(This is a sample cover image for this issue. The actual cover is not yet available at this time.)

Journal of Theoretical Biology 294 (2012) 29–39

Contents lists available at SciVerse ScienceDirect

Journal of Theoretical Biology

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

A potential in silico antibody–antigen based diagnostic test for precise identification of *Acinetobacter baumannii*

Mohammad Reza Rahbar [a], Iraj Rasooli [a,*], Seyed Latif Mousavi Gargari [a], Gunnar Sandstrom [b], Jafar Amani [c], Yaser Fattahian [a], Abolfazl Jahangiri [a], Marzyeh Jalali [a]

[a] Department of Biology, Shahed University, Tehran-Qom Express Way, Opposite Imam Khomeini's shrine, Tehran 3319118651, Iran
[b] Karolinska Institutet and Karolinska, University Hospital, Huddinge, Department of Laboratory Medicine, Division of Clinical Microbiology, SE-141 86 Stockholm, Sweden
[c] Applied Biotechnology Research Center, Baqiyatallah Medical Science University, Tehran, Iran

ARTICLE INFO

Article history:
Received 6 July 2011
Received in revised form
7 September 2011
Accepted 18 October 2011
Available online 6 November 2011

Keywords:
A. baumannii
Antigenicity
Biofilm associated protein
Bioinformatics

ABSTRACT

Acinetobacter baumannii is a problematic nosocomial pathogen. The resistance to a wide range of antimicrobial agents, attributable to its biofilm phenotype, makes the treatment very difficult. Biofilm is a common feature of most pathogens. Biofilm associated proteins (Bap) are cellular surface components directly involved in biofilm formation process. The dearth of a fast precise diagnostic test and versatility of Bap sequences in *A. baumannii* were intuitions to design this study. In silico analysis is a reliable alternative to laborious experimental work in this connection. Databases were searched for an antigenic conserved region of Bap specific to *A. baumannii.* The region was selected based on alignments and propensity scales. Tertiary structure for this region was built and predicted B-cell epitopes were mapped on the surface of the built model. Our protein subunit was found to be a potential antigen, possessing several antigenic determinants, eliciting antibody. Hence this subunit could be used as a suitable agent for antibody–antigen based diagnostic test. This specific antigen can minimize laboratory errors in identification of *A. baumannii* and thus help clinicians to quick and precise diagnosis of the bacteria and initiatives to the treatment of the infection. Antigenicity of the region could also be explored for elicitation of antibody to protect the individuals exposed to *A. baumannii*.

1. Introduction

Acinetobacter baumannii, is a Gram-negative, non-fermentative bacillus has gained considerable importance in the last decade due to its ability to cause a variety of nosocomial infections, such as respiratory tract, bloodstream and skin and soft-tissue infections (Fournier and Richet, 2006; Katsaragakis et al., 2008). This pathogen is the most common bacterial species isolated from the wounds of soldiers injured in combat zones in Iraq and Afghanistan (Elston et al., 2008; Tien et al., 2007). Several large genomic islands containing multiple resistance genes thought to be acquired from other Gram-negative species are found in *A. baumannii* (Gordon and Wareham, 2010). Resistance of *A. baumannii* to almost every clinically used antibiotics makes treatment of infections very complicated (Jain and Danziger, 2004). Beta-lactamase production (Hujer et al., 2006), presence of aminoglycoside-modifying enzymes (Akers et al., 2010), target site mutations (Hamouda and Amyes, 2004) and multidrug efflux systems (Lin et al., 2009), are of known mechanisms of resistance to antibiotics in this organism. This remarkable resistant phenotype could be attributed to the ability of *A. baumannii* clinical strains to form biofilms on abiotic and biotic surfaces (Loehfelm et al., 2008; Tomaras et al., 2003; Vidal et al., 1996, 1997). *A. baumannii* easily adheres both to biological and abiotic surfaces, on which it is able to form biofilms (Cevahir et al., 2008; Lee et al., 2008). Biofilm phenotype is known as a vital pathogenic feature of most pathogens, in which microorganisms enclosed in a polymeric matrix called exopolysaccharide (EPS) to form a sessile community (Monds and O'Toole, 2009). Biofilm formation involves a variety of pathways that are regulated by quorum sensing and a number of two-component regulatory systems (Gaddy and Actis, 2009). Also several cellular components are directly involved in biofilm formation (Stoodley et al., 2002) one of these apparatuses is surface exposed proteins known as biofilm associated protein (Bap) (Lasa and Penadés, 2006; Latasa et al., 2006). Sequence similarities, global structural organization and functional criteria of Bap-related proteins suggest that these proteins constitute a new family of surface proteins (Lasa and Penadés, 2006). A homolog to staphylococcal biofilm-associated protein (Bap) (Cucarella et al., 2001) was identified in a bloodstream isolate of *A. baumannii.* It was the first detection of a specific cell surface protein directly involved in biofilm formation by *A. baumannii*

* Corresponding author. Tel.: +98 21 51212600; fax: +98 21 51212601.
E-mail addresses: rasooli@shahed.ac.ir (I. Rasooli), Gunnar.Sandstrom@ki.se (G. Sandstrom), Jafar.amani@gmail.com (J. Amani).

0022-5193/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jtbi.2011.10.026

(Loehfelm et al., 2008). It has been suggested that $Bap_{A.\ baumannii}$ is involved in intercellular adhesion within the mature biofilm (Loehfelm et al., 2008). Bap-related proteins are present on the bacterial surface; confer upon bacteria the capacity to form a biofilm; show a high molecular weight; contain a core domain of tandem repeats; play a relevant role in bacterial infectious processes and can occasionally be contained in mobile elements (Lasa and Penadés, 2006; Latasa et al., 2006, 2005; Loehfelm et al., 2008). In the present study we used bioinformatic tools to locate a conserved specific region of $Bap_{A.\ baumannii}$ to design and develop a specific and sensitive diagnostic test for the presence of *A. baumannii*. Such a test will improve rapid clinical decisions and provide a method to epidemiologically evaluate management and control the infections caused by *A. baumannii*.

Table 1
Accession number and length of 6 biofilm associated proteins in *A. baumannii* species and the construct positions in their amino acid sequences.

Protein ID (Uniprot)	Strain	Length	Construct position
B7I652	*A. baumannii* AB0057	5464	4250–4657
B7GXL8	*A. baumannii* AB307	1315	101–508
B7GXL7	*A. baumannii* AB307	7639	7406–7634[a]
B2HXE6	*A. baumannii* ACICU	2139	388–797
B0UEF9	*A. baumannii* AYE	8200	6986–7393
B0LHN4	*A. baumannii*	8620	7406–7813

[a] The construct covers 229 amino acids of this sequence and contains three repeat modules (see the text).

2. Methods

2.1. Study design

The present study is divided into two stages. In the first stage, database search and analysis were conducted to find a conserved and specific region of Bap leading to selection of a proper region expressed in several species of *A. baumannii*. The second stage includes various analyses of the selected region such as tertiary structure, B-cell epitope prediction, antigenicity, etc.

2.2. Data sources

All sequences were obtained from Uniprot Database at http://www.uniprot.org. To find the members of Bap family among clusters, all sequences were analyzed by dot blot matrices using Jdotter (Sonnhammer and Durbin, 1995) and CLC-protein workbench, repeat modules and isoelectric points were considered. In order to find out the presence or absence of Bap across species we used Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) database of physical and functional interactions (Szklarczyk et al., 2011).

2.3. Alignments

Multiple sequence alignments were generated by MUSCLE (Edgar, 2004) at www.ebi.ac.uk/Tools/psa/emboss_needle/ using selected Bap sequences (Table 1). Alignment file was then edited in CLC Protein Work Bench. ESPript (Gouet et al., 2003) at

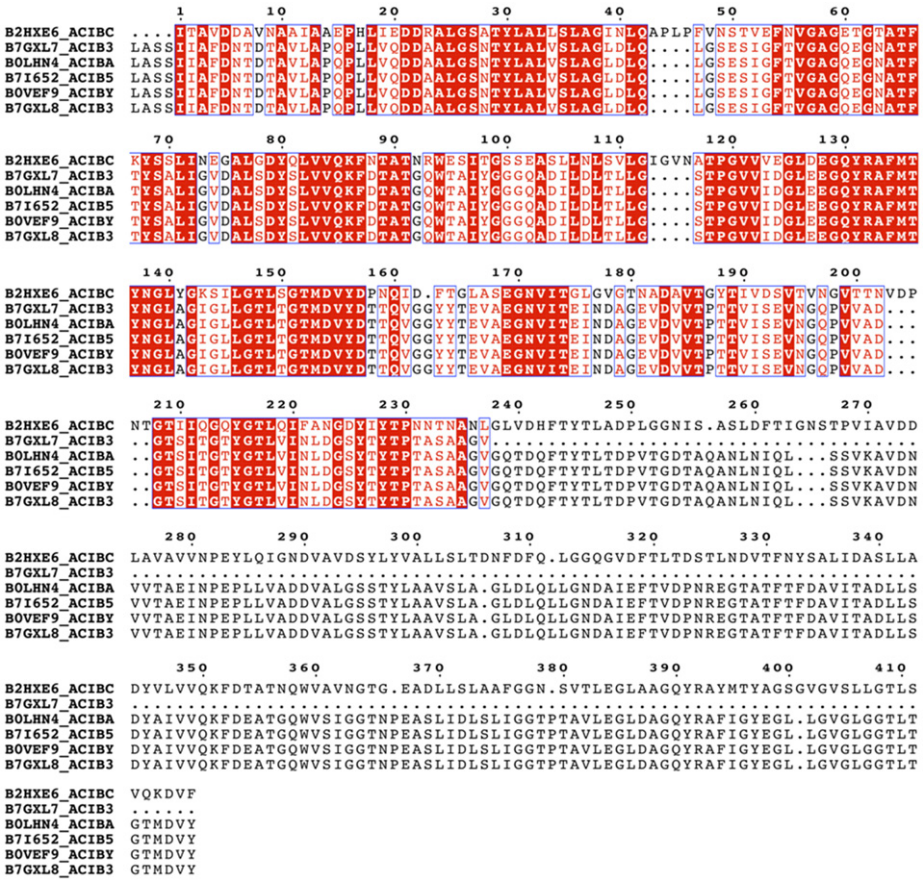

Fig. 1. Graphical display of 6 Bap multi sequence alignment. Except for (B7GXL7) sequence identity and coverage of construct is clear in the picture.

http://espript.ibcp.fr/ESPript/ESPript/ was used for presenting aligned sequences with graphical enhancements. Needle program at http://www.ebi.ac.uk/Tools/psa/emboss_needle/ was employed for pairwise comparison of sequences.

2.4. Tertiary structure prediction and validation

Various softwares such as Phyre (Gouet et al., 2003), CPHmodels (Nielsen et al., 2010) and LOOPP (Teodorescu et al., 2004) were employed to build three dimensional structures for construct. Finally the whole structures were built by ab initio modeling using I-Tasser (Roy et al., 2010) at http://zhanglab.ccmb.med.umich.edu/I-TASSER/. To recognize the errors in the generated models, coordinates were supplied by uploading 3D structures in PDB format into ProSA-web, which is frequently employed in protein structure validation (Wiederstein and Sippl, 2007) at https://prosa.services.came.sbg.ac.at/prosa.php.

2.5. Localizing and quantifying the energetic frustration

The degree of local frustration manifested in protein molecules was predicted using Protein Frustratometer Server at http://bioinf.qb.fcen.uba.ar/frustra/ for gaining insight to the proteins' biological behavior by analyzing how the energy is distributed in predicted structures.

2.6. Ligand binding sites

The I-Tasser software automatically defined structural analogs with binding sites similar to the built model of I-Tasser. 3DligandSite (Wass and Sternberg, 2009) at http://www.sbg.bio.ic.ac.uk/~3dligandsite was another software employed for predicting the binding sites on 3D structures.

2.7. Immunoinformatic assay

2.7.1. Primary sequence analysis

In order to find differences in propensity scales of Bap sequences, several parameters including hydrophilicity (Black and Mould, 1991), flexibility (Bhaskaran and Ponnuswamy, 1988), turns (Levitt, 1978), polarity and antigenic propensity (Kolaskar and Tongaonkar, 1990) of Bap sequences were assessed at www.expasy.ch/protscale (Gasteiger et al., 2005).

2.7.2. Predicting the discontinuous and continuous B-cell epitopes

Linear B-cell epitopes of construct were predicted using Ellipro. Briefly, this method was performed based upon solvent-accessibility and flexibility when the PDB file of 3D structure uploaded to the software. A combination of hidden Markov model (HMM) with two best propensity scale methods (Parker and Levitt) (Levitt, 1978; Parker et al., 1986) was employed for predicting linear B-cell epitopes using BepiPred (Larsen et al., 2006) the threshold sets as 0.35. The application of this method to a large number of proteins allowed an

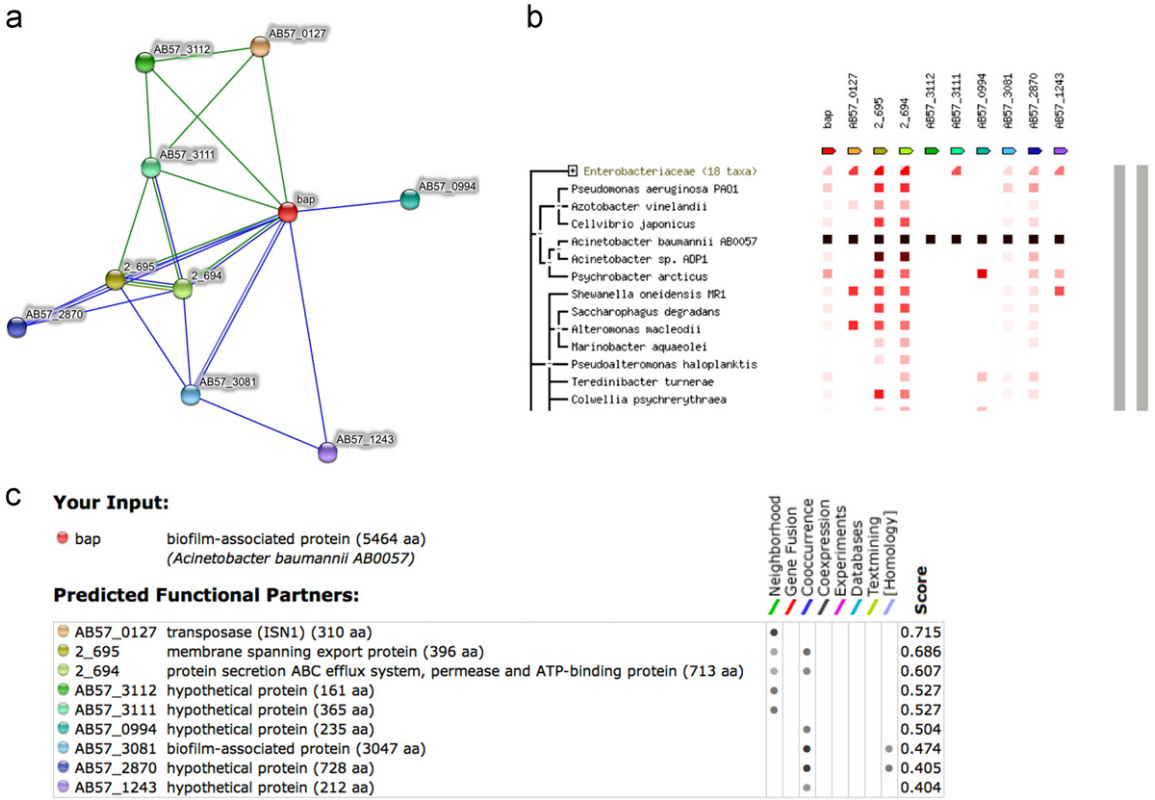

Fig. 2. Sample of protein network visualization on the STRING website. The interaction network for genetically interacting proteins possibly related in function with *A. baumannii* Bap is shown. The network nodes are proteins. Edge lines represent the existence of types of evidence used in predicting the associations. A red line indicates the presence of fusion evidence; a green line—neighborhood evidence; a blue line—co-occurrence evidence; a purple line—experimental evidence; a yellow line—textmining evidence; a light blue line—database evidence. (b) The occurrence view shows the presence or absence of linked proteins across species. Proteins are listed across the top of the page and a phylogenetic tree with species names is listed down the left hand side. In the subsequent grid, the presence of the protein in a species is marked with a red square and absence with a white space. The intensity of the color of the red square reflects the amount of conservation of the homologous protein in the specie. (c) Summary view of predicted associations, sorted by score. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

accuracy of 75%. The degree of conservancy of linear epitopes was calculated within PSI-BLAST (Altschul et al., 1997) extracted protein sequence set at 80% of sequence identity (when construct served as a query for this BLAST search). The degree of conservation is defined as the fraction of protein sequences containing the epitope at an 80% identity level. Discontinuous B-cell epitopes were predicted using 3D structure by Ellipro using default software parameters (minimum score of 0.5 and maximum distance of 6 Å). All these approaches were performed at www.immuneepitope.org.

2.7.3. Mapping the epitopes on the protein surface

The conformational epitopes on the protein surface were predicted by an automated sequence analysis of all phage display sequences and a comparison to the distribution of amino acids on 3D patches on the protein surface. In short, linear B-cell sequences obtained from Ellipro software, were submitted into Episearch software at http://curie.utmb.edu/episearch.html as well as PDB file of 3D structure. The software compared and quantified the amino acid composition of the linear and 3D profiles in a score function for each patch on the protein surface and the software defines the location of conformational epitopes correctly (Negi and Braun, 2009).

2.7.4. Important properties of construct

Protein solubility was evaluated using recombinant protein solubility prediction at www.biotech.ou.edu (Davis et al., 1999; Harrison, 2000; Wilkinson and Harrison, 1991). Vaxijen (Doytchinova and

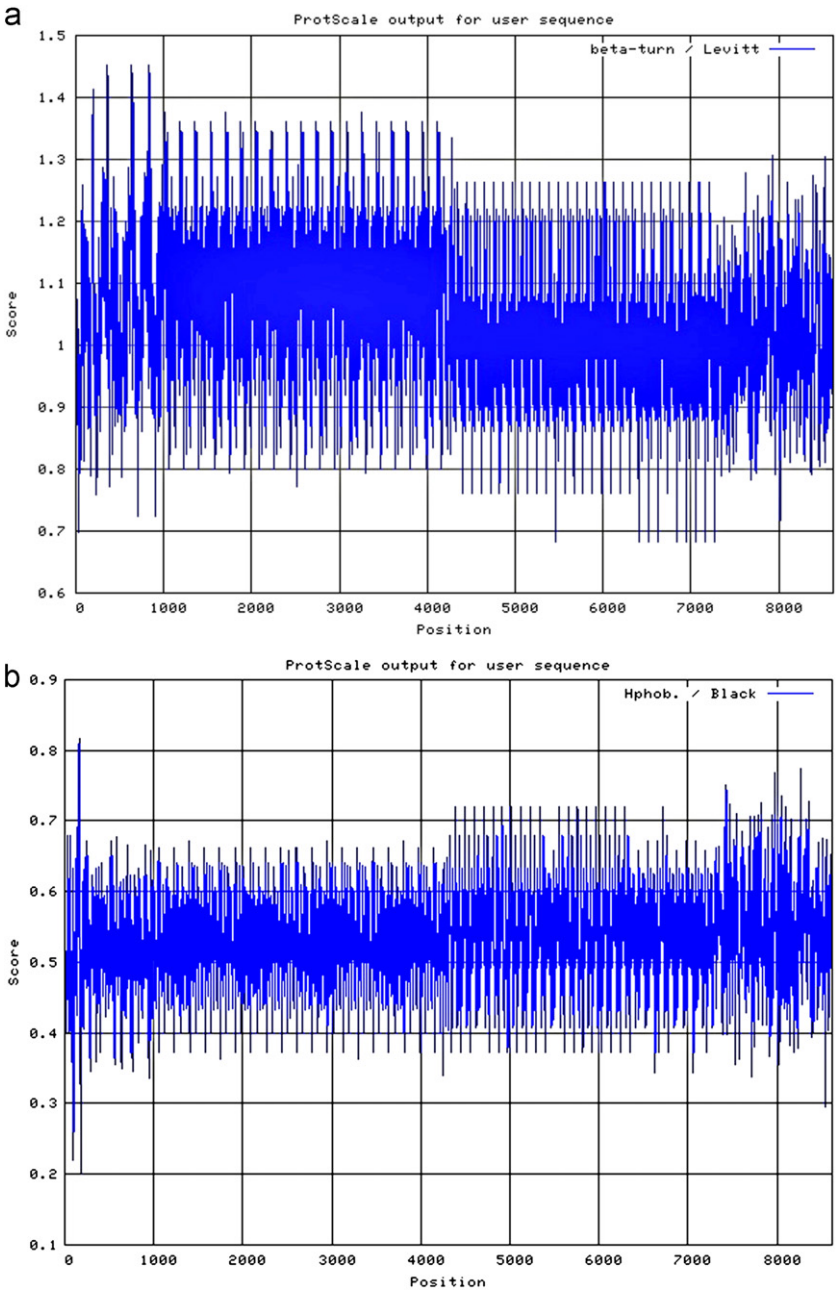

Fig. 3. Samples of propensity scales for one of the longest Bap sequences (Uniprot Acc. no. B0LNH4). (a) Beta turn prediction (a) beta-turn/Levitt, the individual values for the 20 amino acids are: Ala: 0.770; Arg: 0.880; Asn: 1.280; Asp: 1.410; Cys: 0.810; Gln: 0.980; Glu: 0.990; Gly: 1.640; His: 0.680; Ile: 0.510; Leu: 0.580; Lys: 0.960; Met: 0.410; Phe: 0.590; Pro: 1.910; Ser: 1.320; Thr: 1.040; Trp: 0.760; Tyr: 1.050; Val: 0.470, 1.345, 0.985, 0.952. (b) Hydrophobicity using the scale Hphob./Black, the individual values for the 20 amino acids are: Ala: 0.616; Arg: 0.000; Asn: 0.236; Asp: 0.028; Cys: 0.680; Gln: 0.251; Glu: 0.043; Gly: 0.501; His: 0.165; Ile: 0.943; Leu: 0.943; Lys: 0.283; Met: 0.738; Phe: 1.000; Pro: 0.711; Ser: 0.359; Thr: 0.450; Trp: 0.878; Tyr: 0.880; Val: 0.825, 0.132, 0.147, 0.526.

Flower, 2007) software was used for estimating the probability of antigenicity of construct at http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html. This software allows users to assess a protein's ability to induce protection. The server deals with single proteins as well as whole proteomes submitted in fasta format. The leave-one-out cross-validation (LOO-CV) of the software has 82% accuracy, 91% sensitivity and 72% specificity.

2.8. Validating the method

In the present study we have used The UniRef databases that provide clustered sets of sequences from UniProt Knowledgebase (including splice variants and isoforms) and selected UniParc records, in order to obtain complete coverage of sequence space at several resolutions while hiding redundant sequences (but not their descriptions) from view (Suzek et al., 2007). We have also used Version 9.0 of STRING with high confidence (0.70), which covers more than 1100 completely sequenced organisms (Szklarczyk et al., 2011). Some important parts of the present study was specially based on 3D structure of the construct. The iterative threading assembly refinement (I-TASSER), an integrated platform for automated protein structure and function prediction based on the sequence-to-structure-to-function paradigm, was employed for this purpose. Since determining the 3D structure of proteins is a major biological problem (Roy et al., 2010) and the biological usefulness of the predicted protein models relies on the accuracy of the structure prediction (Zhang, 2009). Z-scores and RMSD as well as TM-score helped us to quantitatively and qualitatively estimate the accuracy of 3D model. Software parameters were specified precisely to achieve the best and reliable results.

3. Results

3.1. Primary sequence analysis

Sequences are belonging to cluster of 50% identity to biofilm-associated protein (Acc. no. UniRef50_B7I652). The favorite sequences are listed in Table 1. These sequences serve all criteria to be a member of Bap family. They are large proteins with tandem repeat modules and low isoelectric points. Alignment of these proteins revealed the conservation of several locations in sequences, mainly repeat modules. Within these conserved regions, a specific region was selected based on PSI-BLAST search against non-redundant protein database (Fig. 1). Majority of PSI-BLAST hits corresponded to *A. baumannii* when the selected region was served as a query. Minority of sequences with lower score and higher *e*-value belonged to *Salmonella enterica.* Co-occurrence with biofilm associated proteins of protein secretion ABC efflux system, permease and ATP-binding protein (Uniprot ACC no. B7IBV0) across several *A. baumannii* genomes is notable (Fig. 2b). Sample of STRING network view of protein interaction is presented in Fig. 2a. Propensity plots were indicative of a long region of Bap that vastly differs from whole protein sequences in all scales viz. hydrophilicity, flexibility, accessibility, beta-turns, exposed surface, polarity and antigenicity. Plot samples are illustrated in Fig. 3.

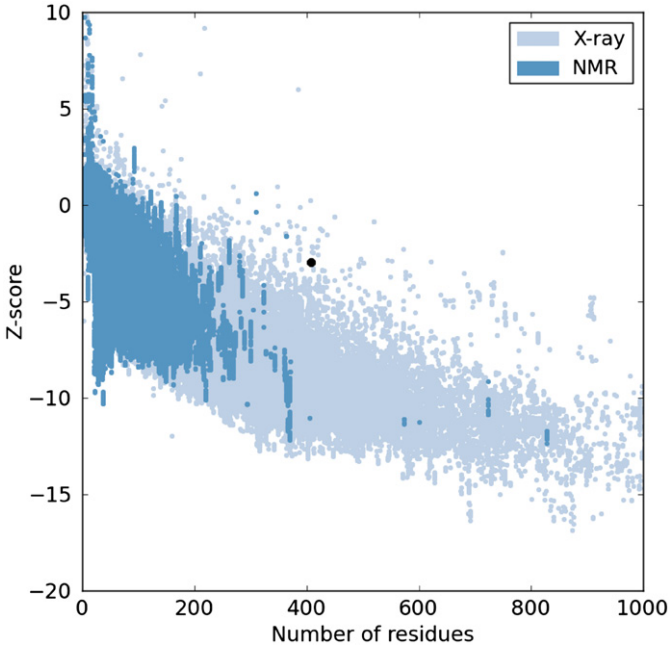

Fig. 5. Z-score plot for tertiary structure of construct. The z-score indicates overall model quality. Its value is displayed in a plot that contains the z-scores of all experimentally determined protein chains in current PDB. In this plot, groups of structures from different sources (X-ray, NMR) are distinguished by different colors. It can be used to check whether the z-score of the input structure is within the range of scores typically found for native proteins of similar size.

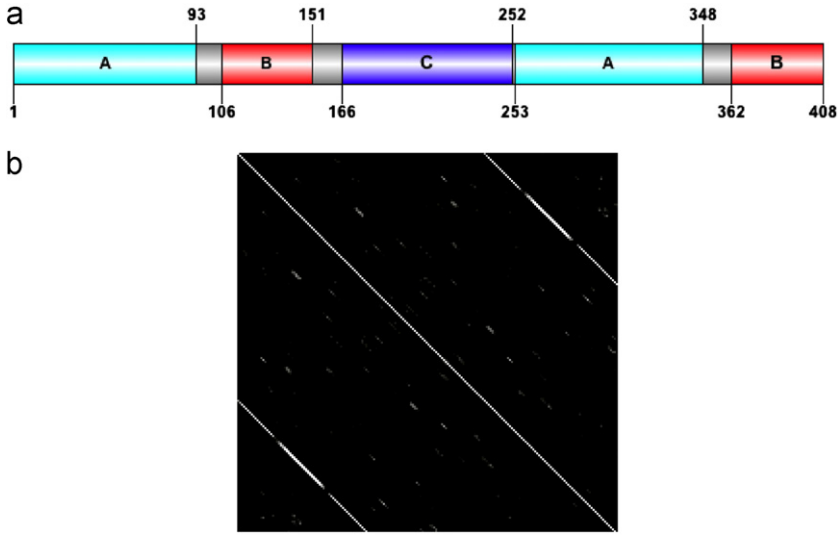

Fig. 4. Graphical display of construct and its repeat containing modules. The length and position of repeat modules are graphically illustrated on a gray rod (a). Three repeat units are clearly evident in dot blot matrix (b) horizontal and vertical axes are both construct sequence.

The percentage identities determination of the Bap proteins of *A. baumannii* strains and construct was evaluated using the Needle program, the construct sequence, used as the reference sequence. The construct was found to share 100% identity with the B0LHN4-ACIB, B7GXL8, B0VEF9, B7I652; 56% identity and 56% similarity with B7GXL7; 48% identity and 69% similarity with B2HXE6.

3.2. Appropriate region selection

Based on alignment and primary sequence analyses and distribution of antigenic determinant, an appropriate region that serves our purpose was selected and named as construct. The corresponding position of this region to Bap sequences is shown in Table 1. A graphical display of the construct is presented in Fig. 4a.

3.3. Tertiary structures

No significant hit was observed when construct served as a query for BLAST search against Protein Data Bank proteins. No template protein of similar folds (or super-secondary structures) was retrieved from the PDB library by LOMETS (Wu and Zhang, 2007), a locally installed meta-threading approach. Hence, threading and homology modeling would completely fail for building the 3D structure for the construct. In contrast ab initio approach was found to be the best way of constructing 3D structure. All tertiary structures were analyzed for overall model quality estimation. The z-score of the model built by I-TASSER for selected region was higher than the others (Fig. 5). The z-score of the input structure was within the range of scores typically found for native proteins of similar size. The confidence score (C-score) for estimating the quality of predicted models by I-TASSER was -1.9. (C-score is typically in the range of $[-5$ to 2], where a C-score of higher value signifies the model with a high confidence.) Also the expected TM-score for this model was 0.48 ± 0.15. The expected RMSD was 11.5 ± 4.5.

3.4. Localizing and quantifying the energetic frustration

The highest degree of frustration is located at small stretches of amino acids particularly at position 90–150 and C-terminal of the construct. These highly frustrated domains are strongly cross-linked by minimally frustrated contact networks (Fig. 6a).

3.5. Protein–ligand binding site predictions by global structure match and local geometry

Predicted binding sites ranked based on the size of the clusters formed after local superposition of ligands of the templates onto the query structure. Binding sites of higher confidence are presented in

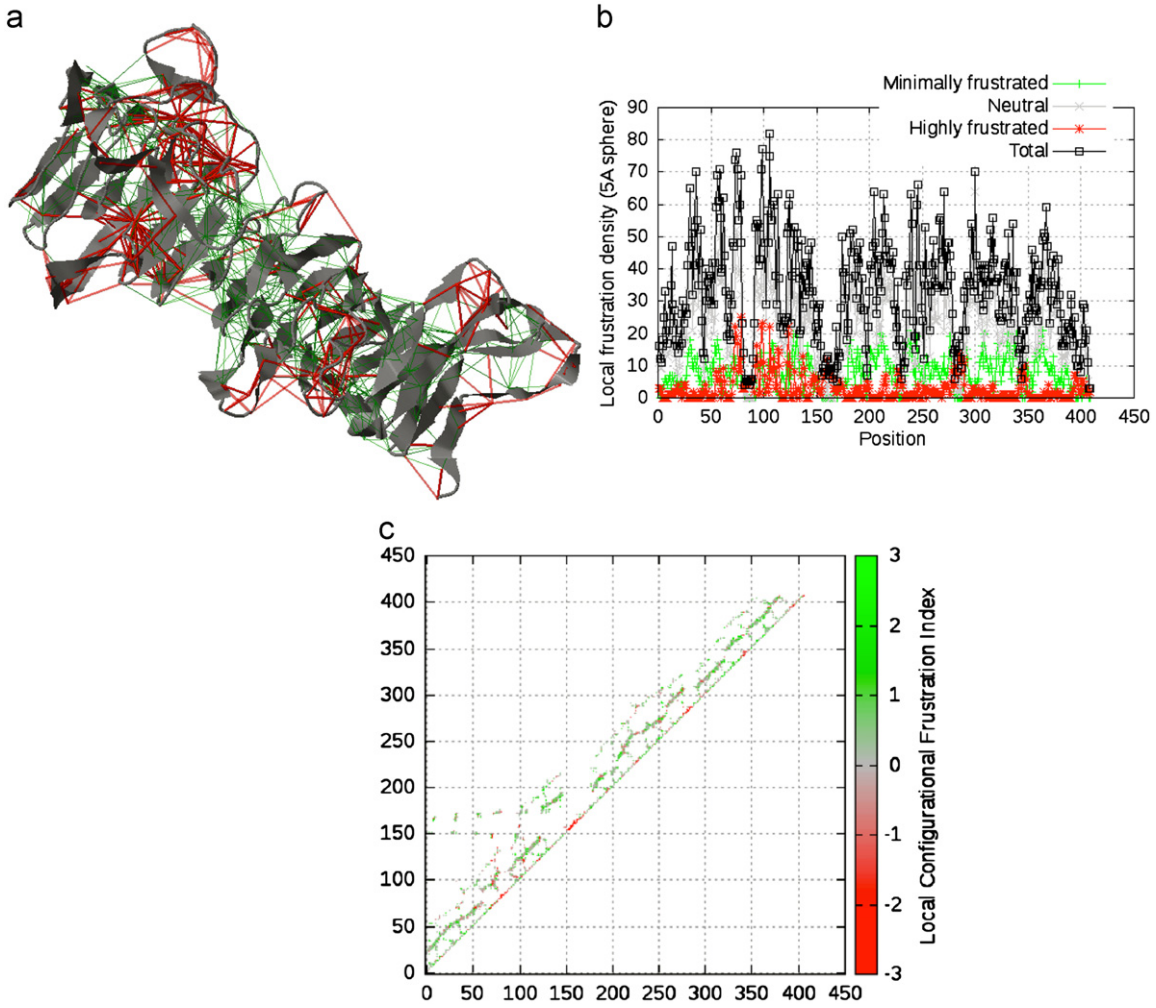

Fig. 6. Localization of frustration of protein structure. Cartoonish display of the localized frustration and minimally frustrated networks in construct structure (a). The protein backbone is displayed as blue ribbons, the direct minimally frustrated interactions are shown in green lines, and highly frustrated contacts in red lines, neutral contacts are not drawn. A 'contact map' of residues (b), each dot represents pair-interactions between amino acids, numbered on the axis. Projections of the contact information on the sequence space (c). Colors are in accordance with their frustration index. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Fig. 7 and the details are given in Table 2. Based on large scale benchmarking analysis, binding site predictions with BS-score, a measure of local sequence and structure similarity between template's binding site and predicted binding site in the query structure, more than 1.1 signify predictions with high confidence. Due to lack of sufficient homologous structures, 3DLigandSite could not predict corresponding binding sites and ligands.

3.6. Immunoinformatic assay

15 potential linear B-cell epitopes were predicted for primary sequence of construct (Table 3). These linear epitopes were conserved within BLAST sequences and mostly belong to *A. baumannii* (Table 4). Analyzing the 3D structure also represented 10 linear B-cell epitopes (Table 5) and three sets of discontinuous B-cell epitopes containing 95 (Score: 0.7), 121 (Score: 0.66) and 34 (Score: 0.63) residues. The set of highest score is listed in Table 6. When the 10 linear antigenic determinants were mapped on protein surface by EpiSearch, high scoring patches were predicted in six locations. The patch with the highest score (8.7) was found with the center at residue F130 (Fig. 8a), and the second highest score was found centered at residue L366 (Score: 8.69) (Fig. 8b). Patches at the third, fourth, fifth and sixth rank were found at L35 with a score 8.460 (Fig. 8c), at A261 (Score: 7.77) (Fig. 8d), at T186 (Score: 7.7) (Fig. 8e), and finally at I188 with score of 7.65 (Fig. 8f).

3.7. Other properties of construct

The construct sequence has an 88.4% chance of solubility when over expressed in *Escherichia coli*. The overall prediction for the antigenicity of the construct was 0.68.

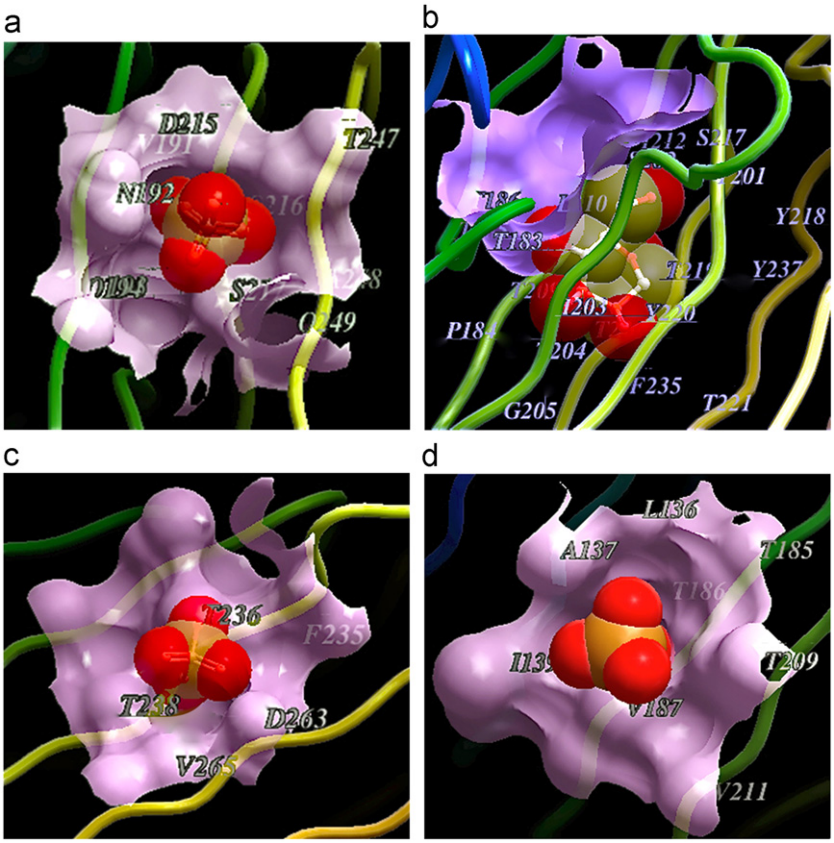

Fig. 7. Graphical representation of interaction of ligands and construct. Protein surface is colored in light purple, corresponding ligands are shown in space fill format, and near contact residues are labeled. Ligands in figure subparts are: (a) PO_4, (b) BMA, (c) POL and (d) PO_4. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2
Predicted binding sites and identified structural analogs with similar binding site.

PDB Hit	TM-score	RMSD[a]	Cov.[b]	BS-score[c]	Lig. name[d]	Predicted binding site residues in the model
2vbeA	0.6131	4.4700	0.7917	1.4420	PO_4	215,216,217
1rmgA	0.5382	4.3500	0.6789	1.3481	BMA	200,201,217,218,219
1oflA	0.5696	4.4700	0.7304	1.0525	POL	220,221,255,256,257,277,278,279
2vbkA	0.6132	4.4600	0.7917	1.4966	PO4	138,186,187

[a] The RMSD between residues that are structurally aligned by TM-align.
[b] The coverage of the alignment by TM-align and is equal to the number of structurally aligned residues divided by length of the model.
[c] BS-score is a measure of local sequence and structure similarity between template's binding site and predicted binding site in the query structure. Based on large scale benchmarking analysis, binding site predictions with BS-score > 1.1 signify predictions with high confidence.
[d] PO_4: phosphate, BMA: beta-D-mannose, POL: *N*-Propanol.

Table 3
Linear B-cell epitopes predicted based on combination of HMM and propensity scale from amino acid sequence of construct.

Peptide length	Peptide	End position	Start position	No.
6	TDTAVL	16	11	1
4	AALG	30	27	2
9	GAGQEGNAT	65	57	3
16	DTATGQWTAIYGGGQA	103	88	4
4	LEEG	125	122	5
17	DVYDTTQVGGYYTEVAE	166	150	6
10	NDAGEVDVVT	183	174	7
20	TVISEVNGQPVVADGTSITG	205	186	8
18	SYTYTPTASAAGVGQTDQ	234	217	9
11	TDPVTGDTAQA	250	240	10
6	TAEINP	272	267	11
11	TVDPNREGTAT	320	310	12
17	DEATGQWVSIGGTNPEA	359	343	13
6	GTPTAV	374	369	14
3	DAG	381	379	15

4. Discussion

Bioinformatics tools are now standard and promising method in the choice of specific and immunogenic epitopes. In silico epitope mapping, combined with in vitro and in vivo verification, accelerates the discovery process by approximately 10-20-fold (Amani et al., 2009). The benefit of in silico studies in clinical trial could provide profound insight into the design of clinical trials (Clermont et al., 2004), ranging from vaccine design (Davies and Flower, 2007; De Groot et al., 2001; Jahangiri et al., 2011; Korber et al., 2006) to find cancer or infectious diseases diagnostic agents (Aagaard et al., 2003; Bannantine et al., 2002; Leerkes et al., 2002; Roukos, 2009). Zhang et al. (2002) identified several intertype specific epitopes on a human adenovirus antigen using genomic alignment tools, antigenicity and 3-D structure prediction programs. Most of the predicted epitopes were originated in the prepared synthetic peptides and recombinant proteins. In another study Olfa et al. (2008) in silico analyzed the OmcB protein of *Chlamydia tracomatis* in order to find specific and immunogenic

Table 4
Predicted linear B-cell epitopes for 3D structure of construct.

Score	Number of residues	Peptide	End position	Start position	No.
0.769	27	TMDVYDTTQVGGYYTEVAEGNVITEIN	174	148	1
0.759	70	VQKFDEATGQWVSIGGTNPEASLIDLSLIGGTPTAVLEGLDAGQYRAFIGYEGLLGVGLGGTLTGTMDVY	408	339	2
0.744	14	ADDVALGSSTYLAA	291	278	3
0.727	36	LASSIIAFDNTDTAVLAPQPLLVQDDAALGSNTYLA	36	1	4
0.713	14	DYSLVVQKFDTATG	92	79	5
0.691	11	TASAAGVGQTD	233	223	6
0.646	18	LAGLDLQLGSESIGFTVG	57	40	7
0.623	6	DLLSDY	335	330	8
0.504	6	TDPVTG	245	240	9
0.503	12	GNATFTYSALIG	73	62	10

Table 5
Epitope conservancy among 36 BALST extracted sequences at identity level of $\geq 80\%$.

Maximum identity (%)	Minimum identity (%)	Percent of protein sequence matches at identity $\geq 80\%$	Epitope length	Epitope sequence	Epitope no.
100	66.67	94.44% (34/36)	6	TDTAVL	1
100	75.00	66.67% (24/36)	4	AALG	2
100	44.44	47.22% (17/36)	9	GAGQEGNAT	3
100	31.25	25.00% (9/36)	16	DTATGQWTAIYGGGQA	4
100	75.00	30.56% (11/36)	4	LEEG	5
100	29.41	27.78% (10/36)	17	DVYDTTQVGGYYTEVAE	6
100	40.00	19.44% (7/36)	10	NDAGEVDVVT	7
100	35.00	19.44% (7/36)	20	TVISEVNGQPVVADGTSITG	8
100	33.33	25.00% (9/36)	18	SYTYTPTASAAGVGQTDQ	9
100	36.36	27.78% (10/36)	11	TDPVTGDTAQA	10
100	50.00	30.56% (11/36)	6	TAEINP	11
100	45.45	27.78% (10/36)	11	TVDPNREGTAT	12
100	35.29	27.78% (10/36)	17	DEATGQWVSIGGTNPEA	13
100	66.67	27.78% (10/36)	6	GTPTAV	14
100	100.00	100.00% (36/36)	3	DAG	15

Table 6
Discontinuous B-cell epitopes of higher score, predicted based on 3D structure.

Score	Number of residues	Residues
0.691	95	L214, T240, D241, P242, V243, T244, G245, D246, T267, A268, E269, I270, D298, L299, R315, E316, G317, T320, D324, I327, D330, L331, L332, S333, D334, Y335, V339, Q340, K341, F342, D343, E344, A345, T346, G347, Q348, W349, V350, I352, G353, G354, T355, N356, P357, E358, A359, S360, L361, I362, D363, L364, S365, L366, I367, G368, G369, T370, P371, T372, A373, V374, L375, E376, G377, L378, D379, A380, G381, Q382, Y383, R384, A385, F386, I387, G388, Y389, E390, G391, L392, L393, G394, V395, G396, L397, G398, G399, T400, L401, T402, G403, T404, M405, D406, V407, Y408

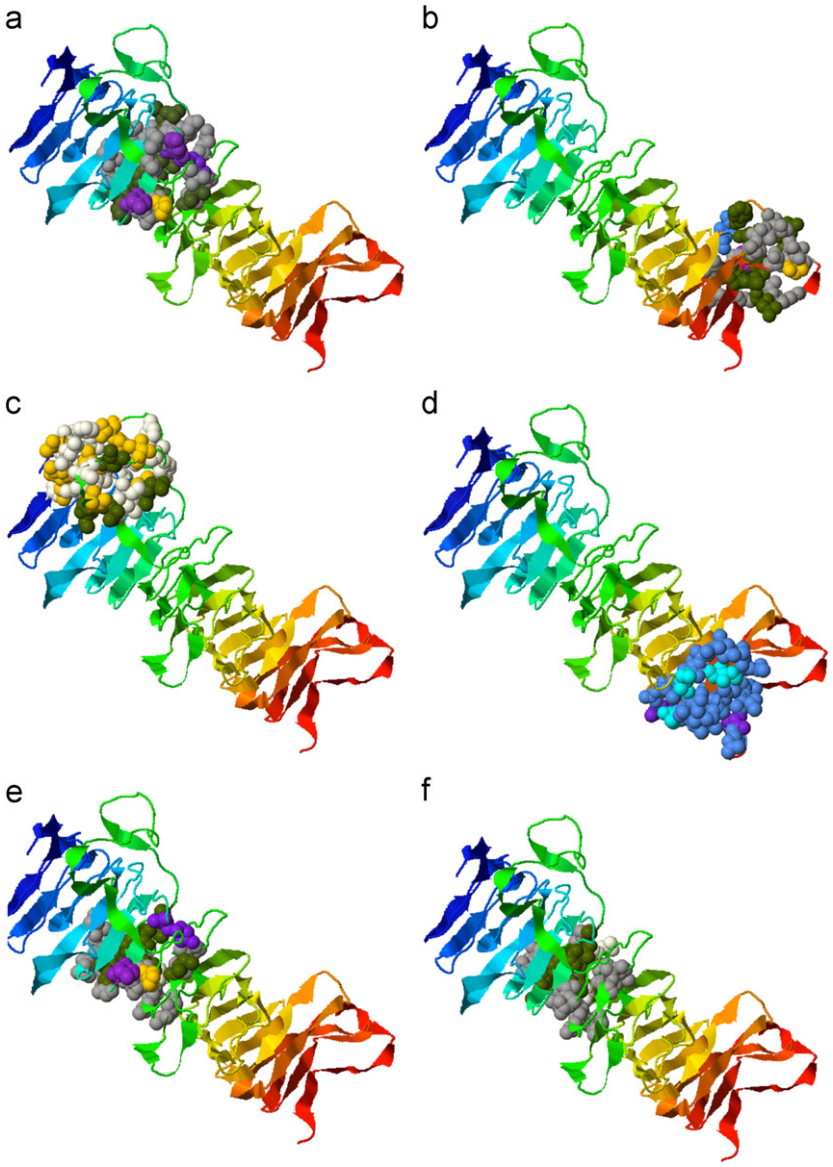

Fig. 8. Localization of potential B-cell epitopes on the protein structure.

antigen for the serodiagnosis of *C. trachomatis* infections. They validated their in silico finding by performing ELISA tests with several clinical specimens; their results indicated that the use of bioinformatic tools might be a useful method for identifying species specific regions in an immunodominant antigen. The species-specific probes and the DNA microarray-based assay system were developed for the detection of nosocomial *A. baumannii* and *P. aeruginosa* strains from a variety of clinical specimens (Keum et al., 2006). More recently, outer membrane proteins (OMPs) were introduced as prime candidates for recognition by host antibodies (Islam et al., 2011). A new, sensitive, precise, and cost-effective test is required for the fast detection of *A. baumannii* infections. Development of such a test would require finding specific antigenic proteins of *A. baumannii*. Biofilm associated proteins are widespread in bacteria. The most important feature of these proteins is their participation in the biofilm formation process (Lasa and Penadés, 2006). Some members of this group of proteins: Bap (Cucarella et al., 2001), Esp (Shankar et al., 1999), LapA (Hinsa et al., 2003), BapA (Latasa et al., 2005), $Bap_{A.\ baumannii}$ (Loehfelm et al., 2008; Rahbar et al., 2010) are well characterized. Homologous proteins to Bap can be found in several currently available genomes of *A. baumannii*. Existence of surface proteins related to Bap in several *A. baumannii* strains genomes (Fig. 2b) makes it a proper candidate for our purpose in finding a diagnostic test agent. Higher scores and the lowest BLAST hits for this selection indicate the specificity of the region to *A. baumannii*. In contrast, homologs to other regions exist in other organisms. The conservancy of construct is clearly evident in Fig. 1 in which one sequence (B7GXL7) does not cover the whole length of the construct while the remainder shows high identity. As previously mentioned (Rahbar et al., 2010), number and arrangement of repeat units are the main differences between Bap homologs. These modules of Bap proteins are important functional constituents (Rahbar et al., 2010). The need for maintaining functionally associated genes together can become visible as an agreement in occurrence-patterns across several genomes (Huynen and Bork, 1998; Pellegrini et al., 1999). Co-occurrence of protein secretion, ABC efflux system, permease and ATP-binding protein (Uniprot ACC no. B7IBV0) with biofilm associated proteins across several *A. baumannii* genomes is notable (Fig. 2b and c) suggesting the role of efflux system in biofilm specific resistance of this organism. There is sufficient

evidence that expression of efflux pumps in biofilms might lead to antibiotic resistance in these surface-attached communities (De Kievit et al., 2001). Higher expression of efflux components in biofilm phenotype is well documented in other organisms (Lynch et al., 2007; Zhang and Mah, 2008). The identification of conformational epitopes in antibody–antigen interaction is a crucial step for the rational design of novel drugs and vaccines (Negi and Braun, 2009). This requirement necessitated the construction of tertiary structure for desired construct. Expected TM-score (Zhang and Skolnick, 2004) of 0.48 ± 0.15 validates the accuracy of the 3D model (a TM-score > 0.5 indicates a model of correct topology). Other scores including z-score and C-score suggest its confidence. Since proteins in native state, comprise 40% of minimally frustrated, 45% of neutral, and only a small fraction of highly frustrated (Ferreiro et al., 2007), therefore localization of highly frustrated region in 3D model is another support for validity of the predicted structure and in vivo stability of the construct. Presence of some energetic frustration in a folded protein is not ruled out (Ferreiro et al., 2007) and is in agreement with our prediction of localization of conformational epitopes in highly frustrated regions (Fig. 8). Highly frustrated regions often correspond to sites that bind other macromolecules, in this case antibodies, or small ligands (Ferreiro et al., 2007). Ligand binding sites are matched with frustrated residues (Figs. 6b and 7). Parameters such as hydrophilicity, flexibility, accessibility, beta-turns, exposed surface, polarity and antigenic propensity of polypeptides chains have been correlated with the location of continuous epitopes (Larsen et al., 2006; Pellequer et al., 1991). Identification of such linear peptide segments will often be the initial step in the search for antigenic determinants in pathogenic organisms (Larsen et al., 2006). Fifteen potential linear B-cell epitopes were predicted by Bepipred software with sensitivity of 0.49 and specificity of 0.75, as the threshold value was set as 0.35. Conservancy of predicted linear B-cell epitopes among PSI-BLAST extracted sequences and their relevance to *A. baumannii* is another support for specificity of these epitopes to the pathogen (Table 5), conservancy of epitopes allowed diminishing the extent of probable cross reactions of antibodies raise against the construct, which may lower the specificity of serodiagnostic test. Location of predicted B-cell epitopes on several positions of the construct (Fig. 8) is identifiable by the antibodies against our construct in any epitope presenting pattern. While homologous proteins from different species may have a high degree of sequence identity, they have markedly different epitope presentations (Brady et al., 2007). In recent years the problem of *A. baumannii* grows significantly in clinical practice owing to its resistance to all commonly used Gram-negative antimicrobials (Falagas et al., 2006; Morgan et al., 2009; Rossolini and Mantengoli, 2008). Conventional time consuming and costly culture methods or biochemical tests are still used for detection of this organism (Islam et al., 2011). Its rapid emergence and prevalence is indicative of a vital need for a diagnostic test that is more rapid, sensitive and specific compared to conventional detection methods. Antigenicity of the region could also be explored for elicitation of antibody as a prophylactic measure in individuals exposed to *A. baumannii*. The construct could also serve for production of a recombinant vaccine candidate. Our construct fulfills these needs. Laboratory tests with the serum of infected patients are in progress in our research laboratory for practical confirmation of the predictions.

5. Conclusion

A construct serving as a potential agent for diagnostic test based on antibody–antigen interaction is introduced herein. The antigen is predicted to be soluble, possessing several sites for antibody binding.

Acknowledgment

The authors wish to thank Shahed University for the sanction of grants to conduct the present study. We also declare no conflict of interests.

References

Aagaard, C., Govaerts, M., Meng Okkels, L., Andersen, P., Pollock, J.M., 2003. Genomic approach to identification of *Mycobacterium bovis* diagnostic antigens in cattle. J. Clin. Microbiol. 41, 3719.

Akers, K.S., Chaney, C., Barsoumian, A., Beckius, M., Zera, W., Yu, X., Guymon, C., Keen III, E.F., Robinson, B.J., Mende, K., 2010. Aminoglycoside resistance and susceptibility testing errors in *Acinetobacter baumannii*–calcoaceticus complex. J. Clin. Microbiol. 48, 1132.

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang2, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402.

Amani, J., Mousavi, S.L., Rafati, S., Salmanian, A.H., 2009. In silico analysis of chimeric espA, eae and tir fragments of *Escherichia coli* O157:H7 for oral immunogenic applications. Theor. Biol. Med. Model. 6, 28.

Bannantine, J.P., Baechler, E., Zhang, Q., Li, L.L., Kapur, V., 2002. Genome scale comparison of *Mycobacterium avium* subsp. *paratuberculosis* with *Mycobacterium avium* subsp. *avium* reveals potential diagnostic sequences. J. Clin. Microbiol. 40, 1303.

Bhaskaran, R., Ponnuswamy, P., 1988. Positional flexibilities of amino acid residues in globular proteins. Int. J. Peptide Protein Res. 32, 241–255.

Black, S., Mould, D., 1991. Amino acid scale: hydrophobicity of physiological L-alpha amino acids. Anal. Biochem. 193, 72–82.

Brady, R.A., Leid, J.G., Kofonow, J., Costerton, J.W., Shirtliff, M.E., 2007. Immunoglobulins to surface-associated biofilm immunogens provide a novel means of visualization of methicillin-resistant *Staphylococcus aureus* biofilms. Appl. Environ. Microbiol. 73, 6612–6619.

Cevahir, N., Demir, M., Kaleli, I., Gurbuz, M., Tikvesli, S., 2008. Evaluation of biofilm production, gelatinase activity, and mannose-resistant hemagglutination in *Acinetobacter baumannii* strains. J. Microbiol. Immunol. Infect. 41, 513–518.

Clermont, G., Bartels, J., Kumar, R., Constantine, G., Vodovotz, Y., Chow, C., 2004. In silico design of clinical trials: a method coming of age. Crit. Care Med. 32, 2061.

Cucarella, C., Solano, C., Valle, J., Amorena, B., Lasa, I., Penades, J., 2001. Bap, a *Staphylococcus aureus* surface protein involved in biofilm formation. J. Bacteriol. 183, 2888.

Davies, M.N., Flower, D.R., 2007. Harnessing bioinformatics to discover new vaccines. Drug Discovery Today 12, 389–395.

Davis, G.D., Elisee, C., Newham, D.M., Harrison, R.G., 1999. New fusion protein systems designed to give soluble expression in *Escherichia coli*. Biotechnol. Bioeng. 65, 382–388.

De Groot, A.S., Bosma, A., Chinai, N., Frost, J., Jesdale, B.M., Gonzalez, M.A., Martin, W., Saint-Aubin, C., 2001. From genome to vaccine: in silico predictions. ex vivo verification. Vaccine 19, 4385–4395.

De Kievit, T.R., Parkins, M.D., Gillis, R.J., Srikumar, R., Ceri, H., Poole, K., Iglewski, B.H., Storey, D.G., 2001. Multidrug efflux pumps: expression patterns and contribution to antibiotic resistance in *Pseudomonas aeruginosa* biofilms. Antimicrob. Agents Chemother. 45, 1761.

Doytchinova, I., Flower, D., 2007. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. BMC Bioinformatics 8, 4.

Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res., 32.

Elston, J.W.T., Bannan, C.L., Chih, D.T., Boutlis, C.S., 2008. *Acinetobacter* spp. in gunshot injuries. Emerging Infect. Dis. 14, 178.

Falagas, M.E., Koletsi, P.K., Bliziotis, I.A., 2006. The diversity of definitions of multidrug-resistant (MDR) and pandrug-resistant (PDR) *Acinetobacter baumannii* and *Pseudomonas aeruginosa*. J. Med. Microbiol. 55, 1619–1629.

Ferreiro, D.U., Hegler, J.A., Komives, E.A., Wolynes, P.G., 2007. Localizing frustration in native proteins and protein assemblies. Proc. Natl. Acad. Sci. 104, 19819.

Fournier, P., Richet, H., 2006. The epidemiology and control of *Acinetobacter baumannii* in health care facilities. Clin. Infect. Dis. 42, 692–699.

Gaddy, J.A., Actis, L.A., 2009. Regulation of *Acinetobacter baumannii* biofilm formation. Future Microbiol. 4, 273–278.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M., Appel, R., Bairoch, A., 2005. Protein identification and analysis tools on the expasy server. in: Walker, John M. (Ed.), The Proteomics Protocols Handbook, Humana Press, Full text-Copyright Humana Press.

Gordon, N.C., Wareham, D.W., 2010. Multidrug-resistant *Acinetobacter baumannii*: mechanisms of virulence and resistance. Int. J. Antimicrob. Agents 35, 219–226.

Gouet, P., Robert, X., Courcelle, E., 2003. ESPript/ENDscript: extracting and rendering sequence and 3D information from atomic structures of proteins. Nucleic Acids Res. 31, 3320.

Hamouda, A., Amyes, S.G.B., 2004. Novel gyrA and parC point mutations in two strains of *Acinetobacter baumannii* resistant to ciprofloxacin. J. Antimicrob. Chemother. 54, 695.

Harrison, R.G., 2000. Expression of soluble heterologous Proteins via fusion with NusA protein. inNovation 11, 4–7.

Hinsa, S.M., Espinosa-Urgel, M., Ramos, J.L., O'Toole, G.A., 2003. Transition from reversible to irreversible attachment during biofilm formation by *Pseudomonas fluorescens* WCS365 requires an ABC transporter and a large secreted protein. Mol. Microbiol. 49, 905–918.

Hujer, K.M., Hujer, A.M., Hulten, E.A., Bajaksouzian, S., Adams, J.M., Donskey, C.J., Ecker, D.J., Massire, C., Eshoo, M.W., Sampath, R., 2006. Analysis of antibiotic resistance genes in multidrug-resistant *Acinetobacter* sp. isolates from military and civilian patients treated at the Walter Reed Army Medical Center. Antimicrob. Agents Chemother. 50, 4114.

Huynen, M.A., Bork, P., 1998. Measuring genome evolution. Proc. Natl. Acad. Sci. 95, 5849.

Islam, A.H.M.S., Singh, K.K.B., Ismail, A., 2011. Demonstration of an outer membrane protein that is antigenically specific for *Acinetobacter baumannii*. Diagn. Microbiol. Infect. Dis. 69, 38–44.

Jahangiri, A., Rasooli, I., Gargari, S.L.M., Owlia, P., Rahbar, M.R., Amani, J., Khalili, S., 2011. An in silico DNA vaccine against Listeria monocytogenes. Vaccine 29, 6948–6958.

Jain, R., Danziger, L.H., 2004. Multidrug-resistant Acinetobacter infections: an emerging challenge to clinicians. Ann. Pharmacother. 38, 1449.

Katsaragakis, S., Markogiannakis, H., Toutouzas, K., Drimousis, P., Larentzakis, A., Theodoraki, E.M., 2008. Theodorou D: *Acinetobacter baumannii* infections in a surgical intensive care unit: predictors of multi-drug resistance. World J. Surg. 32, 1194–1202.

Keum, K.C., Yoo, S.M., Lee, S.Y., Chang, K.H., Yoo, N.C., Yoo, W.M., Kim, J.M., Choi, J.Y., Kim, J.S., Lee, G., 2006. DNA microarray-based detection of nosocomial pathogenic *Pseudomonas aeruginosa* and *Acinetobacter baumannii*. Molec. Cell. Probes 20, 42–50.

Kolaskar, A.S., Tongaonkar, P.C., 1990. A semi-empirical method for prediction of antigenic determinants on protein antigens. FEBS Lett. 276, 172–174.

Korber, B., LaBute, M., Yusim, K., 2006. Immunoinformatics comes of age. PLoS Computat. Biol. 2, e71.

Larsen, J.E.P., Lund, O., Nielsen, M., 2006. Improved method for predicting linear B-cell epitopes. Immunome Res. 2, 2.

Lasa, I., Penadés, J.R., 2006. Bap: A family of surface Proteins involved in biofilm formation. Res. Microbiol. 157, 99–107.

Latasa, C., Solano, C., Penadés, J.R., Lasa, I., 2006. Biofilm-associated proteins. C.R. Biol. 329, 849–857.

Latasa, C., Roux, A., Toledo-Arana, A., Ghigo, J.M., Gamazo, C., Penadés, J.R., Lasa, I., 2005. BapA, a large secreted protein required for biofilm formation and host colonization of *Salmonella enterica* serovar Enteritidis. Molec. Microbiol. 58, 1322–1339.

Lee, H.W., Koh, Y.M., Kim, J., Lee, J.C., Lee, Y.C., Seol, S.Y., Cho, D.T., 2008. Capacity of multidrug-resistant clinical isolates of *Acinetobacter baumannii* to form biofilm and adhere to epithelial cell surfaces. Clin. Microbiol. Infect. 14, 49–54.

Leerkes, M.R., Caballero, O.L., Mackay, A., Torloni, H., O'Hare, M.J., Simpson, A.J.G., de Souza, S.J., 2002. In silico comparison of the transcriptome derived from purified normal breast cells and breast tumor cell lines reveals candidate upregulated genes in breast tumor cells. Genomics 79, 257–265.

Levitt, M., 1978. Conformational preferences of amino acids in globular proteins. Biochemistry 17, 4277–4285.

Lin, L., Ling, B.D., Li, X.Z., 2009. Distribution of the multidrug efflux pump genes, adeABC, adeDE and adeIJK, and class 1 integron genes in multiple-antimicrobial-resistant clinical isolates of *Acinetobacter baumannii–Acinetobacter calcoaceticus* complex. Int. J. Antimicrob. Agents 33, 27–32.

Loehfelm, T.W., Luke, N.R., Campagnari, A.A., 2008. Identification and characterization of an *Acinetobacter baumannii* biofilm-associated protein. J. Bacteriol. 190, 1036–1044.

Lynch, S., Dixon, L., Benoit, M., Brodie, E., Keyhan, M., Hu, P., Ackerley, D., Andersen, G., Matin, A., 2007. Role of the rapA gene in controlling antibiotic resistance of *Escherichia coli* biofilms. Antimicrob. Agents Chemother. 51, 3650.

Monds, R., O'Toole, G., 2009. The developmental model of microbial biofilms: ten years of a paradigm up for review. Trends Microbiol. 17, 73–87.

Morgan, D.J., Weisenberg, S.A., Augenbraun, M.H., Calfee, D.P., Currie, B.P., Furuya, E.Y., Holzman, R., Montecalvo, M.C., Phillips, M., Polsky, B., Sepkowitz, K.A., 2009. Multidrug-resistant *Acinetobacter baumannii* in New York City—10 years into the epidemic. Infect. Control Hosp. Epidemiol. 30, 196–197.

Negi, S.S., Braun, W., 2009. Automated detection of conformational epitopes using phage display peptide sequences. Bioinformatics Biol. Insights 3, 71.

Nielsen, M., Lundegaard, C., Lund, O., Petersen, T.N., 2010. CPHmodels-3.0óremote homology modeling using structure-guided sequence profiles. Nucleic Acids Res. 38, W576.

Olfa, F.G., Radhouane, G., Abir, Z., Boutheina, G., Jalel, G., Ahmed, R., Adnene, H., 2008. Evaluation of an in silico predicted specific and immunogenic antigen from the OmcB protein for the serodiagnosis of *Chlamydia trachomatis* infections. BMC Microbiol. 8.

Parker, J., Guo, D., Hodges, R., 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.

Pellegrini, M., Marcotte, E.M., Thompson, M.J., Eisenberg, D., Yeates, T.O., 1999. Assigning protein functions by comparative genome analysis: protein phylogenetic profiles. Proc. Natl. Acad. Sci. USA 96, 4285.

Pellequer, J., Westhof, E., Van Regenmortel, M., 1991. Predicting location of continuous epitopes in proteins from their primary structures. Meth. Enzymol. 203, 176–201.

Rahbar, M.R., Rasooli, I., Mousavi Gargari, S.L., Amani, J., Fattahian, Y., 2010. In silico analysis of antibody triggering biofilm associated protein in *Acinetobacter baumannii*. J. Theor. Biol. 266, 275–290.

Rossolini, G.M., Mantengoli, E., 2008. Antimicrobial resistance in Europe and its potential impact on empirical therapy. Clin. Microbiol. Infect. 14, 2–8.

Roukos, D.H., 2009. Personalized cancer diagnostics and therapeutics. Expert Rev. Molec. Diagn. 9, 227–229.

Roy, A., Kucukural, A., Zhang, Y., 2010. I-TASSER: a unified platform for automated protein structure and function prediction. Nat. Protoc. 5, 725–738.

Shankar, V., Baghdayan, A.S., Huycke, M.M., Lindahl, G., Gilmore, M.S., 1999. Infection-derived *Enterococcus faecalis* strains are enriched in esp, a gene encoding a novel surface protein. Infect. Immunity 67, 193–200.

Sonnhammer, E.L.L., Durbin, R., 1995. A dot-matrix program with dynamic threshold control suited for genomic DNA and protein sequence analysis. Gene 167, GC1–GC10.

Stoodley, P., Sauer, K., Davies, D.G., Costerton, J.W., 2002. Biofilms as complex differentiated communities. Ann. Rev. Microbiol. 56, 187–209.

Suzek, B.E., Huang, H., McGarvey, P., Mazumder, R., Wu, C.H., 2007. UniRef: comprehensive and non-redundant UniProt reference clusters. Bioinformatics 23, 1282.

Szklarczyk, D., Franceschini, A., Kuhn, M., Simonovic, M., Roth, A., Minguez, P., Doerks, T., Stark, M., Muller, J., Bork, P., 2011. The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. Nucleic Acids Res. 39, D561.

Teodorescu, O., Galor, T., Pillardy, J., Elber, R., 2004. Enriching the sequence substitution matrix by structural information. Proteins Struct. Funct. Bioinformatics 54, 41–48.

Tien, H.C., Battad, A., Bryce, E.A., Fuller, J., Mulvey, M., Bernard, K., Brisebois, R., Doucet, J.J., Rizoli, S.B., Fowler, R., 2007. Multi-drug resistant Acinetobacter infections in critically injured Canadian forces soldiers. BMC Infect. Dis. 7, 95.

Tomaras, A.P., Dorsey, C.W., Edelmann, R.E., Actis, L.A., 2003. Attachment to and biofilm formation on abiotic surfaces by *Acinetobacter baumannii*: involvement of a novel chaperone–usher pili assembly system. Microbiology 149, 3473–3484.

Vidal, R., Dominguez, M., Urrutia, H., Bello, H., Gonzalez, G., Garcia, A., Zemelman, R., 1996. Biofilm formation by *Acinetobacter baumannii*. Microbios 86, 49–58.

Vidal, R., Dominguez, M., Urrutia, H., Bello, H., Garcia, A., Gonzalez, G., Zemelman, R., 1997. Effect of imipenem and sulbactam on sessile cells of *Acinetobacter baumannii* growing in biofilm. Microbios 91, 79–87.

Wass, M.N., Sternberg, M.J., 2009. Prediction of ligand binding sites using homologous structures and conservation at CASP8. Proteins 77 (suppl. 9), 147–151.

Wiederstein, M., Sippl, M.J., 2007. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res. 35, W407.

Wilkinson, D.L., Harrison, R.G., 1991. Predicting the solubility of recombinant proteins in *Escherichia coli*. Bio/technology (Nature Publishing Company) 9, 443–448.

Wu, S., Zhang, Y., 2007. LOMETS: a local meta-threading-server for protein structure prediction. Nucleic Acids Res. 35, 3375.

Zhang, L., Mah, T.F., 2008. Involvement of a novel efflux system in biofilm-specific resistance to antibiotics. J. Bacteriol. 190, 4447.

Zhang, T., Xin, R., Zhang, Q., Bao, Y., Wu, J., 2002. Characterization of intertype specific epitopes on adenoviruses hexon. [Zhonghua shi yan he lin chuang bing du xue za zhi=Zhonghua shiyan he linchuang bingduxue zazhi=Chinese]. J. Exp. Clin. Virol. 16, 44.

Zhang, Y., 2009. Protein structure prediction: when is it useful? Curr. Opinion Struct. Biol. 19, 145–155.

Zhang, Y., Skolnick, J., 2004. Scoring function for automated assessment of protein structure template quality. Proteins—N.Y. 57, 702–710.