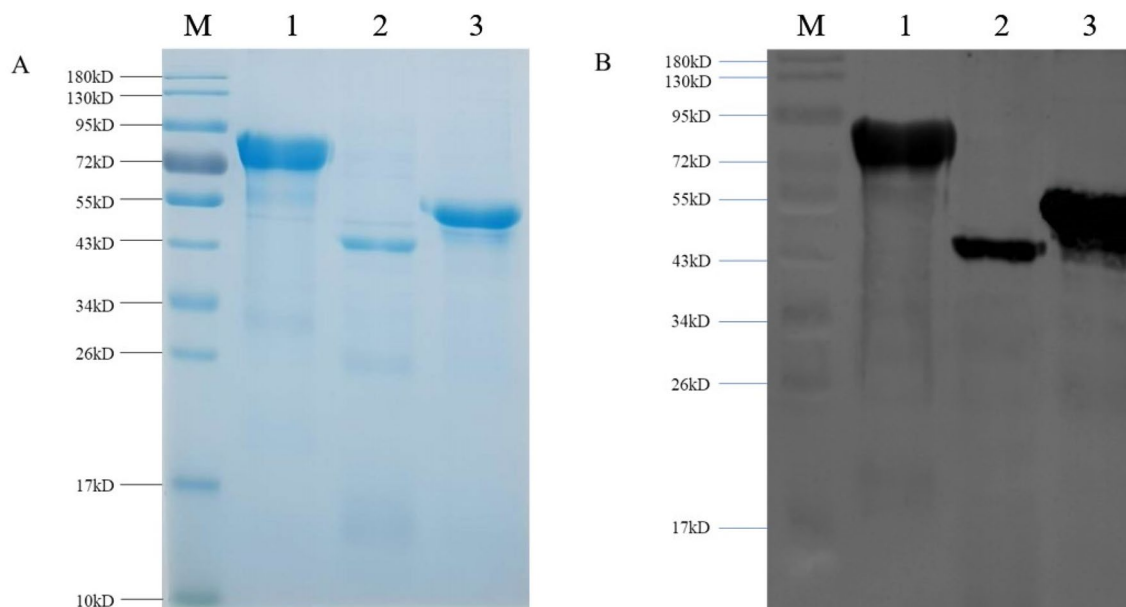
**Fig. 1** The phylogenetic tree of gene *ureB* (a), HP0499 (b) and HP0229 (c) from *H. pylori* SS1 and 12 clinical isolates *H. pylori* A-L

S3, and S6 by ProSA-web, RAMPAGE, and ERRAT servers were reasonable and they were consistent with C-score evaluation from I-TASSER. This also illustrated that the structures of these four constructs are more stable and reasonable. Nevertheless the structure validations of Constructs S4 and S5 were poor, which means that their two predicted structures were unreasonable and unstable.

### Solvent Accessible Surface Area (SASA) Analysis of B Cell Epitopes in Multi-epitope Constructs

Only when the B cell epitopes are distributed on the surface of the antigen can the host immune system recognize them and generate specific antibodies against them. In the

different multi-epitope antigen constructs we constructed, the constructs with highly exposed B cell epitopes are more likely to provoke specific antibody responses as we designed. SASA analysis of B cell epitopes was performed in six constructs. The B cell epitopes SASA analysis displays of Constructs S1–S6 were shown in Fig. 6 and the SASA value were shown in Table 4. Obviously, the SASA of Construct S1 is 9600.67 Å<sup>2</sup>, the largest of all constructs based on its high structure quality. So Construct S1 could be considered as the best choice. Further, the binding affinity of T cell epitopes in Construct S1 was verified again using NetMHCIIpan. The selected 6 T cell epitopes in their original antigens were all belong to strong binding levels. After being constructed, 4 T cell epitopes in Construct S1 were



**Fig. 2** Expression, purification, and antigen characteristics of UreB, PLA1 and Omp6. **a** Visualization of the purified UreB, PLA1, and Omp6. The proteins purified from *E. coli* BL21(DE3)/pET30a-UreB, pET30a-PLA1, and pET30a-Omp6 were resolved in 12% SDS-PAGE gel and stained with Coomassie Blue. (M) Maker; (Lanes 1, 2, and 3)

the purified proteins of UreB, PLA1, and Omp6. **b** Immunoreactivity of UreB, PLA1 and Omp6 analyzed by Western blot. (M) Maker; UreB (Lane 1), PLA1 (Lane 2), and Omp6 (Lane 3) resolved in 12% SDS-PAGE gels were probed with Rabbit anti-*H. pylori* polyclonal antibody

**Table 2** B-cell epitopes of UreB, PLA1, and Omp6 from servers and databases

Antigen	Number	Position	Sequence	Source	Score <sup>a</sup>
UreB	B1	211–225	IEAGAIGFKIHEDWG	IEDB	–
	B2	321–339	CHHLDKSIKEDVQFADSRI	IEDB	–
	B3	349–363	TLHDMGIFISITSSDS	IEDB	–
PLA1	B4	52–63	YLPPQKAPKEHS	BepiPred	1.30
	B5	189–213	NGVGGAAQCYQPFNKEGNPENQFPGQ	BepiPred	1.19
Omp6	B6	171–188	CPGIDGGNGKPWGINASG	BepiPred	1.32
	B7	209–231	KKAAADARRTAPESPNQPSAFNN	BepiPred	1.49

<sup>a</sup>It is the average of all the amino acid BepiPred scores in the epitope sequence

**Table 3** T-cell epitopes of UreB, PLA1, and Omp6 from servers and databases

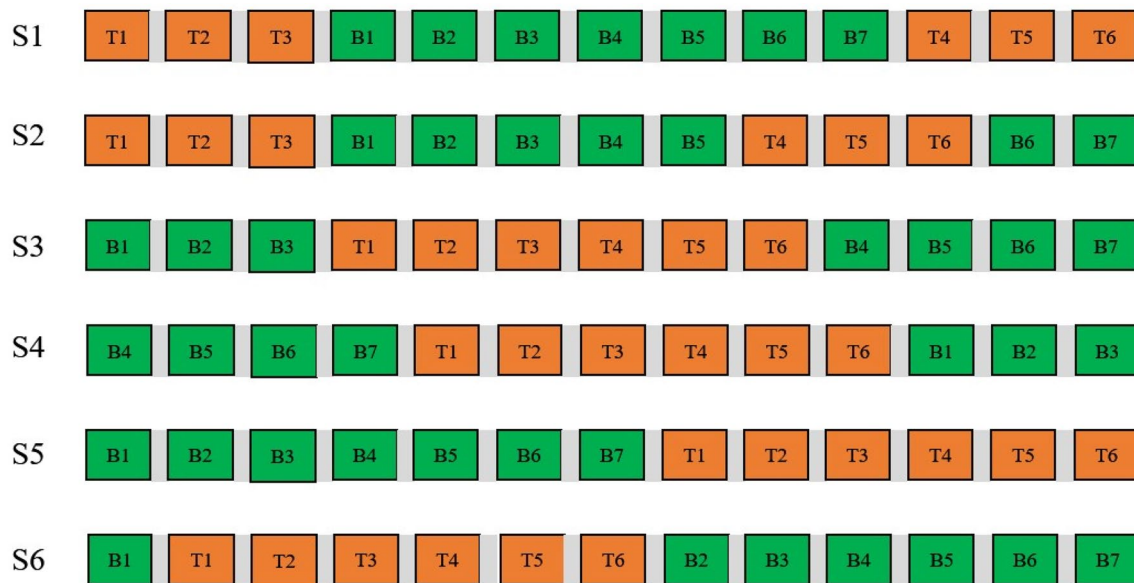
Antigen	Number	Type	Position	Sequence	Source	%Rank
UreB	T1	H-2-IAd	229–244	SAINHALDVADKYDVQ	IEDB	–
	T2	H-2-IEd	409–421	YTINPAIAHGISE	IEDB	–
	T3	H-2-IEd	548–561	DGKEVTSKPANKVS	IEDB	–
PLA1	T4	H-2-IEd	105–119	RNEFKFQISFRVPVF	NetMHCIIpan	0.05
Omp6	T5	H-2-IAd	93–107	SANQAVSSALSSAVA	NetMHCIIpan	0.9
	T6	H-2-IEd	428–442	NFQFLFNTGLRLQGI	NetMHCIIpan	0.25

still belong to strong binding levels and the other two turned into weak binding levels (Table S2), which showed that the epitope assembly method of Construct S1 was feasible.

### Molecular Docking with Toll-Like Receptors (TLRs)

The innate immune response principally depends on the pattern recognition receptors (PRRs) that recognizing the pathogen-associated molecular patterns (PAMP) of the pathogen (Chang et al. 2004). Toll-like receptors are the





**Fig. 3** The design and construction of multi-epitope vaccines with six random permutations. B cell epitopes filled with Color Green and T cell epitopes filled with Color Orange

**Table 4** The 3D structure evaluation and B cell epitopes exposure calculation of six constructs

Construct	VaxiJen score	C-score	Structure validation			SASA of B cell epitopes (Å <sup>2</sup> )
			Z-score	Residues in favoured regions (%)	Overall quality factor	
S1	1.1150	− 0.20	− 2.39	74.9	86.250	9600.67
S2	1.0974	0.23	− 3.09	73.7	87.967	7616.18
S3	1.1189	− 0.40	− 2.56	71.7	83.830	9393.13
S4	1.1261	− 0.73	− 0.71	63.6	67.089	8181.05
S5	1.1247	− 2.76	− 0.61	66.4	77.729	9037.90
S6	1.1250	− 0.23	− 2.25	68.0	89.540	8879.76

most studied PRRs, which could identify invading microorganisms to activate immune response (Netea et al. 2004). According to previous studies, *H. pylori* elicits the expression of pro-inflammatory genes via different receptors (TLR2, TLR4, and TLR9) (Chang et al. 2004; Ding et al. 2005; Uno et al. 2007). For the evaluation of the interaction between the vaccine Construct S1 and mouse TLR2, TLR4, and TLR9, ZDOCK server was employed. Their complex structures were given in Fig. 7 and the specific data of docking was shown in Table 5. Under the premise of reliable  $\Delta^iG$  P-value ( $P < 0.5$ ), the three docking interface surfaces are all interaction-specific based on the negative  $\Delta^iG$ .

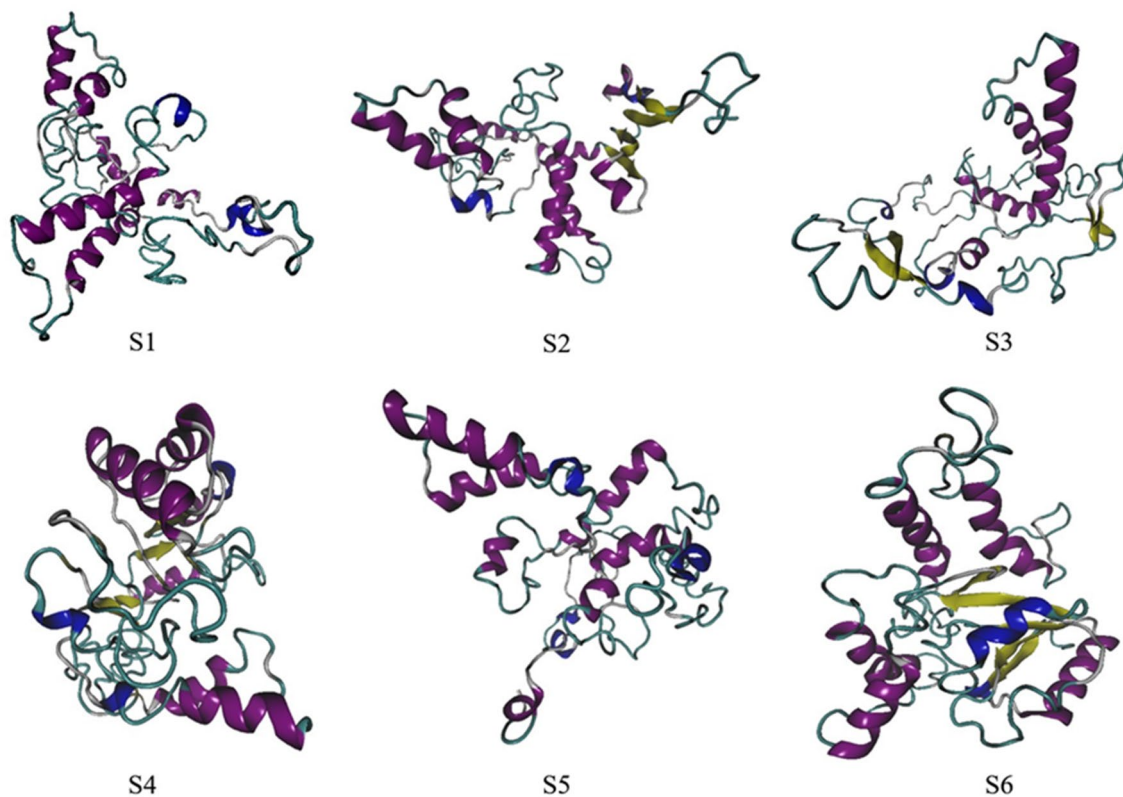
### Evaluation of In-Silico Cloning of Vaccine Constructs

For this study, Jcat was used to improve the expression of the multi-epitope vaccine in *E. coli* K12. A total of 249 amino acids were submitted. After codon optimization, the Codon

Adaptation Index (CAI) could reach 1.0 with GC contents of 48.46%, which is close to the GC contents of *E. coli* K12 (50.73%). This indicated the final multi-epitope vaccine with optimized sequence will be largely expressed in *E. coli* K12. The in-silico cloning map of multi-epitope vaccine construct was shown in Fig. 8.

### Discussion

To avoid the problems in antibiotic regimens against *H. pylori* infection, including relapses, increased resistance, and flora disturbances, vaccines could be a better alternative for its safety and effectiveness. Compared with the low effectiveness and high costs in terms of budget and time about conventional vaccine approaches, reverse vaccinology has become a new strategy for subunit vaccines development. Based on the information about the genome and proteome of



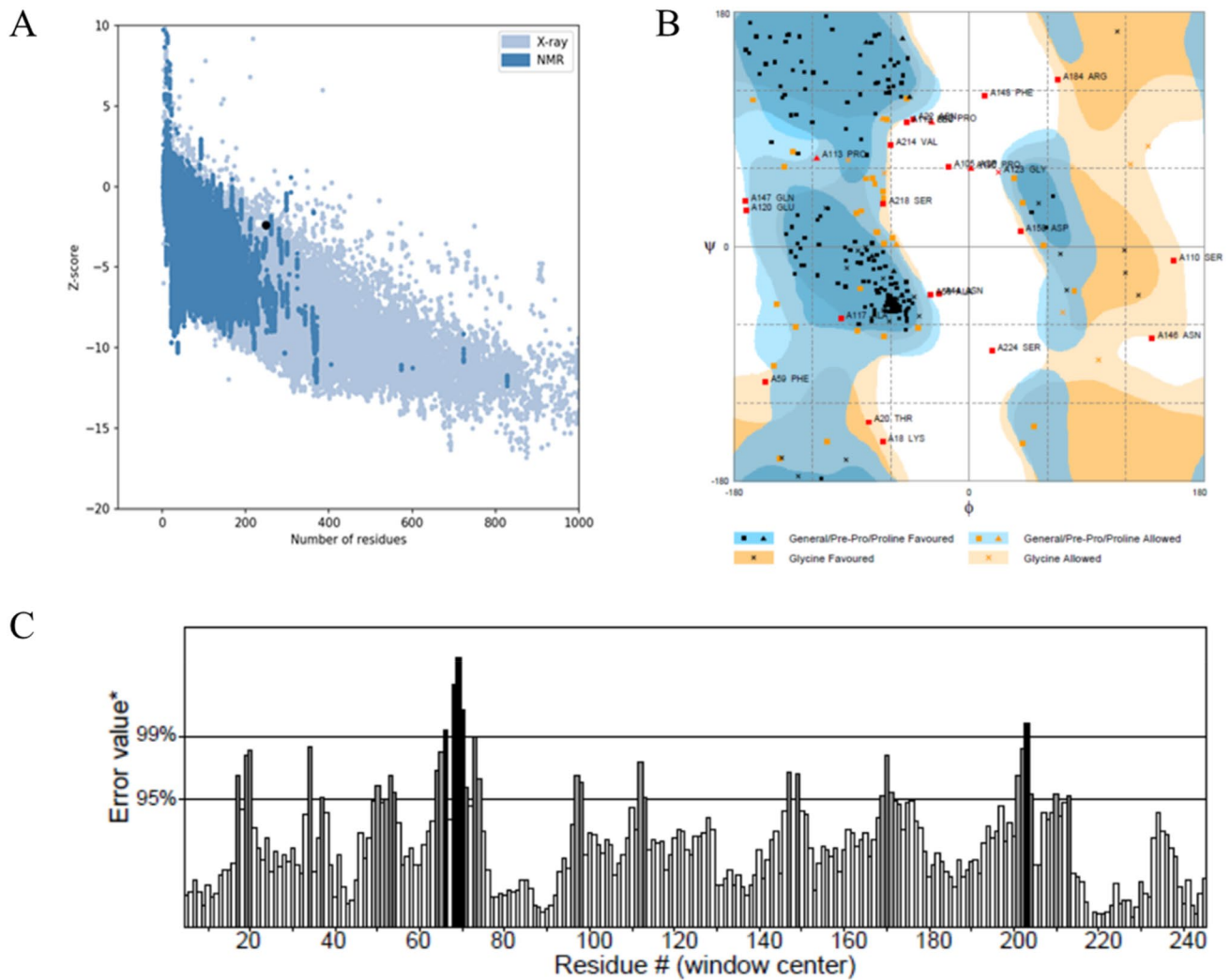
**Fig. 4** The 3D structures S1–S6 predicted by I-TASSER

microbes, the screening and evaluation of vaccine antigens have become more practically possible.

In this study, 381 protein sequences were screened from 14,702 protein sequences including outer membrane proteins and secreted proteins of *H. pylori* based on the immunogenicity score. MHC-I binding prediction, MHC-II binding prediction along with B cell linear epitope prediction were comprehensively considered in the evaluation of protein epitopes. PLA1, Omp6 and UreB were selected by virtue of the immunogenicity score and distribution in clinical strains. Besides, other untested proteins may have some potential as candidate antigens, since the six screened antigens were selected at random.

The development of subunit vaccine is no longer confined to a single antigen. Multi-epitope vaccine, which could stimulate a broader neutralizing antibody response and block multiple pathogenic channels, shows more effective results than a single antigen. Based on the new antigens PLA1, Omp6, and widely used antigen UreB, their epitopes were

combined into several constructs with different orders. The constructs, although same in amino acid composition, show the diverse antigen stability along with different position of key domain. In this study, 3D structure reliability and B cell epitope SASA analysis were used for evaluating the stability of the vaccine construct and exposure of B cell epitopes. Compared with homology modeling server, I-TASSER Server reflects better superiority in structure prediction of multi-epitope vaccines for the fragmented composition. The scores provided by I-TASSER and further structure validation by ProSA-web, RAMPAGE, and ERRAT sever could reflect the rationality of protein folding and the certainty of protein expression to some extent. Besides, based on the preliminary screening of VaxiJen Server, six constructs with high scores were performed for next structural prediction because of the time consumption of I-TASSER Server. In fact, a reasonable number of structures could be predicted to test based on the needs and resources until an accepted multi-epitope vaccine construct was obtained.

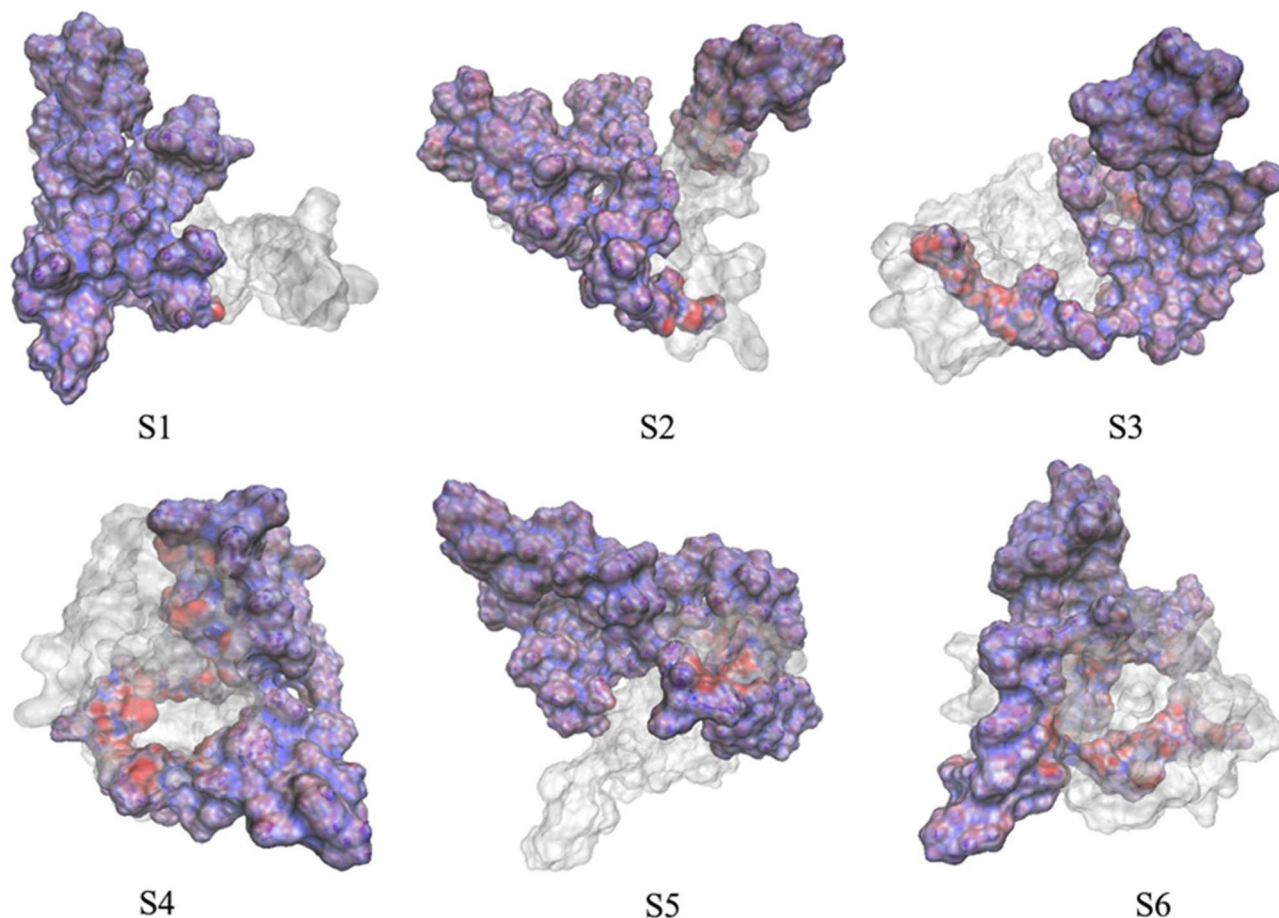


**Fig. 5** Structure validation of Construct S1. **a** The Z-score of Construct S1 calculated by ProSA web server was  $-2.39$ , which is in a reasonable range. **b** Ramachandran plot validation of Construct S1. Residues are in favoured regions 74.9%; residues are in allowed regions 15.8%; residues are in outlier regions 9.3%. **c** ERRAT analy-

sis of Construct S1. The overall quality factor of Combination N1 was 86.250. On the error axis, two lines of 95% and 99% are drawn to indicate the confidence with which it is possible to reject regions that exceed that error value

Cholera Toxin Subunit B (CTB), a widely used mucosal adjuvant in the construction of multi-epitope vaccine against *H. pylori*, enhanced the immunogenicity and antigenicity of the vaccine constructs in actual animal experiments (Guo et al. 2017b; Pan et al. 2018). It is doubtful that CTB or other adjuvant was added into the structure prediction of

multi-epitope vaccines in the design phase in some studies (Hugo Urrutia-Baca et al. 2019; Nosrati et al. 2019). The combination of CTB and multi-epitope antigen is a fusion protein rather than a new independent protein. Since only one parsed protein structure could be used as a template, the structure prediction platforms based on parsed



**Fig. 6** B cell epitopes SASA analysis displays of Constructs S1–S6. Exposed area of B cell epitopes filled with color blue, non-exposed one filled with color red. T cell epitopes filled with color gray

structures were not suitable for predicting the structures of fusion proteins. On the one hand, the structure of CTB has great changes in the predicted structure of fusion protein of CTB and multi-epitope antigen, which may result in loss of adjuvant effect. On the other hand, the addition of CTB may result in some changes in the terminal structure of the antigen, which affects the accuracy of the prediction to a certain extent. It may be more appropriate to add adjuvants after completing the design of multi-epitope vaccines.

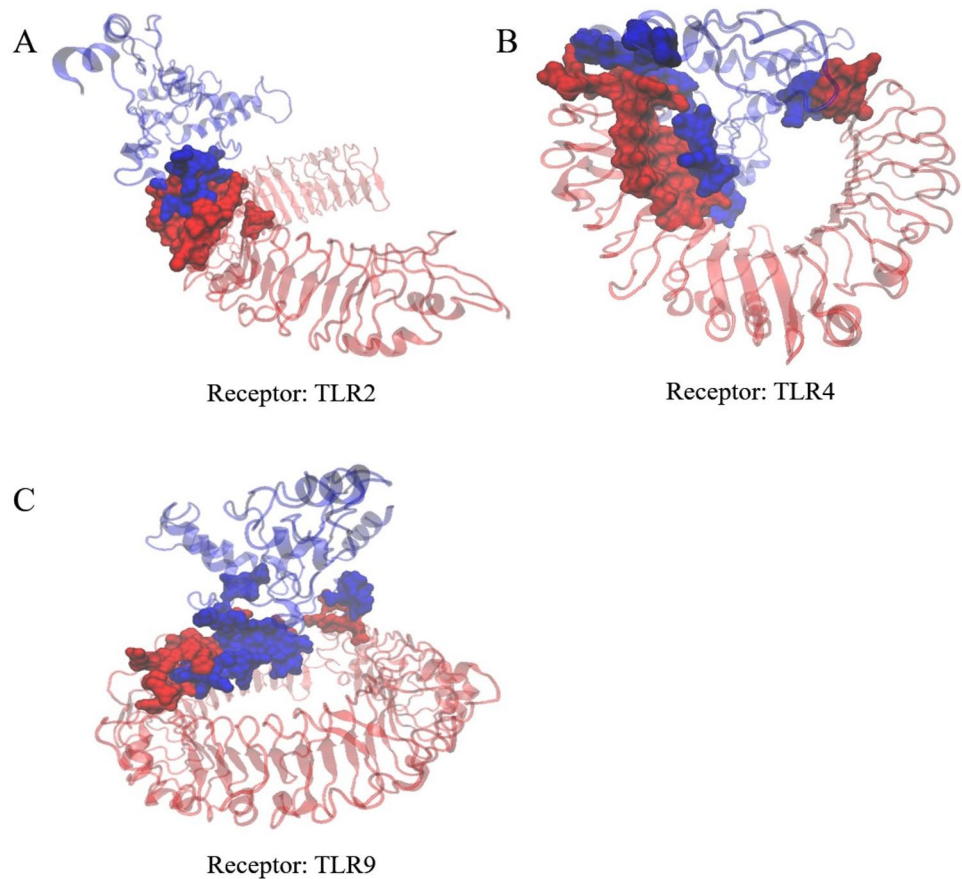
In the construction of multi-epitope vaccines, the optimal result is that the inserted B cell epitope could be located on the surface of the antigen protein. But in the real case, some amino acids of the B cell epitopes could be folded inside the protein. To some extent, the SASA analysis of B cell epitope could reflect the degree of B cell epitope exposure. More exposed ones could be selected according to the SASA calculation. Especially this is the first time that SASA analysis

was used for assessing the exposure of the B cell epitopes on the vaccine surface in multi-epitope vaccine construction. Nevertheless, the exposure to water, which is the analysis principle of SASA, could not represent the exposure to antibody. So there is a certain gap between this analysis method and the real situation. It still needs to be perfected.

To evaluate interactions among TLRs (TLR2, TLR4, and TLR9) and the vaccine structure, ZDOCK Server were applied for molecular docking. Docking results demonstrated that the multi-epitope vaccine we designed has the potential to bind TLRs. But the more accurate interactions need further validation of molecular dynamics simulations. After the codon optimization by Jcat Server, the designed vaccine could enhance expression in *E. coli* K12. Also note that further animal experimental verification is required to check if the vaccine we designed can trigger a specific immune response for controlling *H. pylori* infection.



**Fig. 7** The complex structures of Construct S1 and TLRs. The structure of Construct S1 filled with color blue and the structures of TLRs filled with color red. Amino acids at the interface are displayed with “surf” drawing method and the other amino acids are displayed with “Newcartoon” drawing method in VMD



**Table 5** The interface list of docking complex structures

Receptor	Interface area <sup>a</sup> (Å <sup>2</sup> )	$\Delta^iG^b$ (kcal/M)	$\Delta^iG$ P-value <sup>c</sup>
TLR2	1406.7	− 20.7	0.065
TLR4	1824.1	− 16.2	0.412
TLR9	1598.3	− 15.1	0.303

<sup>a</sup>Interface area calculated as difference in total accessible surface areas of isolated and interfacing structures divided by two

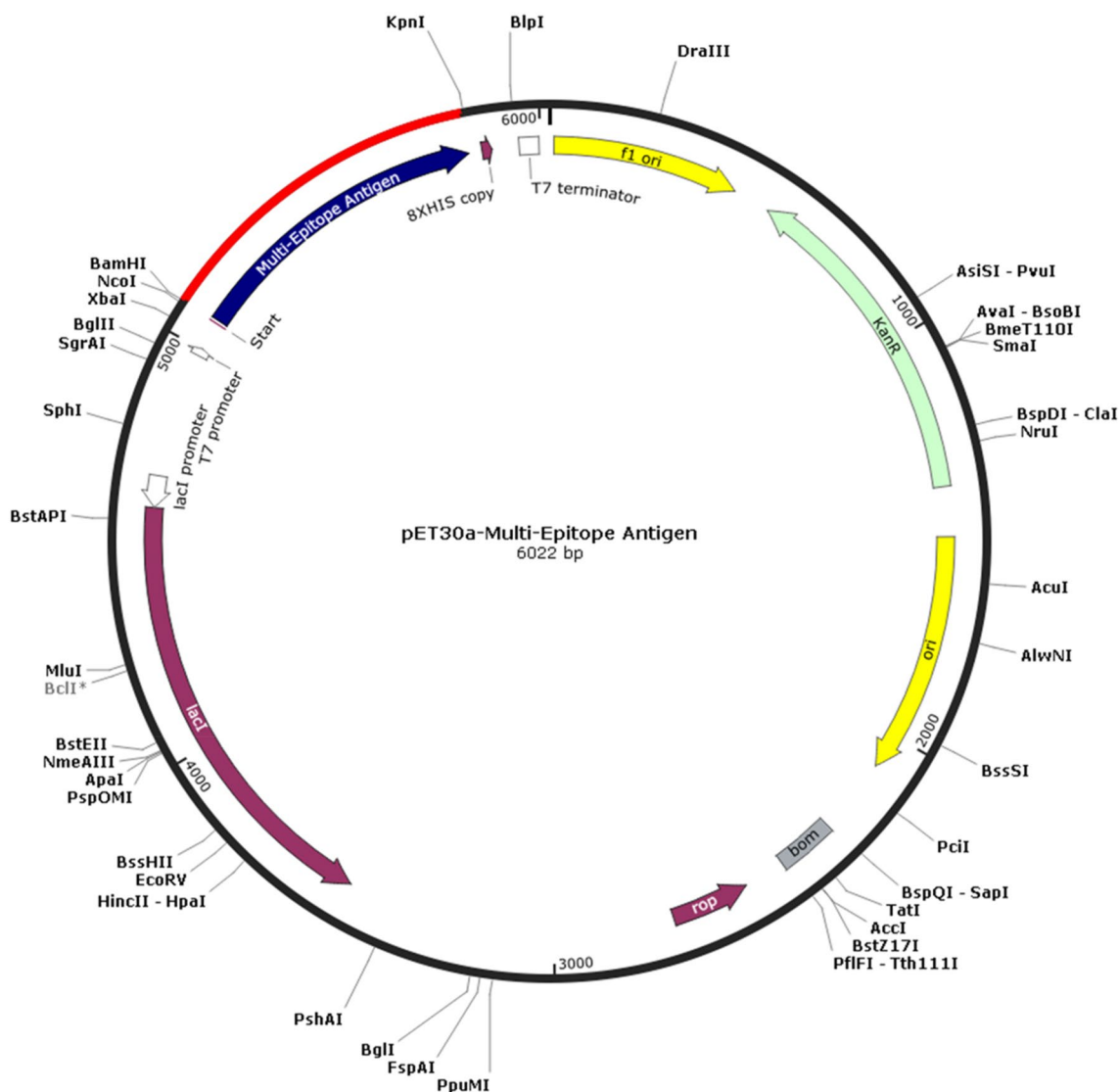
<sup>b</sup> $\Delta^iG$  indicates the solvation free energy gain upon formation of the interface. Negative  $\Delta^iG$  corresponds to hydrophobic interfaces, or positive protein affinity

<sup>c</sup> $\Delta^iG$  P-value indicates the P-value of the observed solvation free energy gain.  $P < 0.5$  indicates interfaces with surprising (higher than would-be-average for given structures) hydrophobicity, implying that the interface surface can be interaction-specific

## Conclusion

Vaccine is an effective and safe measure against pathogen infection. Further, how to improve the effectiveness of vaccines is the key to treat infectious diseases. To develop the vaccine against *H. pylori*, reverse vaccinology based on genomics, proteomics along with bioinformatics was applied for screening the new antigens, PLA1 and Omp6. To enhance the immunity effect, a new construction strategy was applied to develop multi-epitope vaccines based on the epitope analysis of PLA1, Omp6 and UreB. It assesses the stability of vaccine constructs, the location of key domains and probable interaction with TLRs on the basis of the predicted structures. Finally, the best construct was selected. We believe that this strategy will have great potential in multi-epitope vaccine construction and it is worth considering for other pathogen vaccines.





**Fig. 8** In-silico cloning map of multi-epitope vaccine Construct S1. The sequence of multi-epitope antigen is highlighted in red color

**Funding** This work was supported by the National Key R&D Program of China (Grant No. 2018YFC1602500) and Science and Technology innovation Plan of Shanghai (Grant No. 19391902000).

### Compliance with Ethical Standards

**Conflict of interest** All authors declare that they have no conflict of interest to declare.

### References

- Ansari HR, Flower DR, Raghava GPS (2010) AntigenDB: an immunoinformatics database of pathogen antigens. *Nucleic Acids Res* 38:D847–D853
- Barabino A (2002) *Helicobacter pylori*-related iron deficiency anemia: a review. *Helicobacter* 7(2):71–75
- Chang YJ, Wu MS, Lin JT, Sheu BS, Muta T, Inoue H, Chen CC (2004) Induction of cyclooxygenase-2 overexpression in human gastric epithelial cells by *Helicobacter pylori* involves TLR2/TLR9 and c-Src-dependent nuclear factor-kappa B activation. *Mol Pharmacol* 66(6):1465–1477
- Colovos C, Yeates TO (1993) Verification of protein structures: patterns of nonbonded atomic interactions. *Protein Sci* 2(9):1511–1519
- Corthesy-Theulaz IE, Hopkins S, Bachmann D, Saldinger PF, Porta N, Haas R, Zheng-Xin Y, Meyer T, Bouzourene H, Blum AL, Kraehenbuhl JP (1998) Mice are protected from *Helicobacter pylori* infection by nasal immunization with attenuated *Salmonella*

- typhimurium* phoP(c) expressing urease A and B subunits. *Infect Immun* 66(2):581–586
- Ding SZ, Torok AM, Smith MF, Goldberg JB (2005) Toll-like receptor 2-mediated gene expression in epithelial cells during *Helicobacter pylori* infection. *Helicobacter* 10(3):193–204
- Dormitzer PR, Grandi G, Rappuoli R (2012) Structural vaccinology starts to deliver. *Nat Rev Microbiol* 10(12):807–813
- Doytchinova IA, Flower DR (2007) VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *Bmc Bioinform* 8:4
- Emilia G, Luppi M, Torelli G (2004) *Helicobacter pylori* and idiopathic thrombocytopenic purpura. *Ann Intern Med* 141(10):820–820
- Every AL, Stent A, Moloney MB, Ng GZ, Skene CD, Edwards SJ, Sutton P (2011) Evaluation of superoxide dismutase from *Helicobacter pylori* as a protective vaccine antigen. *Vaccine* 29(7):1514–1518
- Fu L, Niu B, Zhu Z, Wu S, Li W (2012) CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* 28(23):3150–3152
- Giuliani MM, Adu-Bobie J, Comanducci M, Arico B, Savino S, Santini L, Brunelli B, Bambini S, Biolchi A, Capecci B, Cartocci E, Ciocchi L, Di Marcello F, Ferlicca F, Galli B, Luzzi E, Massignani V, Serruto D, Veggi D, Contorni M, Morandi M, Bartalesi A, Cinotti V, Mannucci D, Titta F, Ovidit E, Welsch JA, Granoff D, Rappuoli R, Pizza M (2006) A universal vaccine for serogroup B meningococcus. *Proc Natl Acad Sci USA* 103(29):10834–10839
- Gold MR, Reth MG (2019) Antigen receptor function in the context of the nanoscale organization of the B cell membrane. In: Yokoyama WM (ed) *Annual review of immunology*. Palo Alto, Annual Reviews
- Grote A, Hiller K, Scheer M, Munch R, Nortemann B, Hempel DC, Jahn D (2005) JCat: a novel tool to adapt codon usage of a target gene to its potential expression host. *Nucleic Acids Res* 33:W526–W531
- Guo L, Yang H, Tang F, Yin R, Liu H, Gong X, Wei J, Zhang Y, Xu G, Liu K (2017a) 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
- Guo L, Yin R, Xu G, Gong X, Chang Z, Hong D, Liu H, Ding S, Han X, Li Y, Tang F, Liu K (2017b) Immunologic properties and therapeutic efficacy of a multivalent epitope-based vaccine against four *Helicobacter pylori* adhesins (urease, Lpp20, HpaA, and CagL) in Mongolian gerbils. *Helicobacter*. <https://doi.org/10.1111/hel.12428>
- Hugo Urrutia-Baca V, Gomez-Flores R, Angelica De La Garza-Ramos M, Tamez-Guerra P, Guadalupe Lucio-Sauceda D, Cristina Rodriguez-Padilla M (2019) Immunoinformatics approach to design a novel epitope-based oral vaccine against *Helicobacter pylori*. *J Comput Biol* 26(10):1177–1190
- Jensen KK, Andreatta M, Marcatili P, Buus S, Greenbaum JA, Yan Z, Sette A, Peters B, Nielsen M (2018) Improved methods for predicting peptide binding affinity to MHC class II molecules. *Immunology* 154(3):394–406
- Jespersen MC, Peters B, Nielsen M, Marcatili P (2017) BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. *Nucleic Acids Res* 45(W1):W24–W29
- Jurtz V, Paul S, Andreatta M, Marcatili P, Peters B, Nielsen M (2017) NetMHCpan-4.0: improved peptide-MHC Class I interaction predictions integrating eluted ligand and peptide binding affinity data. *J Immunol* 199(9):3360–3368
- Kavoosi M, Creagh AL, Kilburn DG, Haynes CA (2007) Strategy for selecting and characterizing linker peptides for CBM9-tagged fusion proteins expressed in *Escherichia coli*. *Biotechnol Bioeng* 98(3):599–610
- Klement M, Liu C, Loo BLW, Choo AB-H, Ow DS-W, Lee D-Y (2015) Effect of linker flexibility and length on the functionality of a cytotoxic engineered antibody fragment. *J Biotechnol* 199:90–97
- Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* 372(3):774–797
- Kulp DW, Schief WR (2013) Advances in structure-based vaccine design. *Curr Opin Virol* 3(3):322–331
- Livingston B, Crimi C, Newman M, Higashimoto Y, Appella E, Sidney J, Sette A (2002) A rational strategy to design multi-epitope immunogens based on multiple th lymphocyte epitopes. *J Immunol* 168(11):5499–5506
- Lovell SC, Davis IW, Adrenall WB, de Bakker PIW, Word JM, Prisant MG, Richardson JS, Richardson DC (2003) Structure validation by C alpha geometry: phi, psi and C beta deviation. *Protein-Struct Funct Genet* 50(3):437–450
- Lundberg U, Senn BM, Schueler W, Andreas M, Hanner M (2013) Identification and characterization of antigens as vaccine candidates against *Klebsiella pneumoniae*. *Human Vacc Immunother* 9(3):497–505
- Meza B, Ascencio F, Sierra-Beltran AP, Torres J, Angulo C (2017) A novel design of amulti-antigenic, multistage and multi-epitope vaccine against *Helicobacter pylori*: an in silico approach. *Infect Genet Evol* 49:309–317
- Neefjes J, Jongstra MLM, Paul P, Bakke O (2011) Towards a systems understanding of MHC class I and MHC class II antigen presentation. *Nat Rev Immunol* 11(12):823–836
- Negahdaripour M, Nezafat N, Eslami M, Ghoshoon MB, Shoolian E, Najafipour S, Morowvat MH, Dehshahri A, Erfani N, Ghasemi Y (2018) Structural vaccinology considerations for in silico designing of a multi-epitope vaccine. *Infect Genet Evolut* 58:96–109
- Netea MG, van der Graaf C, Van der Meer JWM, Kullberg BJ (2004) Toll-like receptors and the host defense against microbial pathogens: bringing specificity to the innate-immune system. *J Leukoc Biol* 75(5):749–755
- Nezafat N, Eslami M, Negahdaripour M, Rahbar MR, Ghasemi Y (2017) Designing an efficient multi-epitope oral vaccine against *Helicobacter pylori* using immunoinformatics and structural vaccinology approaches. *Mol BioSyst* 13(4):699–713
- Nosrati M, Behbahani M, Mohabatkar H (2019) Towards the first multi-epitope recombinant vaccine against Crimean-Congo hemorrhagic fever virus: a computer-aided vaccine design approach. *J Biomed Inf* 93
- Okushin K, Takahashi Y, Yamamichi N, Shimamoto T, Enooku K, Fujinaga H, Tsutsumi T, Shintani Y, Sakaguchi Y, Ono S, Kodashima S, Fujishiro M, Moriya K, Yotsuyanagi H, Mitsushima T, Koike K (2015) *Helicobacter pylori* infection is not associated with fatty liver disease including non-alcoholic fatty liver disease: a large-scale cross-sectional study in Japan. *BMC Gastroenterol* 15:25
- Pan X, Ke H, Niu X, Li S, Lv J, Pan L (2018) Protection against *Helicobacter pylori* infection in BALB/c mouse model by oral administration of multivalent epitope-based vaccine of cholera toxin B subunit-HUUC. *Front Immunol* 9:10. <https://doi.org/10.3389/fimmu.2018.01003>
- Peek RM, Blaser MJ (2002) *Helicobacter pylori* and gastrointestinal tract adenocarcinomas. *Nat Rev Cancer* 2(1):28–37
- Peng X, Zhang R, Duan G, Wang C, Sun N, Zhang L, Chen S, Fan Q, Xi Y (2018) Production and delivery of *Helicobacter pylori* NapA in *Lactococcus lactis* and its protective efficacy and immune modulatory activity. *Scientific Reports* 8:6435
- Pierce BG, Wiehe K, Hwang H, Kim B-H, Vreven T, Weng Z (2014) ZDOCK server: interactive docking prediction of protein-protein complexes and symmetric multimers. *Bioinformatics* 30(12):1771–1773
- Rey S, Acab M, Gardy JL, Laird MR, DeFays K, Lambert C, Brinkman FSL (2005) PSORTdb: a protein subcellular localization database for bacteria. *Nucleic Acids Res* 33:D164–D168

- Scheiblhofer S, Laimer J, Machado Y, Weiss R, Thalhamer J (2017) Influence of protein fold stability on immunogenicity and its implications for vaccine design. *Expert Rev Vacc* 16(5):479–489
- Shanmugham B, Pan A (2013) Identification and characterization of potential therapeutic candidates in emerging human pathogen *Mycobacterium abscessus*: a novel hierarchical in silico approach. *PLoS ONE* 8(3):10. <https://doi.org/10.1371/journal.pone.0059126>
- Sigal M, Rothenberg ME, Logan CY, Lee JY, Honaker RW, Cooper RL, Passarelli B, Camorlinga M, Bouley DM, Alvarez G, Nusse R, Torres J, Amieva MR (2015) *Helicobacter pylori* activates and expands Lgr5(+) stem cells through direct colonization of the gastric glands. *Gastroenterology* 148(7):1392
- Singh H, Ansari HR, Raghava GPS (2013) Improved method for linear B-cell epitope prediction using antigen's primary sequence. *PLoS ONE* 8(5):e62216
- Suerbaum S, Michetti P (2002) Medical progress: *Helicobacter pylori* infection. *N Engl J Med* 347(15):1175–1186
- Sutton P, Chionh YT (2013) Why can't we make an effective vaccine against *Helicobacter pylori*? *Expert Rev Vacc* 12(4):433–441
- Suzuki S, Esaki M, Kusano C, Ikehara H, Gotoda T (2019) Development of *Helicobacter pylori* treatment: how do we manage antimicrobial resistance? *World J Gastroenterol* 25(16):1907–1912
- Tsirigos KD, Bagos PG, Hamodrakas SJ (2011) OMPdb: a database of beta-barrel outer membrane proteins from Gram-negative bacteria. *Nucleic Acids Res* 39:D324–D331
- Uno K, Kato K, Atsumi T, Suzuki T, Yoshitake J, Morita H, Ohara S, Kotake Y, Shimosegawa T, Yoshimura T (2007) Toll-like receptor (TLR) 2 induced through TLR4 signaling initiated by *Helicobacter pylori* cooperatively amplifies iNOS induction in gastric epithelial cells. *Am J Physiol* 293(5):G1004–G1012
- Van Regenmortel MHV (2009) What is a B-cell epitope? *Methods Mol Biol* 524:3–20
- Vita R, Mahajan S, Overton JA, Dhanda SK, Martini S, Cantrell JR, Wheeler DK, Sette A, Peters B (2019) The immune epitope database (IEDB): 2018 update. *Nucleic Acids Res* 47(D1):D339–D343
- Wang Y-H, Huang Y (2014) Effect of *Lactobacillus acidophilus* and *Bifidobacterium bifidum* supplementation to standard triple therapy on *Helicobacter pylori* eradication and dynamic changes in intestinal flora. *World J Microbiol Biotechnol* 30(3):847–853
- Wiederstein M, Sippl MJ (2007) ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res* 35:W407–W410
- Wilson KT, Crabtree JE (2007) Immunology of *Helicobacter pylori*: insights into the failure of the immune response and perspectives on vaccine studies. *Gastroenterology* 133(1):288–308
- Yang B, Sayers S, Xiang Z, He Y (2011) Protegen: a web-based protective antigen database and analysis system. *Nucleic Acids Res* 39:D1073–D1078
- Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y (2015) The I-TASSER Suite: protein structure and function prediction. *Nat Methods* 12(1):7–8
- Zhang J, Zhao X, Sun P, Gao B, Ma Z (2014) Conformational B-cell epitopes prediction from sequences using cost-sensitive ensemble classifiers and spatial clustering. *Biomed Res Intl* 2014:689219

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.