



B Cell Epitopes of Four Fimbriae Antigens of *Klebsiella pneumoniae*: A Comprehensive In Silico Study for Vaccine Development

Fatemeh Nemati Zargaran¹ · Alisha Akya¹ · Shahab Rezaeian¹ · Keyghobad Ghadiri¹ · Roya Chegene Lorestani¹ · Hamid Madanchi^{2,3} · Sadegh Safaei^{4,5} · Mosayeb Rostamian¹

Accepted: 28 October 2020
© Springer Nature B.V. 2020

Abstract

Klebsiella pneumoniae is one of the major causes of nosocomial infections worldwide which can cause several diseases in children and adults. The globally dissemination of hyper-virulent strains of *K. pneumoniae* and the emergence of antibiotics-resistant isolates of this pathogen narrows down the treatment options and has renewed interest in its vaccines. Vaccine candidates of *Klebsiella pneumoniae* have not been adequately protective, safe and globally available yet. In *K. pneumoniae* infection, it is well known that B cells that induce robust humoral immunity are necessary for the host complete protection. Identifying the B cell epitopes of antigens is valuable to design novel vaccine candidates. In the present study using immunoinformatics approaches we found B cell epitopes of four *K. pneumoniae* type 1 fimbriae antigens namely FimA, FimF, FimG, and FimH. Linear and conformational B cell epitopes of each antigen were predicted using different programs. Subsequently, many bioinformatics assays were applied to choose the best epitopes including prediction antigenicity, toxicity, human similarity and investigation on experimental records. These assays resulted in final four epitopes (each for one Fim protein). These final epitopes were modeled and their physiochemical properties were estimated to be used as potential vaccine candidates. Altogether, we found four B cell epitopes of *K. pneumoniae* Fim antigens that are immunogen, antigenic, not similar to human peptides, not allergen and not toxic. Also, they have suitable physiochemical properties to administrate as vaccine, although their complete efficacy should be also shown in vitro and in vivo.

Keywords *Klebsiella pneumoniae* · Vaccine · Immunoinformatics · B cell epitopes

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s10989-020-10134-3>) contains supplementary material, which is available to authorized users.

✉ Mosayeb Rostamian
mosayeb.rostamian@gmail.com

- ¹ Infectious Diseases Research Center, Health Institute, Kermanshah University of Medical Sciences, Kermanshah, Iran
- ² Department of Biotechnology, School of Medicine, Semnan University of Medical Sciences, Semnan, Iran
- ³ Drug Design and Bioinformatics Unit, Department of Medical Biotechnology, Biotechnology Research Center, Pasteur Institute of Iran, Tehran, Iran
- ⁴ Department of Molecular Medicine, Faculty of Advanced Technologies in Medicine, Iran University of Medical Sciences, Tehran, Iran
- ⁵ Oncopathology Research Center, Iran University of Medical Sciences, Tehran, Iran

Introduction

Klebsiella pneumoniae is one of the major causes of nosocomial infections worldwide (Podschun and Ullmann 1998). *K. pneumoniae* colonizes the humans gastrointestinal tract and it is able to cause infection if gains entry to the blood and other tissues (Piperaki et al. 2017). It can cause meningitis and sepsis in neonates and infants in pediatric wards as well as serious infections in immunocompromised children and adults. *K. pneumoniae* is also one of the causative agents for community acquired pneumonia (Piperaki et al. 2017). The hospitalization and mortality rate for patients infected by *K. pneumoniae* is varying among different countries. As an instance, the overall hospitalization and mortality rate for patients with pneumonia caused by *K. pneumoniae* in the United States is 3.6 cases/100,000 population and 14.3%, respectively, where the hospitalization rate has increased since 2002 to 2011 (Wuerth et al. 2016). The worldwide pooled mortality rate of patients infected with

carbapenem-resistant and carbapenem-susceptible *K. pneumoniae* was reported to be 42.14% and 21.16%, respectively (Xu et al. 2017).

A significant increase of more than 40% in antibiotic resistance has been reported for *K. pneumoniae* which is mainly observed for the following antibiotics: cefazolin, cefuroxime, amoxicillin-clavulanic acid, ceftriaxone, cefepime, and ceftazidime (Kahraman and Çiftcia 2017). The globally dissemination of hyper-virulent strains of *K. pneumoniae* and the emergence of antibiotics-resistant isolates of this pathogen narrows down the treatment options and has renewed interest in its vaccines. A number of vaccine candidates including inactivated whole cell vaccines, capsular polysaccharides, outer membrane vesicles, and highly conserved immunogenic proteins have been proposed for prevention of *K. pneumoniae* infections. A clinical trial aiming *K. pneumoniae* prevention evaluated a multicomponent vaccine called Immunovac in patients with chronic pulmonary diseases (Egorova et al. 2010). This vaccine contained antigenic complexes derived from *K. pneumoniae*, *P. vulgaris*, *S. aureus*, and *E. coli* and was observed to be immunogenic in elderly patients and preschool children (Egorova et al. 2010). However, in spite of numerous attempts, vaccine candidates for *K. pneumoniae* prevention have not been fully protective, safe and globally available yet, demanding new strategies to find more efficient antigens to develop effective vaccine (Pletz et al. 2016; Choi et al. 2019).

Vaccine studies in vivo and in vitro are time-consuming and costly processes. Using bioinformatics approaches such as in silico epitope prediction can make these processes easier by introducing novel or characterizing previous found vaccine candidates (He et al. 2010).

Fimbriae proteins of *K. pneumoniae* have been shown to be effective protein carriers and immunogens (Li et al. 2010; Staniszewska et al. 2000; Lavender et al. 2005; Witkowska et al. 2005; Babu et al. 2017). However, vaccine candidates based on these proteins are at the primarily stages of development and require more studies (Choi et al. 2019). Type 1 fimbriae is composed of 1000 units of FimA (the major subunit) and single copies of FimF (the adaptor subunit), FimG (the tip subunit), and FimH (the adhesion) proteins (Volkan et al. 2014).

It is accepted that humoral (antibody) responses play a critical role in controlling of *K. pneumoniae* infections. The anti-*K. pneumoniae* antibodies evoked in experimental animals have been shown to cause a protective response (Libon et al. 2002; Lee et al. 2015; Hegerle et al. 2018). B cell-specific epitopes are linear or discontinuous amino acid sequences that induce specific, desirable, and broad-spectrum humoral immunity (Chyau Liang 1998). Defining these epitopes will assist to find and characterize the novel vaccine candidates for this gram-negative pathogen.

In our previous work, we predicted T cell epitopes of four type 1 fimbriae antigens namely FimA, FimF, FimG, and FimH using immunoinformatic, docking and molecular dynamics (Rostamian et al. 2020). Due to the importance of humoral immune responses against *K. pneumoniae*, here we performed an in silico study on these four Fim antigens as vaccine candidates and tried to find and characterize B-cell epitopes of these antigens. Furthermore, the selected B cell epitopes were modeled and their physicochemical properties were estimated to be used as potential vaccine candidates.

Methods

Sequences Preparation

The sequences of the single proteins of FimA (accession number: CDO12578.1), FimF (accession number: CDO12574.1), FimG (accession number: CDO12573.1) and FimH (accession number: CDO12572.1) were retrieved from NCBI protein database (<https://www.ncbi.nlm.nih.gov/protein>). Note that in *K. pneumoniae* type 1 fimbriae a bunch of about 1000 identical FimA subunits are present, but here we only select one single subunit of FimA.

Linear B-Cell Epitope Prediction

For prediction of linear epitopes, the Fim protein sequences were submitted to the IEDB (<https://tools.iedb.org/main/bcell/>), (Fleri et al. 2017), and ABCpred (<https://crdd.osdd.net/raghava/abcpred/>) (Saha and Raghava 2006) servers. The IEDB server contains five tools to predict linear B-cell epitopes that are based on the following scales: hydrophilicity (Parker et al. 1986), beta turn (Chou and Fasman 2006), surface accessibility (Emini et al. 1985), flexibility (Karpus and Schulz 1985), and antigenicity (Kolaskar and Tongaonkar 1990). At the first step, the linear B-cell epitopes of Fim proteins were predicted by these five tools. For each tool, as recommended by the tool, the window size was set on seven amino acid residues and the threshold was set on the defaults provided by the tool. Each tool provided a chart in which regions with scores higher than threshold are more probable to be B-cell epitope. All charts provided by the tools were compared and the regions that predicted to be epitopes in at least three (out of five) tools were selected for further analysis.

Also, two other linear B-cell epitope prediction programs namely BepiPred 1.0 (Larsen et al. 2006) and BepiPred 2.0 (Jespersen et al. 2017) can be found in the IEDB server. BepiPred 1.0 use algorithms that combined hidden Markov model and a propensity scale method, while BepiPred 2.0 use a Random Forest algorithm trained on epitopes and non-epitope amino acid residues found from crystal structures

(Larsen et al. 2006; Jespersen et al. 2017). Similar to above mentioned B-cell epitope prediction tools, these tools also report their analyses by charts in which the regions with scores higher than threshold are defined as epitopes.

To find more probable linear B-cell epitope, ABCpred server was also applied. This server predicts epitopes using artificial neural network (Saha and Raghava 2006). The proteins sequences were submitted and other parameters were remained as default. The server presents the more probable linear epitopes ranked according to their score when higher score means the higher probability to be epitope.

Finally, to find similar predicted epitopes, all the epitopes obtained from IEDB (seven tools) and ABCpred were compared using Epitope Cluster Analysis tool (<https://tools.iedb.org/cluster/>) and the common epitopes were selected as final linear B-cell epitopes.

Modelling of Fim proteins

For prediction of conformational B-cell epitopes, the predictors need tertiary structures of our proteins. Since the 3D structures of our proteins were not available in the PDB bank, the structures of Fim proteins were obtained by modeling using SWISS-MODEL server (<https://swissmodel.expasy.org/>, (Waterhouse et al. 2018)).

The amino acid sequence of Fim proteins were submitted to the server. Following target-template alignment, the server found and used the available 3D structure of proteins similar to Fim (usually Fim proteins of other bacteria) as templates. The model accuracy and modeling errors were estimated by the qualitative model energy analysis¹ (QMEAN) scoring function supplied by the server (Waterhouse et al. 2018). Also, the quality of models was assessed by generation of Ramachandran plot using RAMPAGE server (<https://mordred.bioc.cam.ac.uk/~rapper/rampage.php>, (Lovell et al. 2003)). This server use Ramachandran plots to predict the possibility of amino acids to form a secondary structure and show the qualities of models based on the amino acid percentage in the allowed, favoured and outlier regions (Lovell et al. 2003).

Conformational B-Cell Epitope Prediction

The 3D models of Fim proteins were submitted to DiscoTope 2.0 (Kringelum et al. 2012) and ElliPro (Ponomarenko et al. 2008) to predict conformational epitopes. The parameters were remained as the tools defaults. Each tool predicts one or more regions as epitopes. The regions common in predicted epitopes of two tools were considered as final conformational epitopes.

¹ QMEAN: qualitative model energy analysis.

Toxicity, Human Similarity, and Experimental Records

The linear epitopes and the segment sequences of conformational epitopes were checked for potential toxicity using ToxinPred server (Gupta et al. 2013). The epitopes sequences were entered as FASTA format, the support-vector machine² (SVM) algorithm was selected and other parameters were left as the server defaults.

By searching in BLASTP (<https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins>), selected linear and conformational epitopes were screened for similarity with human proteome (taxid 9606). The epitopes with $\geq 90\%$ similarity with human peptides (at both coverage and identity values) were excluded from the study.

The shortlisted epitopes were investigated for possible experimental records using the home page of IEDB (<https://www.iedb.org/>) by selecting human as the host and applying each epitope as query.

Antigenicity and Allergenicity

Antigenic regions of Fim proteins were predicted using the antigenic program of the European Molecular Biology Open Software Suite³ (EMBOSS) (<https://www.bioinformatics.nl/cgi-bin/emboss/antigenic>) that predicts antigenicity using the method of Kolaskar and Tongaonkar (1990). The amino acid sequences of Fim proteins were entered to the program separately and other parameters were remained as defaults. The output represents amino acid sequences with different lengths and different scores that are potentially antigenic regions. These predicted regions were clustered with previously predicted linear and conformational epitopes using Epitope Cluster Analysis tool (<https://tools.iedb.org/cluster/>). The Fim regions that have been predicted in both linear/conformational epitope prediction and antigenicity prediction were selected as the best B-cell epitopes.

Presence of antigenic sequence in these selected epitopes was also confirmed by Antigenic Peptide Prediction⁴ (APP) software (<https://imed.med.ucm.es/Tools/antigenic.pl>). APP predicts those segments within a protein or a peptide sequence that are likely to be antigenic by eliciting an antibody response by Kolaskar and Tongaonkar method (1990).

The allergenicity of the final peptides was checked by Structural Database of Allergenic Proteins⁵ (SDAP) (Ivančić 2003) and AllerCatPro (Maurer-Stroh et al. 2019) servers.

² SVM: support-vector machine.

³ EMBOSS: European Molecular Biology Open Software Suite.

⁴ APP: Antigenic Peptide Prediction.

⁵ SDAP: Structural Database of Allergenic Proteins.

3D Structure and Other Physicochemical Properties of Final Epitopes

The PEP-Fold3 server (<https://bioserv.rpbs.univ-paris-diderot.fr/services/PEP-FOLD3/2018>) was used to predict the tertiary structure of selected epitopes (Lamiable et al. 2016). To achieve the other physicochemical properties of epitopes, ProtParam from ExPasy (<https://web.expasy.org/protparam/>) was used (Adhikari et al. 2018). These properties included molecular weight, half-life, instability index, net charge, hydrophobicity, grand average of hydropathicity index⁶ (GRAVY), and isoelectric pH⁷ (pI).

Moreover, helical wheel projection was carried out to predict the position of amino acids in peptides by Heliquest server (<https://heliquest.ipmc.cnrs.fr/cgi-bin/ComputParams.py>) (Gautier et al. 2008).

Results

Linear B-Cell Epitope Prediction

A B-cell linear epitope search was performed for all Fim proteins using the IEDB (seven tools) and ABCpred. Each tool in the IEDB provided a chart in which regions with scores higher than threshold were more probable to be B-cell epitope. These charts are presented in Figs. S1, S2, S3, and S4. After comparison of all charts provided by the tools the regions that predicted to be epitopes in at least five (out of seven) tools were selected. Then, these selected epitopes were compared to the results of ABCpred tool and clustered using cluster analysis tool with cut-off set on $\geq 70\%$ similarity. Only one epitope of each cluster was selected as final linear B-cell epitope which was the one that had the highest prediction scores or had been predicted by more prediction tools. The number of epitopes predicted by each tool and the selected linear B-cell epitopes of each antigen is presented in Table 1.

Modelling of Fim Proteins

The 3D structures of Fim proteins were modeled by SWISS-MODEL server based on target-template alignment. Regarding quality estimations provided by the server, the models were created with high quality and accuracy (Fig. 1a–c).

Furthermore, the quality of models was also confirmed by Ramachandran plotting that revealed large portions of amino acid in the favored regions (Fig. 1d).

Conformational B-Cell Epitope Prediction

The created models of Fim proteins were submitted to DiscoTope 2.0 and ElliPro to predict conformational B-cell epitope of each antigen. The tools predicted epitopes with different scores. The outputs were compared and the regions predicted to be epitope in both tools or had gotten the highest score by the tools were considered as final conformational epitopes. The number of predicted epitopes by each tool and the selected conformational epitopes has been represented in Table 2. The locations of selected conformational epitope residues for each Fim protein have also been illustrated in Fig. 2.

Toxicity, Human Similarity, and Experimental Records

The toxicity, human similarity and experimental records investigations were performed for the selected linear epitopes and the segment sequences of conformational epitopes. For predicting toxicity the ToxinPred server compared our peptides with 1805 toxic peptides to find any possible homology of our peptides and these toxic peptides. The results of toxicity prediction showed that none of our selected epitopes was toxic for humans. Comparison of the epitopes with human proteome showed that only two epitopes (TAGSTSS of FimA and TADSTGY of FimG) were completely similar (100% identity and 100% coverage) with human peptides, hence were excluded. No experimental record was found for any of our selected epitopes.

Antigenicity Prediction

The antigenicity prediction resulted in the sequences with different lengths and different scores that were potentially antigenic regions (Table 3). These regions were clustered with our previously predicted linear and conformational epitopes and the common regions between them were selected as the best B-cell epitopes. In case more than one common region existed, only one region was selected which had been predicted by more tools or had higher prediction scores (Table 4). Furthermore, for these final peptides APP software predicted that there are antigenic determinants in all of them confirming their high immunogenicity (Table 4).

The allergenicity prediction results showed that none of our final epitopes are allergen.

Structural and Physicochemical Properties Prediction

To simplify the explanations of the results for physicochemical and structural prediction, the symbols of the epitopes (as Table 4) were used. The physicochemical properties of

⁶ GRAVY: grand average of hydropathicity index.

⁷ pI: isoelectric pH.

Table 1 The number of predicted and the selected linear B-cell epitopes

	Prediction tool	Ep. No.*	Selected epitopes**				
			TTTVNGG	CDTTVAT	TAGSTSS	VNGGTVH	AIDSSNT
FimA	Antigenicity	97	–	+	–	+	–
	Beta turn	78	+	+	+	+	+
	Flexibility	94	+	+	+	+	+
	Hydrophilicity	99	+	+	+	+	+
	Surface accessibility	68	+	–	+	–	+
	BepiPred 1.0	11	+	+	+	+	+
	BepiPred 2.0	7	+	–	+	–	–
	ABCpred	13	+	+	–	+	+
			GESKDFT		PVNAPSS		QSRLPVN
FimF	Antigenicity	87	–	–	+	–	+
	Beta turn	77	+	+	+	+	+
	Flexibility	77	+	+	+	+	+
	Hydrophilicity	82	+	+	+	+	+
	Surface accessibility	63	+	+	+	+	+
	BepiPred 1.0	10	+	+	+	–	+
	BepiPred 2.0	6	+	+	–	–	–
	ABCpred	13	+	+	+	+	+
			TADSTGY	QNEGDTT	YKNQGTA	LNNGSSQ	TGYKYNQ
FimG	Antigenicity	83	–	–	–	–	–
	Beta turn	68	+	+	+	+	+
	Flexibility	69	+	+	+	+	+
	Hydrophilicity	85	+	+	+	+	+
	Surface accessibility	52	+	+	+	+	+
	BepiPred 1.0	7	+	+	+	+	+
	BepiPred 2.0	5	+	+	+	+	+
	ABCpred	18	+	+	+	+	+
			TNNYNSD	SRTDKPW	GTTADSA	QTNNYNS	
FimH	Antigenicity	157	–	–	–	–	–
	Beta turn	143	+	+	+	+	+
	Flexibility	141	+	+	+	+	+
	Hydrophilicity	155	+	+	+	+	+
	Surface accessibility	87	+	+	+	+	+
	BepiPred 1.0	17	–	+	+	+	–
	BepiPred 2.0	13	+	–	+	+	+
	ABCpred	29	+	+	+	+	+

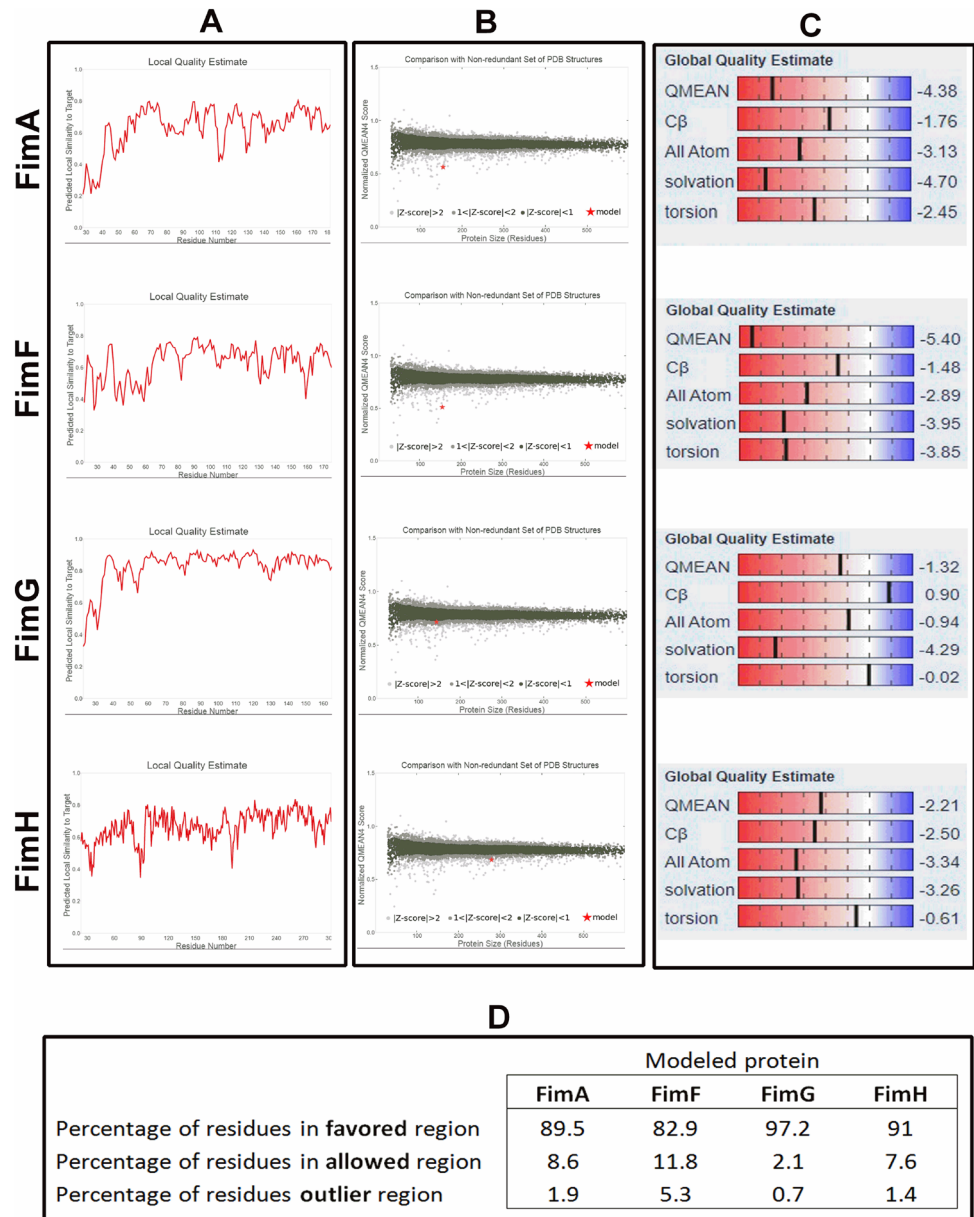
*Number of epitopes predicted by each tool

**The selected epitopes of each antigen (those predicted by a tool are shown by "+")

final epitopes including molecular weight, half-life, instability index, net charge, hydrophobicity, GRAVY, and pI have been presented in Table 5. These results showed that BEF-G and BEF-H were naturally basic, stable, and hydrophilic while BEF-A sequence was acidic, stable and less hydrophilic. BEF-F was also acidic, hydrophilic but unstable. Hydrophobic and hydrophilic faces of each peptide were drawn up by helical wheel projection (Fig. 3).

The prediction results of the 3D structure of the peptides revealed that BEF-A and BEF-F are mixture of helix and coil while BEF-G and BEF-H are a combination of β -sheet, turn and coil (Fig. 4).

Fig. 1 Evaluation of the models quality. After homology modeling of Fim proteins, the model quality was evaluated by Local Quality Estimate (*LQE* Local Quality Estimate) (*LQE*) plots, Comparison with Non-redundant Set of PDB Structures, and Global Quality Estimate (*GQE* Global Quality Estimate) (*GQE*) prepared by SWISS-MODEL server, as well as by Ramachandran plotting by RAMPAGE server. In *LQE* plots (a), residues with a score above 0.6 are expected to be of acceptable quality. In comparison plots (b), whatever the model (indicated by a red star) is near to the dark grey region, the model has a higher quality. In *GQE* plots (c), for each term the Z-score number around zero shows better agreement between the models and experimental structures of similar size. The results of Ramachandran plots (d) are summarized in a table that show the percentage of amino acid residues in favored, allowed, and outlier regions



Discussion

In the present study using immunoinformatics approaches we found B cell epitopes of four Fim proteins as vaccine candidates against *K. pneumoniae*. B cell epitopes may be linear or conformational. Linear epitopes are short sequences that elicit humoral responses even if the antigen administered in the denatured form. Conversely, conformational epitopes consist of atoms from distant amino acid residues joined on the surface of the antigens. These epitopes may be bound to either an immunoglobulin or a B cell receptor and triggers humoral responses (Chyau Liang 1998).

B-cell epitope prediction resulted in the final 5, 3, 5, and 4 linear epitopes with 7-amino acid sequences. Since we

defined the window size of prediction tools on 7, the predicted epitopes are all 7-amino acid sequences. It does not mean that all B-cell epitopes should be restrictively 7-amino acid length, but it means that these seven amino acids are the core of B-cell epitopes and almost always present in antibody-epitope interactions.

In contrast, conformational epitope predictions lead to different numbers of epitopes for each Fim antigen with various sequence lengths. For FimA, a conformational epitope was found (amino acid number 29 to number 35) that had been also found by linear epitope prediction emphasizing the importance of this region as an epitope. For other Fim proteins, a conformational epitope was found for each antigen that composed of discontinues parts. There was no

Table 2 The number of predicted and the selected conformational B-cell epitopes

Antigen	Prediction tool	Ep. No.*	Selected epitope**
			<u>29NGGTVHF35</u>
FimA	DiscoTope 2.0	1	+
	ElliPro	4	+
			23ADSTIAISGYVRDNACA39-76SPCGTSVTA84-137TTLTPGQTN145-176Q [#]
FimF	DiscoTope 2.0	0	–
	ElliPro	3	+
			<u>24DVTLTVNGKVVARP37-77PVG79</u>
FimG	DiscoTope 2.0	1	+
	ElliPro	3	+
			<u>52NVG54-183G183-193VTLPDYPGSMAMP205-211AQSQQLG216-</u> <u>222TTADSANAIFTTASSPAQIG245-249TRNGSAVPANSTVSLGTVGT-</u> <u>SPVNLG274-279YARTTGQVTAGNVQS293-302Q^{&}</u>
FimH	DiscoTope 2.0	1	+
	ElliPro	2	+

*Number of epitopes predicted by each tool

**The selected conformational epitope of each antigen. The conformational epitopes are composed of one or more segments. Each segment is separated by dash. The number of first and last amino acid residue of each segment is represented. The underlined bold sequences are the parts that were predicted by DiscoTope 2.0. The epitope is shown by "+" if predicted by the tool

[#]No epitope was found for FimF using DiscoTope 2.0

[&]This epitope is predicted by DiscoTope 2.0, but many segments of it were also present in a very large epitope predicted by ElliPro (not shown here)

Fig. 2 The locations of selected conformational epitope residues in Fim proteins. The location of selected epitopes are depicted by yellow color. The Fim A, FimF, and FimG epitope location have been extracted from ElliPro server, while the FimH epitope was extracted from DiscoTope 2.0 server

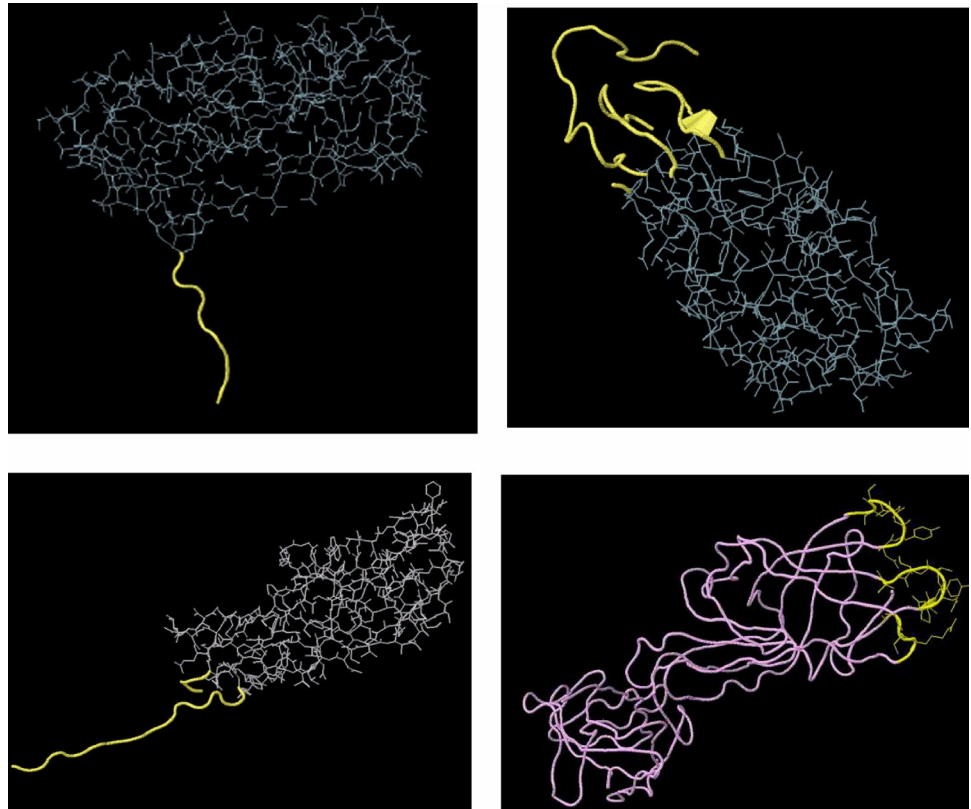


Table 3 The antigenicity prediction of Fim proteins

	Antigenic regions*	Max score**
FimA	KTLAMIVVSALSLSSTAALA	13
	GGTVHFKEVVNAACAVDAG	43
	NVGVQIL	125
	DQTVQLGQVRSACLATA	57
	TTVLALQNS	109
	ATKASVTFAGT	91
	SSAVGFNIQLDDCDTT	78
	PLALDGA	136
	IIPFQARYYATG	158
FimF	KQFYAVGATTPPVPRIVLSPCGTSVTAVKVGFTGVADSVNTSLLKLD	72
	LQYLLGALFTLGAPAALAADSTIAISGY	10
	LNFYARLMATQVPVTAGHVNT	161
	DNACAVAG	40
	GVEILDQQSRLPVNAP	128
	DFTVDLQ	49
FimG	VGWLLAGLLTASASLRAADVTLTVNGKVVARPCTVSTVNATVELGD-LYTFSLIGAG	35
	ARFPLQVRALSVN	143
	SAWHSVALDLSNCPVGTSRVKAT	70
	IQAVINVTY	158
	SQSVQVDET	125
FimH	SMAVPLTVHCAQSQQLGYYLSG	207
	WPAVLYLTPVSTAGGVAITAGSLIAVLILH	151
	ANVYVNLTPAVNVGQNLVVDLSTQIFCHND	56
	IIPLFTLLLLG	11
	NNDVVVPTGGCDVSARDVTVTLPDY	181
	AGNVQSIIGITF	294
	SAVPANSTVSLGTVTGSPVNLGLTAT	263
	YGGVLSFSFGTV	92
	TDYVTLQ	78
	ASPAQGIGVQLT	245
	ARVIYD	117
	SFACKTA	26
FQFIWN	170	

*The antigenic regions predicted by the antigenic program of EMBOSS

**The maximum score given by the antigenic program of EMBOSS. The higher numbers show regions with more antigenicity potency

relationship between discontinues parts and the final predicted linear epitopes. However, it does not make a problem since conformational epitopes may not be necessarily associated with linear ones.

We applied many other assays to choose the best epitopes including prediction antigenicity, allergenicity, toxicity, human similarity and investigation on experimental records. These assays resulted in final four epitopes (each for one Fim protein) that were immunogen, antigenic, not similar to human peptides, not allergen and not toxic.

To find conformational epitops, prediction tools need 3D structure of the antigens.

If the tertiary structure of the antigens is known, the predicted tools can find the conformational epitopes more efficiently. Otherwise, in silico approaches provide a way to model the structure of the antigens and map their conformational epitopes. Accordingly, in the present study we modeled the Fim antigens using in silico approaches. Generally the quality estimations showed that our models are acceptable, although some quality scores were low for some models.

Although no study was found on our Fim antigens, there are some bioinformatics studies on other antigens of *K. pneumoniae* regarding B cell epitope mapping. In a study

Table 4 Final B-cell epitopes of Fim protein

Antigen	Parameter	Sequence	Epitope symbol [#]
FimA	Final region	VNGGTVHFKGEVVNAACAVDAG	BEF-A
	Linear epitope	VNGGTVH	
	Conformational epitope	NGGTVHF	
	Antigenic region*	GGTVHFKGEVVNAACAVDAG	
	Antigenic determinant**	GTVHFKGEVVNAAC	
FimF	Final region	GVEILDQQQSRLPVPNAPSS	BEF-F
	Linear epitope	QSRLPVPN and PVPNAPSS	
	Conformational epitope	None	
	Antigenic region*	GVEILDQQQSRLPVPNAP	
	Antigenic determinant**	ILDQQQSRLPV	
FimG	Final region	SAWHSVALDLSNCPVGTSRVKAT	BEF-G
	Linear epitope	None	
	Conformational epitope	PVG	
	Antigenic region*	SAWHSVALDLSNCPVGTSRVKAT	
	Antigenic determinant**	HSVALDLSNCPVGTS	
FimH	Final region	TRNGSAVPANSTVSLGTVGTSPVNLGLTAT	BEF-H
	Linear epitope	None	
	Conformational epitope	TRNGSAVPANSTVSLGTVGTSPVNLG	
	Antigenic region*	SAVPANSTVSLGTVGTSPVNLGLTAT	
	Antigenic determinant**	GSAVPANSTVSLGTVGTSPVNL	

*Antigenic region predicted by antigenic program of EMBOSS

**Antigenic determinant predicted by Antigenic Peptide Prediction (APP)

[#]Symbols were chosen from first letters of the phrase "B cell Epitope of Fim"**Table 5** Predicted physicochemical properties of epitopes

Epitope	Mw ^a	Estimated Half life	Hyd ^b (%)	Nt ^c	pI ^d	GRAVY ^e	Instability index
BEF-A	2115.4	100 h (mammalian reticulocytes, in vitro) > 20 h (yeast, in vivo) > 10 h (<i>E. coli</i> , in vivo)	50	-0.75	5.32	0.459	-1.23-Stable
BEF-F	2038.2	30 h (mammalian reticulocytes, in vitro) > 20 h (yeast, in vivo) > 10 h (<i>E. coli</i> , in vivo)	31	-1	4.37	-0.484	133.97-Unstable
BEF-G	2399.7	1.9 h (mammalian reticulocytes, in vitro) > 20 h (yeast, in vivo) > 10 h (<i>E.coli</i> , in vivo)	43	1.25	7.97	0.087	35.10-Stable
BEF-H	2843.1	7.2 h (mammalian reticulocytes, in vitro) > 20 h (yeast, in vivo) > 10 h (<i>E.coli</i> , in vivo)	33	1	9.41	0.213	34.19-Stable

^aMolecular weight^bHydrophobicity ratio^cNet charge^dIsoelectric pH^eGrand average of hydropathicity index

done by Dar et al. 222 genomes of *K. pneumoniae* were explored and the core proteome was found and applied for a deep reverse vaccinology. They identified four antigens and predicted their B cell epitopes using ABCPred (Dar

et al. 2019). By applying several bioinformatics approaches and based on conformational B cell and linear CD4+ T cell epitopes, Farhadi et al. engineered a polytopic vaccine using five Omps from *K. pneumoniae* (Farhadi et al. 2015). These

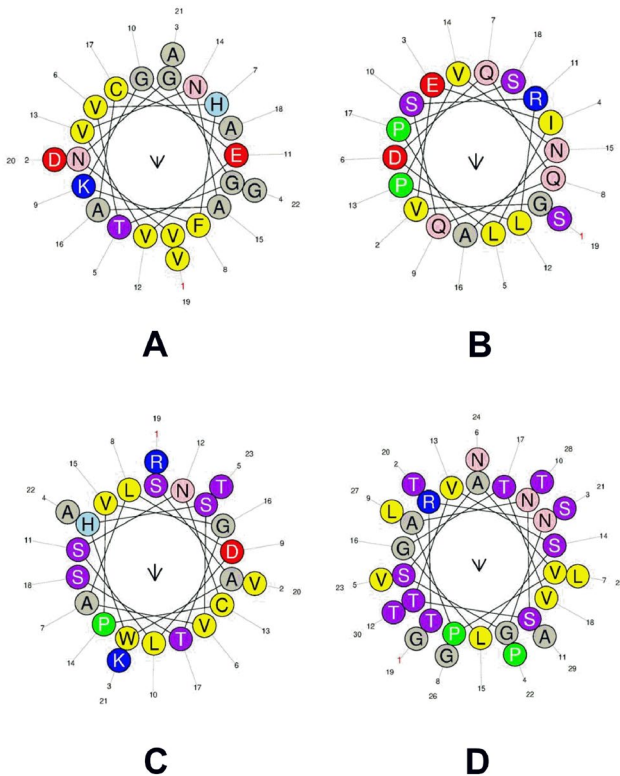


Fig. 3 The helical wheel projection of the final epitopes. The helical wheel projection of BEF-A (a), BEF-F (b), BEF-G (c) and BEF-H (d) are shown. In all epitopes, polar and non-polar amino acids are distributed irregularly and disrupting the uniformity of hydrophobic and hydrophilic faces. The amino acid residues are shown by standard single letters

studies and our study emphasize on the potential ability of prediction tools on B cell epitope-based vaccine candidates for *K. pneumoniae*.

If an epitopic peptide is going to be used as vaccine candidate, it would be essential to elucidate their structures and physicochemical properties. Therefore, in the present study, the physicochemical and structural properties of final four epitopes were investigated. These investigations include molecular weight, half-life, instability index, net charge, hydrophobicity, GRAVY, pI and 3D structure that all of them are valuable and increases the accuracy and precision of the epitope design (Adhikari et al. 2018). Isoelectric point (pI) is an important factor to select a good epitope (Campos-Pinto et al. 2019). The peptides at their own isoelectric point carries no net electrical charge, their immunogenicity is reduced because antibodies and other specific immune receptors such as B cell receptor prefer charged and polar antigen residues (Wang et al. 2018). The pH of the blood and the body tissues is neutral (pH 7.2–7.6), therefore, it is better that pI of the epitopes not be close to the pH of the blood and body tissues. In this study, pI of all epitopes, except BEF-G, is far from the physiological pH of the body.

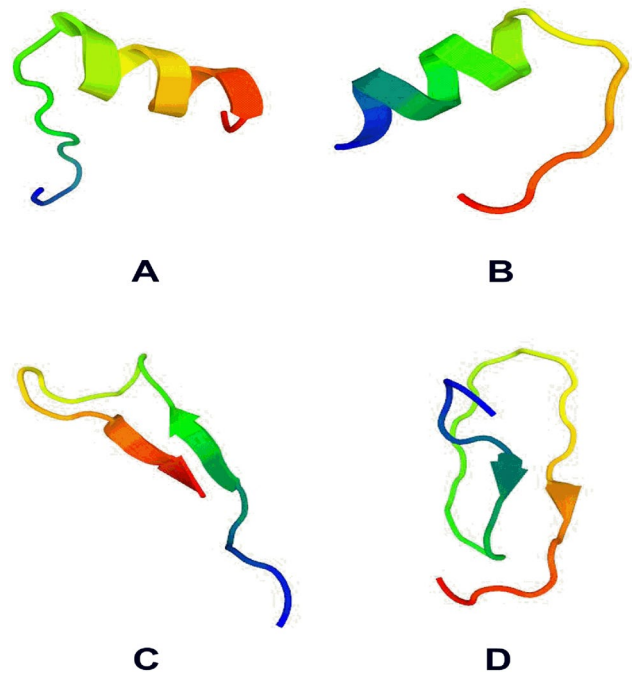


Fig. 4 Structure prediction. 3D structure of epitopes was predicted by PEP-Fold3 server. The 3D structure of BEF-A, BEF-F, BEF-G, and BEF-H are shown in parts A, B, C, and D, respectively

GRAVY indicates the rate of hydrophilicity or hydrophobicity of a protein. Peptides with positive GRAVY are more hydrophobic while peptides with a negative GRAVY value are more hydrophilic and tend to be more water soluble (Kyte and Doolittle 1982; Adhikari et al. 2018). All of our final epitopes were in the suitable range in this regard. Increasing the hydrophobic ratio in the non-polar face of the peptides to increase the symmetry in hydrophobic and hydrophilic faces, can enhanced their toxicity and hemolytic activity (Boland and Separovic 2006; Madanchi et al. 2019). Helical wheel projection showed that in all epitopes, the hydrophilic face became more discontinuous by polar residues, so the symmetry in hydrophobic and hydrophilic faces was not seen in our epitopes. Therefore, this arrangement reduces the possibility of cytotoxicity and hemolytic activity of these peptides to humans. This finding confirms the results of toxicity prediction done by ToxinPred server and showed that our final four epitopic peptides are non toxic.

Altogether, here we found four B cell epitopes (each for one Fim antigen) that are immunogen, antigenic, not similar to human peptides, not allergen and not toxic. They have also suitable physicochemical properties required for vaccine candidates, although their actual efficacy should be investigated using in vitro and in vivo testing.

It should be noted that numerous challenges can be readily predictable in developing *K. pneumoniae* vaccines, but not considered in our study. The first one is the developed

vaccines against *K. pneumoniae* should not have cross-reactivity with the normal flora of intestine where this bacterium colonizes (Ahmad et al. 2012). The second, patients vulnerable to *K. pneumoniae* are often elderly and suffer multiple comorbidities. Consequently, immune responses of these populations to vaccine may be poor and not induce protective immunity (Weinberger et al. 2008).

Acknowledgements The supported of this work by Infectious Diseases Research Center of Kermanshah University of Medical Sciences and Clinical Research Development Center of Imam Reza Hospital is highly appreciated.

Author Contributions FNZ, MR and AA contributed to the study conception and design. Material preparation, data collection and analysis were performed by FNZ, SR, RCL, HM and SS. Funding acquisition and resources preparation was done by KG. The first draft of the manuscript was written by MR and AA. All authors read and approved the final manuscript.

Funding This work was supported by the Kermanshah University of Medical Sciences.

Compliance with Ethical Standards

Conflict of interest The authors declare no conflict of interest.

References

- Adhikari UK, Tayebi M, Rahman MM (2018) Immunoinformatics approach for epitope-based peptide vaccine design and active site prediction against polyprotein of emerging oropouche virus. *J Immunol Res* 2018:6718083. <https://doi.org/10.1155/2018/6718083>
- Ahmad TA, El-Sayed LH, Haroun M et al (2012) Development of immunization trials against *Klebsiella pneumoniae*. *Vaccine* 30:2411–2420
- Babu L, Uppalapati SR, Sripathy MH, Reddy PN (2017) Evaluation of recombinant multi-epitope outer membrane protein-based *Klebsiella pneumoniae* subunit vaccine in mouse model. *Front Microbiol*. <https://doi.org/10.3389/fmicb.2017.01805>
- Boland MP, Separovic F (2006) Membrane interactions of antimicrobial peptides from Australian tree frogs. *Biochim Biophys Acta - Biomembr* 1758:1178–1183
- Campos-Pinto I, Méndez L, Schouten J et al (2019) Epitope mapping and characterization of 4-hydroxy-2-nonenal modified-human serum albumin using two different polyclonal antibodies. *Free Radic Biol Med* 144:234–244. <https://doi.org/10.1016/j.freeradbiomed.2019.05.008>
- Choi M, Tennant SM, Simon R, Cross AS (2019) Progress towards the development of *Klebsiella* vaccines. *Expert Rev Vaccines* 18:681–691
- Chou PY, Fasman GD (2006) Prediction of the secondary structure of proteins from their amino acid sequence. In: Meister A (ed) *Advances in enzymology and related areas of molecular biology*. Wiley Blackwell, Hoboken, pp 45–148
- Chyau Liang T (1998) *Encyclopedia of immunology*, 2nd edn. Elsevier, Amsterdam
- Dar HA, Zaheer T, Shehroz M et al (2019) Immunoinformatics-aided design and evaluation of a potential multi-epitope vaccine against *Klebsiella pneumoniae*. *Vaccines*. <https://doi.org/10.3390/vaccines7030088>
- Egorova NB, Kurbatova EA, Gruber IM, Semenova IB (2010) Immunovac-VP-4 for prevention of acute respiratory diseases in children's communities. *Zh Mikrobiol Epidemiol Immunobiol* 6:35–40
- Emini EA, Hughes JV, Perlow DS, Boger J (1985) Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide. *J Virol* 55:836–839. <https://doi.org/10.1128/jvi.55.3.836-839.1985>
- Farhadi T, Nezafat N, Ghasemi Y et al (2015) Designing of complex multi-epitope peptide vaccine based on Omps of *Klebsiella pneumoniae*: an in silico approach. *Int J Pept Res Ther* 21:325–341. <https://doi.org/10.1007/s10989-015-9461-0>
- Fleri W, Paul S, Dhanda SK et al (2017) the immune epitope database and analysis resource in epitope discovery and synthetic vaccine design. *Front Immunol* 8:278. <https://doi.org/10.3389/fimmu.2017.00278>
- Gautier R, Douguet D, Antony B, Drin G (2008) HELIQUEST: a web server to screen sequences with specific alpha-helical properties. *Bioinformatics* 24:2101–2102. <https://doi.org/10.1093/bioinformatics/btn392>
- Gupta S, Kapoor P, Chaudhary K et al (2013) In silico approach for predicting toxicity of peptides and proteins. *PLoS ONE*. <https://doi.org/10.1371/journal.pone.0073957>
- He Y, Rappuoli R, De Groot AS, Chen RT (2010) Emerging vaccine informatics. *J Biomed Biotechnol* 2010:1–26. <https://doi.org/10.1155/2010/218590>
- Hegerle N, Choi M, Sinclair J et al (2018) Development of a broad spectrum glycoconjugate vaccine to prevent wound and disseminated infections with *Klebsiella pneumoniae* and *Pseudomonas aeruginosa*. *PLoS ONE*. <https://doi.org/10.1371/journal.pone.0203143>
- Ivanciu O (2003) SDAP: database and computational tools for allergenic proteins. *Nucleic Acids Res* 31:359–362. <https://doi.org/10.1093/nar/gkg010>
- Jespersen MC, Peters B, Nielsen M, Marcotili P (2017) BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. *Nucleic Acids Res* 45:W24–W29. <https://doi.org/10.1093/nar/gkx346>
- Kahraman E, Çiftcia I (2017) The antibiotic resistance patterns of *Klebsiella pneumoniae* clinic isolates: a comprehensive meta-analysis. *Open J Bacteriol* 1:021–026. <https://doi.org/10.17352/ojb.000005>
- Karplus PA, Schulz GE (1985) Prediction of chain flexibility in proteins - a tool for the selection of peptide antigens. *Naturwissenschaften* 72:212–213. <https://doi.org/10.1007/BF01195768>
- Kolaskar AS, Tongaonkar PC (1990) A semi-empirical method for prediction of antigenic determinants on protein antigens. *FEBS Lett* 276:172–174. [https://doi.org/10.1016/0014-5793\(90\)80535-Q](https://doi.org/10.1016/0014-5793(90)80535-Q)
- Kringelum JV, Lundegaard C, Lund O, Nielsen M (2012) Reliable B cell epitope predictions: impacts of method development and improved benchmarking. *PLoS Comput Biol* 8:e1002829. <https://doi.org/10.1371/journal.pcbi.1002829>
- Kyte J, Doolittle RF (1982) A simple method for displaying the hydrophobic character of a protein. *J Mol Biol* 157:105–132. [https://doi.org/10.1016/0022-2836\(82\)90515-0](https://doi.org/10.1016/0022-2836(82)90515-0)
- Lamiab A, Thévenet P, Rey J et al (2016) PEP-FOLD3: faster de novo structure prediction for linear peptides in solution and in complex. *Nucleic Acids Res* 44:W449–W454. <https://doi.org/10.1093/nar/gkw329>
- Larsen JEP, Lund O, Nielsen M (2006) Improved method for predicting linear B-cell epitopes. *Immunome Res* 2:2. <https://doi.org/10.1186/1745-7580-2-2>
- Lavender H, Jagnow JJ, Clegg S (2005) *Klebsiella pneumoniae* type 3 fimbria-mediated immunity to infection in the murine model of

- respiratory disease. *Int J Med Microbiol* 295:153–159. <https://doi.org/10.1016/j.ijmm.2005.04.001>
- Lee WH, Il CH, Hong SW et al (2015) Vaccination with *Klebsiella pneumoniae*-derived extracellular vesicles protects against bacteria-induced lethality via both humoral and cellular immunity. *Exp Mol Med* 47:e183. <https://doi.org/10.1038/emm.2015.59>
- Li Y, Li Z, Han W et al (2010) Identification and characterization of Th cell epitopes in MrkD adhesin of *Klebsiella pneumoniae*. *Microb Pathog* 49:8–13. <https://doi.org/10.1016/j.micpath.2010.03.009>
- Libon C, Haeuw JF, Crouzet F et al (2002) Streptococcus pneumoniae polysaccharides conjugated to the outer membrane protein A from *Klebsiella pneumoniae* elicit protective antibodies. *Vaccine* 20:2174–2180. [https://doi.org/10.1016/S0264-410X\(02\)00149-4](https://doi.org/10.1016/S0264-410X(02)00149-4)
- Lovell SC, Davis IW, Arendall WB 3rd et al (2003) Structure validation by Calpha geometry: phi, psi and Cbeta deviation. *Proteins* 50:437–450. <https://doi.org/10.1002/prot.10286>
- Madanchi H, Akbari S, Shabani AA et al (2019) Alignment-based design and synthesis of new antimicrobial Aurein-derived peptides with improved activity against Gram-negative bacteria and evaluation of their toxicity on human cells. *Drug Dev Res* 80:162–170. <https://doi.org/10.1002/ddr.21503>
- Maurer-Stroh S, Krutz NL, Kern PS et al (2019) AllerCatPro—prediction of protein allergenicity potential from the protein sequence. *Bioinformatics* 35:3020–3027. <https://doi.org/10.1093/bioinformatics/btz029>
- Parker JMR, Guo D, Hodges RS (1986) New hydrophilicity scale derived from high-performance liquid chromatography peptide retention data: correlation of predicted surface residues with antigenicity and X-ray-derived accessible sites. *Biochemistry* 25:5425–5432. <https://doi.org/10.1021/bi00367a013>
- Piperaki ET, Syrogiannopoulos GA, Tzouveleki LS, Daikos GL (2017) *Klebsiella pneumoniae*: virulence, biofilm and antimicrobial resistance. *Pediatr Infect Dis J* 36:1002–1005. <https://doi.org/10.1097/INF.0000000000001675>
- Pletz MW, Uebele J, Gotz K et al (2016) Vaccines against major ICU pathogens: where do we stand? *Curr Opin Crit Care* 22:470–476
- Podschn R, Ullmann U (1998) *Klebsiella* spp. as nosocomial pathogens: epidemiology, taxonomy, typing methods, and pathogenicity factors. *Clin Microbiol Rev* 11:589–603
- Ponomarenko J, Bui HH, Li W et al (2008) ElliPro: a new structure-based tool for the prediction of antibody epitopes. *BMC Bioinform* 9:514. <https://doi.org/10.1186/1471-2105-9-514>
- Rostamian M, Farasat A, Cheghe Lorestani R et al (2020) Immunoinformatics and molecular dynamics studies to predict T-cell-specific epitopes of four *Klebsiella pneumoniae* fimbriae antigens. *J Biomol Struct Dyn*. <https://doi.org/10.1080/07391102.2020.1810126>
- Saha S, Raghava GPS (2006) Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. *Proteins Struct Funct Genet* 65:40–48. <https://doi.org/10.1002/prot.21078>
- Staniszewska M, Witkowska D, Gamian A (2000) Fimbriae as a pathologic factor of bacteria and a carrier in conjugate vaccines. *Postep Hig Med Dosw* 54:727–747
- Volkan E, Kalas V, Hultgren S (2014) Pili and Fimbriae of Gram-negative bacteria. In: Sussman M (ed) *Molecular medical microbiology*, 2nd edn. Elsevier Ltd, Amsterdam, pp 147–162
- Wang M, Zhu D, Zhu J et al (2018) Local and global anatomy of antibody-protein antigen recognition. *J Mol Recognit* 31:e2693. <https://doi.org/10.1002/jmr.2693>
- Waterhouse A, Bertoni M, Bienert S et al (2018) SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res* 46:W296–W303. <https://doi.org/10.1093/nar/gky427>
- Weinberger B, Herndler-Brandstetter D, Schwanninger A et al (2008) Biology of immune responses to vaccines in elderly persons. *Clin Infect Dis* 46:1078–1084. <https://doi.org/10.1086/529197>
- Witkowska D, Mieszala M, Gamian A et al (2005) Major structural proteins of type 1 and type 3 *Klebsiella fimbriae* are effective protein carriers and immunogens in conjugates as revealed from their immunochemical characterization. *FEMS Immunol Med Microbiol* 45:221–230. <https://doi.org/10.1016/j.femsim.2005.04.005>
- Wuerth BA, Bonnewell JP, Wiemken TL, Arnold FW (2016) Trends in pneumonia mortality rates and hospitalizations by organism, United States, 2002–2011. *Emerg Infect Dis* 22:1624–1627. <https://doi.org/10.3201/eid2209.150680>
- Xu L, Sun X, Ma X (2017) Systematic review and meta-analysis of mortality of patients infected with carbapenem-resistant *Klebsiella pneumoniae*. *Ann Clin Microbiol Antimicrob* 16:18. <https://doi.org/10.1186/s12941-017-0191-3>

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