

## *In silico* Characterization of Biofilm-Associated Protein (Bap) Identified in a Multi-drug Resistant *Acinetobacter baumannii* Clinical Isolate

Mohammad Reza Shakibaie<sup>1, 2\*</sup> 

<sup>1</sup>Department of Microbiology and Virology, Kerman University of Medical Sciences, Iran; <sup>2</sup>Environmental Health Engineering Research Center, Kerman University of Medical Sciences, Kerman, Iran

### ARTICLE INFO

#### Original Article

**Keywords:** *Acinetobacter baumannii*, Gene ontology, Protein modeling, Model validity, Amino acids alignments

Received: 10 Sep. 2021

Received in revised form: 29 Nov. 2021

Accepted: 04 Dec. 2021

DOI: 10.52547/JoMMID.9.4.210

#### \*Correspondence

Email: mr\_shakibaie@kmu.ac.ir

Tel: +987616913555

Fax: +983431325829

© The Author(s)



### ABSTRACT

**Introduction:** *Acinetobacter baumannii* (AB) is a Gram-negative bacteria associated with various hospital infections. The present study deals with *in silico* analysis of the biofilm-associated protein (Bap) in this pathogen. **Method:** Sixty-eight multi-drug resistant (MDR) AB were isolated from two hospitals in Kerman, Iran. Biofilm-formation was investigated using the microtiter method and PCR followed by sequencing to detect the *bap* gene in the strongest biofilm-forming isolate. The physicochemical parameters of Bap protein were determined by the ProtParam tool using the ExPasy program. The 3D models from the primary amino acid sequence were constructed using the I-TASSER modeling platform based on multiple-threading alignments by LOMETS. Nevertheless, to ensure the correct initial structure, the protein was minimized in energy through the 3DRefine software of the deep learning system. For the accuracy of predicted models, calculation of the orientation of dihedral angles, including the phi ( $\phi$ ) and psi ( $\psi$ ) and backbone conformation using the PROCHECK module of the PDB Sum server was performed. The domains and key amino acids involved in protein structure were studied by the Pfam and Interpro softwares. **Results:** Analysis of the amino acid content of the Bap protein revealed the absence of Arg and Cys in the protein structure. Our Bap protein exhibited ~99.6% identity with other Bap sequences in the GenBank database. Stereochemical simulation identified 19 antiparallel  $\beta$ -sheets with two small  $\alpha$ -helices. The N-terminal of Bap protein formed oligomers that mediate cellular adhesion. **Conclusion:** This study adds considerable information about Bap protein 3D structure, its conformation, domain analysis, and amino acids involved in cellular attachment.

### INTRODUCTION

*Acinetobacter baumannii* is a Gram-negative pathogen associated with various hospital infections, particularly in intensive care units (ICU) of hospitals worldwide [1]. This species accounts for ~80% of reported *Acinetobacter* infections, including ventilator-associated pneumonia, bacteremia, meningitis, peritonitis, urinary tract infections, and wound infections [2]. With the emergence of multi, extensive, and pandrug-resistant strains of this bacterium, the treatment process has become challenging, resulting in increased death rates among immunocompromised patients [3]. One of the outstanding *A. baumannii* features is the biofilm formation on medical devices and living tissue, representing a crucial microbial virulence factor [4]. The biofilm-forming *A. baumannii* isolates are intrinsically resistant to many antibiotics, including carbapenems, fluoroquinolones, and third-

generation cephalosporins [3]. One essential protein contributing to biofilm-formation in this bacterium is the biofilm-associated protein (Bap) [5].

The genus *Acinetobacter* hosts various large multidomain Baps proteins, all exhibiting similarity in structure, as they are internally repetitious and feature multiple (3-50 copies) immunoglobulin-like (Ig-like) domains. These domains have a particular three-dimensional structure known as Ig fold, comprising 70–100 amino acid (aa) residues in seven antiparallel  $\beta$ -strands organized in two  $\beta$ -sheets packed against each other in a sandwich structure [5]. Furthermore, size heterogeneity indicated changes in the number of amino acids due to recA-independent slipped-strand mispairing during DNA replication. Nevertheless, in the *A.*

*baumannii* hospital isolates, the core unit of the Bap protein, is highly conserved [7]. Considering that Bap protein promotes adhesion to biotic and abiotic surfaces, it seems reasonable that the fragments containing the N-terminus of the protein become aggregation-prone and self-assemble into amyloid-like structures under acidic pH and low concentrations of calcium [8].

Our previous study detected a *bap* gene in a potent biofilm-forming *A. baumannii* isolate and analyzed its expression [9]. However, information on the structure of a large part of Bap protein and its domains activities is still poorly understood. The present study deals with the physicochemical characterization, stereochemical simulation, protein modeling, amino acid alignment, and phylogenetic tree analysis of a Bap protein obtained from the strongest biofilm-forming MDR *A. baumannii* clinical isolate identified herein.

## MATERIALS AND METHODS

**Bacterial sources and antimicrobial susceptibility testing.** Bacterial source, gene identification, and antimicrobial susceptibility testing of 68 MDR clinical isolates of *A. baumannii* were determined as previously reported elsewhere [4].

**Biofilm formation.** Biofilm-formation was performed by microtiter assay as described previously [12]. The quantification of cells in the biofilms was carried out by solubilization of dye with 200  $\mu$ L of absolute ethanol at optical density (OD) of 480 nm. The OD cut-off value (OD<sub>c</sub>) was established as three standard deviations (SD) above the mean optical density (OD) of the negative control as shown in the formula: OD<sub>c</sub>=average OD of negative control +(3×SD of negative control). According to the optical density, the isolates were divided into four categories: (1) strong biofilm producer (4×OD<sub>c</sub><OD), (2) medium biofilm producer (2×OD<sub>c</sub><OD<4×OD<sub>c</sub>), (3) weak biofilm producer (OD<sub>c</sub><OD<2×OD<sub>c</sub>), and (4) no biofilm formation (OD<OD<sub>c</sub>). Among the strong biofilm producers, one that showed the highest OD was selected for further studies. The assays were performed in duplicates, and *Pseudomonas aeruginosa* PAO1 was used as biofilm positive strain.

**Detection and sequencing of *bap* gene.** A clinical *A. baumannii* isolate that showed the strongest biofilm-formation was selected and subjected to PCR using specific primer sequences to detect the *bap* gene [13]. The PCR product was sequenced by ABI Prism 373 DNA Sequencer (Applied Biosystems 373/3730XL, Bioneer, Korea) as described previously [4]. The generated sequence was compared with similar sequences deposited in the National Centre for Biotechnology Information (NCBI) database using BLAST comparison algorithm with default settings (https://www.ncbi.nlm.nih.gov/genbank/blast), and similarities were obtained.

**Nucleotide sequence accession number.** The *bap* gene and protein sequences were deposited in the GenBank database under the accession numbers KR080550 and AKL78797, respectively.

**Physicochemical characterization of the Bap protein.** Physicochemical properties of the Bap protein such as molecular weight, theoretical isoelectric point, atomic composition, extinction coefficient, instability index, aliphatic index, and grand average of hydropathicity were determined using the ExPasy ProtParam server (https://web.expasy.org/cgi-bin/protparam/protparam). Furthermore, the amino acid compositions based on percent quantity were determined by ExPasy ProtParam software.

**Classification, functional analysis, and phylogenetic tree of Bap protein.** The primary classification and functional analysis of Bap protein were determined using the InterPro software ver. 5.51-85 (www.Ebi.ac.uk/interpro); it provides an integrative classification of protein sequences into families and domains. The biological functions of the PmrA protein, including ligand-binding site, molecular and cellular functions in terms of Gene Ontology (GO), were inferred from the PDP database based on comparing sequence and structural similarity using the I-TASSER server. Nevertheless, the phylogenetic tree was constructed using the unweighted pair group method with arithmetic mean (UPGMA). The sequence was subjected to the amino acid homology analysis using the BLAST tool obtained from the UniProtKB database (https://www.uniprot.org/). The Bap amino acid sequences displaying the maximum identity to the query sequence were used for alignment and construction of phylogeny [15]. The dendrogram from the distance matrix was generated in MEGA 6 software using the maximum likelihood (ML) method based on the JTT matrix-based model. The bootstrap test of phylogeny was assessed with 1000 repetitions [15].

**Three-dimensional structure of Bap protein.** For constructing the correct 3D protein model from the primary Bap amino acid sequence, the iterative threading assembly refinement (I-TASSER) online software based on multiple-threading alignments LOMETS was used (https://zhanglab.dcmf.med.umich.edu/I-TASSER/). Five closest stereochemical structures were obtained, and the confidence of each model was quantitatively measured by confidence (C)-score. The C-score was calculated based on the significance of threading template alignments and the convergence parameters of the structure assembly simulations. The C-score is generally within the -5 to +2 range. Higher C-score signifies higher confidence in the model. Subsequently, the quality of the structure was estimated by the C, TM, and RMDS scores. In addition, the stability of the models was estimated using the structural validation algorithm ProSA [16]. For confirming the accuracy of the protein secondary structure, the PSSpred algorithm

was used as described previously [17]. The orientation of dihedral angles (torsion) angles phi ( $\Phi$ ) and psi ( $\psi$ ) were determined using a logistic representation Z-score of the Ramachandran plot and the SAVES program (<https://saves.mbi.ucla.edu/>). The predicted models were then subjected to the backbone conformation using the PROCHECK module of the PDBSum server (<https://www.ebi.ac.uk/thornton-srv/software/PROCHECK/>). Finally, a template with the highest quality was selected for the model building based on the confidence score of the Deep Learning Potential [18, 19].

**Structural similarity search.** For searching the structural similarity of the Bap protein, the MADOKA webserver engine (<http://madoka.denglab.org/>) was used. The server consisted of two-phase, searching similar protein structures by aligning the input with the whole PDB library.

## RESULTS

**Bacterial isolates and biofilm formation.** Sixty-eight *A. baumannii* isolates were collected from different wards of two hospitals in Kerman, Iran. All isolates exhibited MDR phenotypes resistance to the third cephalosporins generation, aminoglycosides, and quinolones; 24% of the isolates showed strong, 58% intermediate, and 18% weak or no biofilm activities.

**Physicochemical properties of Bap protein/** Template search by ExPasy ProtParam tools revealed that the Bap protein was a membrane-bound monomeric nonredundant joint protein. The physicochemical parameters of the Bap protein are listed in Table 1. The instability index (II) 15.48 classifies the protein as stable; theoretical pI 3.46 indicated the acidic nature of Bap protein with Ext. Coefficient 41940 M<sup>-1</sup> cm<sup>-1</sup> and grand average of hydropathicity (GRAVY) - 0.016.

**Table 1.** The physicochemical properties of Bap protein obtained in this study.

Parameter	Results
Molecular weight	39557.83
Number of amino acids	396
Theoretical pI	3.46
Total number of negatively charged residues (Asp + Glu)	56
Total number of positively charged residues (Arg + Lys)	7
<b>Atomic composition</b>	
Carbon C	1718
Hydrogen H	2685
Nitrogen N	445
Oxygen O	624
Sulfur S	0
<b>Formula</b>	718H2685N445O624
<b>Total number of atoms</b>	5472
<b>Aliphatic index</b>	85.25
<b>Ext. coefficient</b>	41940
<b>Abs 0.1% (=1 g/l)</b>	1.060
<b>Instability index</b>	15.48
<b>Grand average of hydropathicity (GRAVY)</b>	- 0.016

Extinction coefficients are in M<sup>-1</sup> cm<sup>-1</sup>, at 280 nm measured in water. Abs 0.1% (=1 g/L) 1.060. The aliphatic index of a protein was calculated according to the following formula: Aliphatic Index = X(Ala) + a X(Val) + b (X(Ile) + X(Leu) where X(Ala), X(Val), X(Ile), and X(Leu) are mole percent (100 X mole fraction) of alanine, valine, isoleucine, and leucine.

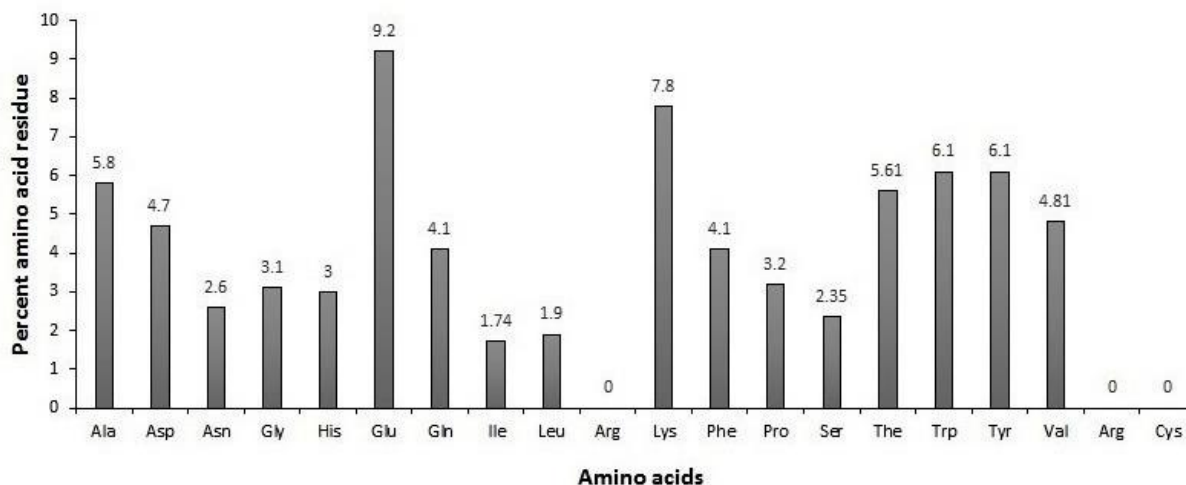
**Amino acids compositions of Bap protein.** The primary analysis of the amino acid content of the Bap protein revealed an absence of Arg and Cys, whereas Glu with 9.2%, Lys with 7.8%, and Trp and Tyr with 6.1% were the most abundant amino acids residues in the protein structure (Fig. 1).

**The classification, structural diversity, and phylogenetic tree.** The InterPro tool classification of the generated Bap protein sequence included the superfamily Ca\_tandamer, containing three regions, namely Big\_3 and

Big\_16 and Ig-like sequences. The sequences inserted under each domain indicated order predictability (Fig. 2A). Interestingly in our case, the Ig-like sequence begins with Gly at position 111, while Big\_3 and Big\_16 sequences begin with Ala at position 217 and Asp at position 304, respectively. The Bap protein belonged to the Ig-like homologs superfamily with cell adhesion property and calcium-binding tandamer activity. Moreover, hierarchical information based on the

structural diversity of this superfamily was compared to other superfamilies in the CATH database, which is a protein structures classification tool in the Protein Data Bank (Fig. 2B); protein structures are chopped up into domains and grouped into superfamilies, and the evolutionary relationships among them are determined through a combination of automated algorithms and

expert manual curation. The sequence diversity was ~73%, indicating a relatively low diversity (Fig. 2B). This result was supported by phylogenetic tree analysis of Bap protein which showed 99.6% identity with other Bap sequences (Fig. 2C). The analysis showed 99.6% identity with *A. baumannii* MBO4071943.1 and 99.4% with *A. baumannii* WP\_140983092.1, respectively.



**Fig. 1.** Amino acid composition of Bap protein characterized in this study based on ExPasy ProtParam software. The Glu and Lys had the highest percent numbers, and Arg and Cys were absent from protein structure.

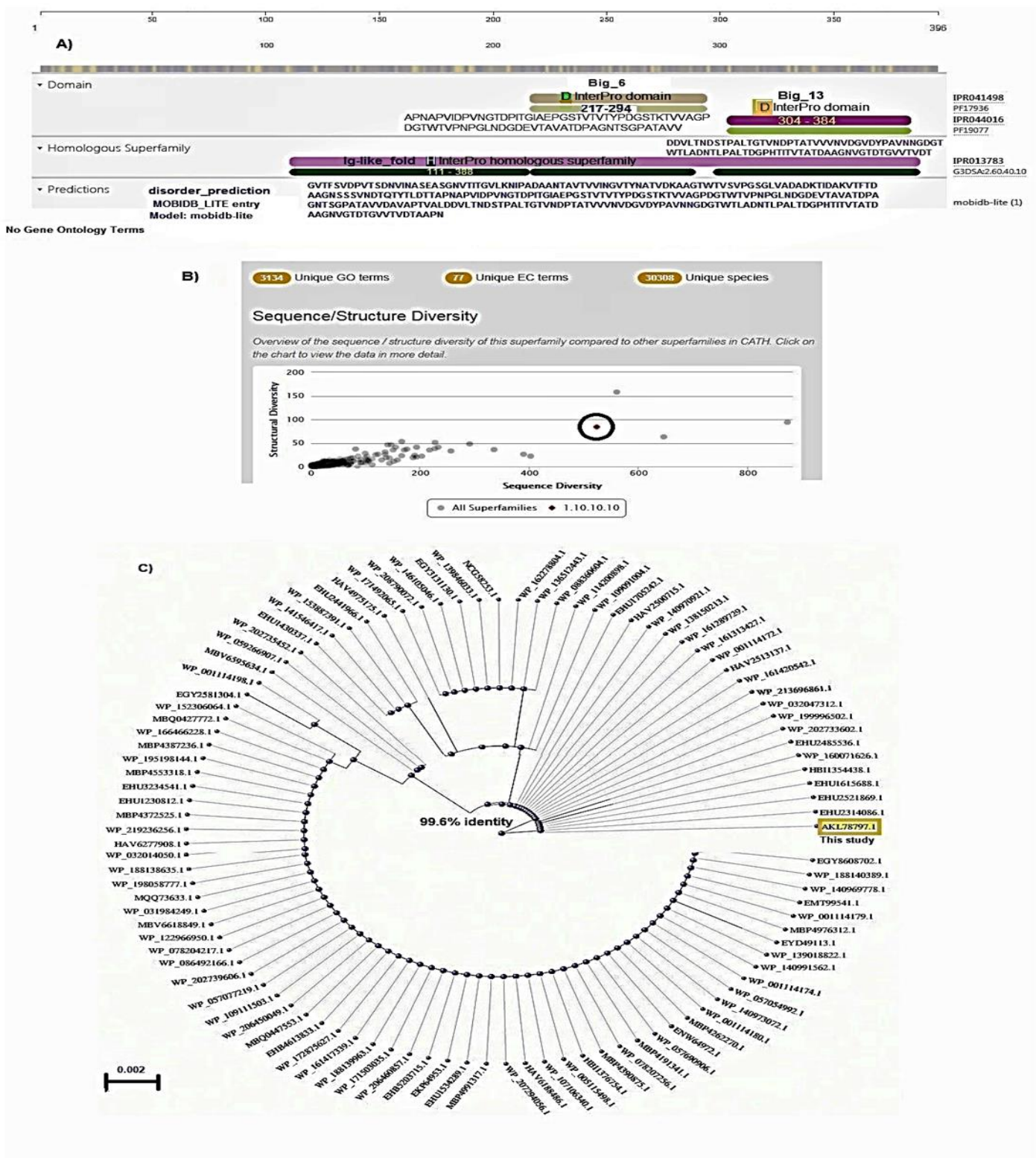
**Sequence Alignments.** Multiple amino acid sequence alignments of the Bap protein (Fig. 3) indicated that the "IWAKFKDAENDADADSDADADADS" consensus sequence was responsible for extracellular adhesion, bringing together bacterial cells. Similarly, the surface binding domain sequence is demonstrated by yellow color. This region seems to be conserved in many Bap proteins. No amino acid substitution and mutation were detected in this study.

**Structure and quality of Bap protein.** The cartoon structure of Bap protein obtained by the I-TASSER webserver (Fig. 4A) contained 19 antiparallel  $\beta$ -sheets with two small  $\alpha$ -helices. The stereochemical trajectory of the protein exhibited three main regions and one surface N-terminal oligomers that mediate biofilm development. This sequence was conserved in many Bap proteins. In contrast, the C-terminal region was membrane bound and variable in amino acid composition. The region-A consisted of 5 antiparallel  $\beta$ -sheets while region-B had six antiparallel  $\beta$ -sheets (Fig. 5A). Interestingly, region-D was situated in the C-terminal site of protein consisting of 6  $\beta$ -sheets and one small  $\alpha$ -helix.

**Quality estimation.** The results of the Ramachandran plot (Fig. 5A) indicated a clash score of 13.68%, Ramachandran favored 86.08%, and mole probability

score = 2.49. The predicted contact and distance probability matrix of Bap protein (Fig. 5B) was displayed by darker color as a higher probability. Furthermore, the contact map of the bap protein defined by Cb-Cb distance showed  $\leq 8$  angstroms. Furthermore, ASP 273, ASN 149, and THR 246 were situated in the non-favored region. The predicted contact and distance probability matrix of Bap protein maps were also determined in the LOMETS platform (Fig. 5B). The darker color and increased yellow color intensity indicated more accuracy of model structure. Furthermore, to predict the exact secondary structure of Bap protein, five highly similar models were obtained. The highest C- score of 0.75 was selected for further study.

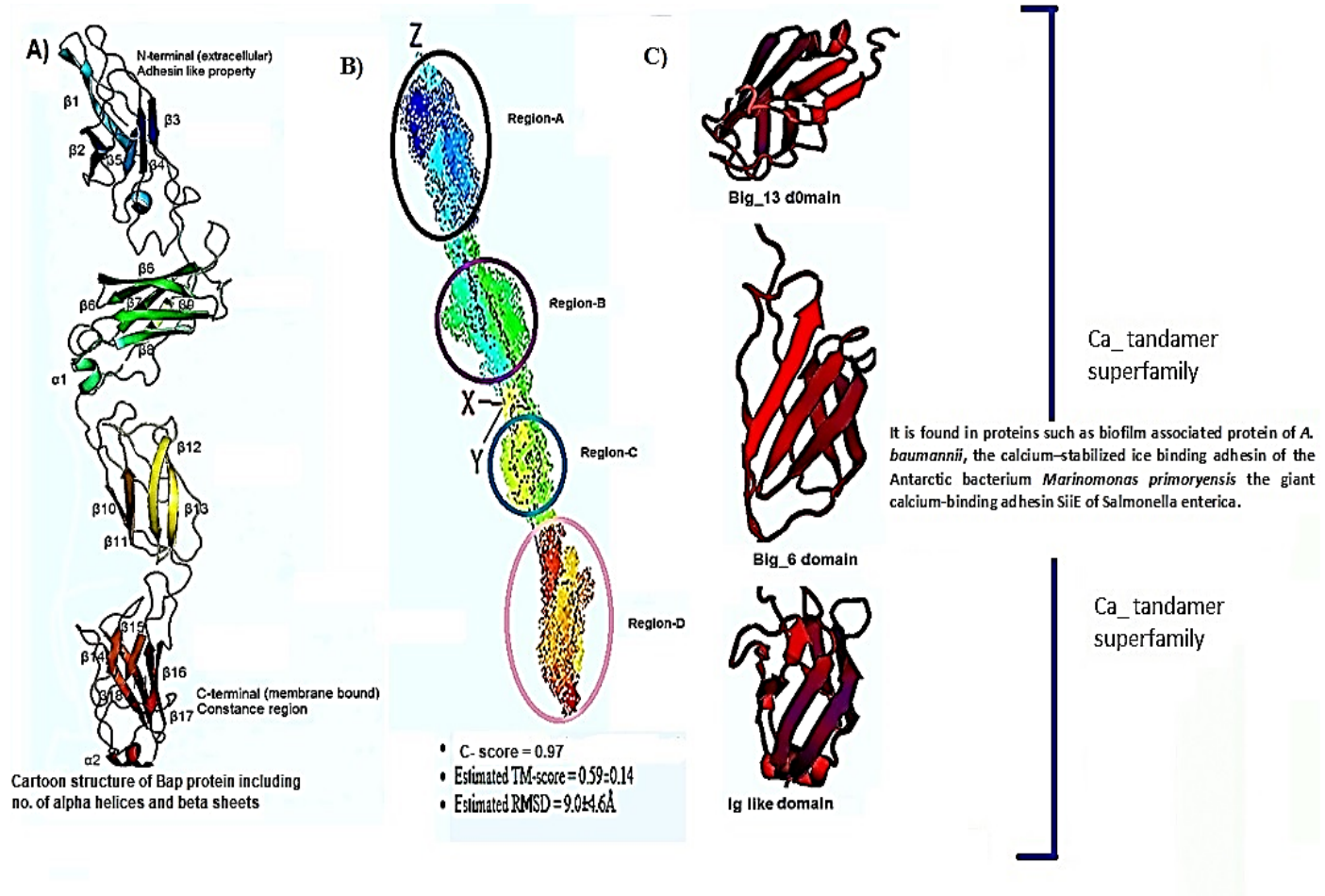
**Amino acids involved in adhesion.** The schematic representation of the amino acids involved in cellular adhesion of Bap protein (Fig. 6) showed that Pro144, Ile 143, Val 138, Tyr 161, Asp 146, Ile 136, Asn 142, and Lys 141 play an important role in biofilm formation.



**Fig. 2.** A) Classification of the submitted sequence of the Bap protein in InterPro tool software. The Ig-like domains have three motifs, fitting the consensus TDnAGN, found in many bacterial surface proteins. B) Overview of the sequence and structural diversity of Bap superfamily by CATH system. C) The phylogenetic tree analysis of Bap protein. The analysis showed a close identity of Bap protein with other similar proteins reported in the GenBank database.

Annotation	A0A0G3VGQ3_ACIBA	1	-----	0
✓ Non-terminal residue	A0A506JXT6_ACIBA	1	MPEIQIIAKDSHKTLVITTEGTSAKLSEASVVLVKVAASDVLVNNREGTNAVIRLKNGETI	60
✓ Domain	A0A5K6CN10_ACIB3	1	MPEIQIIAKDNHKTILVITTEGTSAKLSEASVVLVKVAASDVLVNNREGTNAVIRLKNGETI	60
✓ Region	A0A0D5YEF7_ACIBA	1	MPEIQIIAKDNHKTILVITTEGTSAKLSEASVVLVKVAASDVLVNNREGTNAVIRLKNGETI	60
	A0A2Z0LC85_ACIBA	1	MPEIQIIAKDNHKTILVITTEGTSAKLSEASVVLVKVAASDVLVNNREGTNAVIRLKNGETI	60
	B0LHN4_ACIBA	1	MPEIQIIAKDNHKTILVITTEGTSAKLSEASVVLVKVAASDVLVNNREGTNAVIRLKNGETI	60
Extracellular adhesin like sequence				
	A0A0G3VGQ3_ACIBA	1	-----TWAKFKD <del>FNDAADSDADADADSD</del> DVEPQALLGE	34
	A0A506JXT6_ACIBA	61	VIEGEFSGTAEFKDNLVLFQDENGQLIWAFFKDAENDADADSDADADADSDVEPQALLGE	120
	A0A5K6CN10_ACIB3	61	VIEGEFSGTAEFKDNLVLFQDENGQLIWAFFKDAENDADADSDADADADSDVEPQALLGE	120
	A0A0D5YEF7_ACIBA	61	VIEGEFSGTAEFKDNLVLFQDENGQLIWAFFKDAENDADADSDADADADSDVEPQALLGE	120
	A0A2Z0LC85_ACIBA	61	VIEGEFSGTAEFKDNLVLFQDENGQLIWAFFKDAENDADADSDADADADSDVEPQALLGE	120
	B0LHN4_ACIBA	61	VIEGEFSGTAEFKDNLVLFQDENGQLIWAFFKDAENDADADSDADADADSDVEPQALLGE	120
*****				
	A0A0G3VGQ3_ACIBA	35	DLPAALPAEAPQELVSDVIYQPISSIEPLLYHDAGVNFPLWAAIPLVAGGIIAAASNNHDS	94
	A0A506JXT6_ACIBA	121	DLPAALPAEAPQELVSDVIYQPISSIEPLLYHDAGVNFPLWAAIPLVAGGIIAAASNNHDS	180
	A0A5K6CN10_ACIB3	121	DLPAALPAEAPQELVSDVIYQPISSIEPLLYHDAGVNFPLWAAIPLVAGGIIAAASNNHDS	180
	A0A0D5YEF7_ACIBA	121	DLPAALPAEAPQELVSDVIYQPISSIEPLLYHDAGVNFPLWAAIPLVAGGIIAAASNNHDS	180
	A0A2Z0LC85_ACIBA	121	DLPAALPAEAPQELVSDVIYQPISSIEPLLYHDAGVNFPLWAAIPLVAGGIIAAASNNHDS	180
	B0LHN4_ACIBA	121	DLPAALPAEAPQELVSDVIYQPISSIEPLLYHDAGVNFPLWAAIPLVAGGIIAAASNNHDS	180
*****				
Intracellular signal				
	A0A0G3VGQ3_ACIBA	95	<del>NDDSSAPADITTPSTIDGVTFESVDPVTS</del> SDNVINASEASGNVTITGVLNIPADAANTAVIV	154
	A0A506JXT6_ACIBA	181	<del>NDDSSAPADITTPSTIDGVTFESVDPVTS</del> SDNVINASEASGNVTITGVLNIPADAANTAVIV	240
	A0A5K6CN10_ACIB3	181	<del>NDDSSAPADITTPSTIDGVTFESVDPVTS</del> SDNVINASEASGNVTITGVLNIPADAANTAVIV	240
	A0A0D5YEF7_ACIBA	181	<del>NDDSSAPADITTPSTIDGVTFESVDPVTS</del> SDNVINASEASGNVTITGVLNIPADAANTAVIV	240
	A0A2Z0LC85_ACIBA	181	<del>NDDSSAPADITTPSTIDGVTFESVDPVTS</del> SDNVINASEASGNVTITGVLNIPADAANTAVIV	240
	B0LHN4_ACIBA	181	<del>NDDSSAPADITTPSTIDGVTFESVDPVTS</del> SDNVINASEASGNVTITGVLNIPADAANTAVIV	240
*****				
	A0A0G3VGQ3_ACIBA	155	VINGVTYNATVDKAAAGTWIVSVPGSLVADADKTIIDAKVFTTDAAGNSSSVNDTQTYTLD	214
	A0A506JXT6_ACIBA	241	VINGVTYNATVDKAAAGTWIVSVPGSLVADADKTIIDAKVFTTDAAGNSSSVNDTQTYTLD	300
	A0A5K6CN10_ACIB3	241	VINGVTYNATVDKAAAGTWIVSVPGSLVADADKTIIDAKVFTTDAAGNSSSVNDTQTYTLD	300
	A0A0D5YEF7_ACIBA	241	VINGVTYNATVDKAAAGTWIVSVPGSLVADADKTIIDAKVFTTDAAGNSSSVNDTQTYTLD	300
	A0A2Z0LC85_ACIBA	241	VINGVTYNATVDKAAAGTWIVSVPGSLVADADKTIIDAKVFTTDAAGNSSSVNDTQTYTLD	300
	B0LHN4_ACIBA	241	VINGVTYNATVDKAAAGTWIVSVPGSLVADADKTIIDAKVFTTDAAGNSSSVNDTQTYTLD	300
*****				
Surface binding domain				
	A0A0G3VGQ3_ACIBA	215	TTAPNAPVIDPVNGTDPITGTAEFGSTIVITVYPDGSTKTIVVAGSPDGTWIVPNPGLNDGDE	274
	A0A506JXT6_ACIBA	301	TTAPNAPVIDPVNGTDPITGTAEFGSTIVITVYPDGSTKTIVVAGSPDGTWIVPNPGLNDGDE	360
	A0A5K6CN10_ACIB3	301	TTAPNAPVIDPVNGTDPITGTAEFGSTIVITVYPDGSTKTIVVAGSPDGTWIVPNPGLNDGDE	360
	A0A0D5YEF7_ACIBA	301	TTAPNAPVIDPVNGTDPITGTAEFGSTIVITVYPDGSTKTIVVAGSPDGTWIVPNPGLNDGDE	360
	A0A2Z0LC85_ACIBA	301	TTAPNAPVIDPVNGTDPITGTAEFGSTIVITVYPDGSTKTIVVAGSPDGTWIVPNPGLNDGDE	360
	B0LHN4_ACIBA	301	TTAPNAPVIDPVNGTDPITGTAEFGSTIVITVYPDGSTKTIVVAGSPDGTWIVPNPGLNDGDE	360
*****				
	A0A0G3VGQ3_ACIBA	275	VTAVATDPAGNTSGPATAVVDVAVPTVALDDVLTNDSTPALTGTVNDPTATVVVNVGVD	334
	A0A506JXT6_ACIBA	361	VTAVATDPAGNTSGPATAVVDVAVPTVALDDVLTNDSTPALTGTVNDPTATVVVNVGVD	420
	A0A5K6CN10_ACIB3	361	VTAVATDPAGNTSGPATAVVDVAVPTVALDDVLTNDSTPALTGTVNDPTATVVVNVGVD	420
	A0A0D5YEF7_ACIBA	361	VTAVATDPAGNTSGPATAVVDVAVPTVALDDVLTNDSTPALTGTVNDPTATVVVNVGVD	420
	A0A2Z0LC85_ACIBA	361	VTAVATDPAGNTSGPATAVVDVAVPTVALDDVLTNDSTPALTGTVNDPTATVVVNVGVD	420
	B0LHN4_ACIBA	361	VTAVATDPAGNTSGPATAVVDVAVPTVALDDVLTNDSTPALTGTVNDPTATVVVNVGVD	420
*****				
	A0A0G3VGQ3_ACIBA	335	YPAVNNNGDGTWTLADNLTPLADGPHITITVATDAAAGNVGTDIGVTVDTAAPNTAGVTF	394
	A0A506JXT6_ACIBA	421	YPAVNNNGDGTWTLADNLTPLADGPHITITVATDAAAGNVGTDIGVTVDTAAPNTAGVTF	480
	A0A5K6CN10_ACIB3	421	YPAVNNNGDGTWTLADNLTPLADGPHITITVATDAAAGNVGTDIGVTVDTAAPNTAGVTF	480
	A0A0D5YEF7_ACIBA	421	YPAVNNNGDGTWTLADNLTPLADGPHITITVATDAAAGNVGTDIGVTVDTAAPNTAGVTF	480
	A0A2Z0LC85_ACIBA	421	YPAVNNNGDGTWTLADNLTPLADGPHITITVATDAAAGNVGTDIGVTVDTAAPNTAGVTF	480
	B0LHN4_ACIBA	421	YPAVNNNGDGTWTLADNLTPLADGPHITITVATDAAAGNVGTDIGVTVDTAAPNTAGVTF	480
*****				
Nonterminal residues				
	A0A0G3VGQ3_ACIBA	395	-----	396
	A0A506JXT6_ACIBA	481	-----	480
	A0A5K6CN10_ACIB3	481	TIDSVTADNVINASEAAGNVITITGVLNIPADATNTAVTVVINGVTYNATVDKTAGTWIV	540
	A0A0D5YEF7_ACIBA	481	TIDSVTADNVINASEAAGNVITITGVLNIPADATNTAVTVVINGVTYNATVDKTAGTWIV	540
	A0A2Z0LC85_ACIBA	481	TIDSVTADNVINASEAAGNVITITGVLNIPADATNTAVTVVINGVTYNATVDKTAGTWIV	540
	B0LHN4_ACIBA	481	TIDSVTADNVINASEAAGNVITITGVLNIPADATNTAVTVVINGVTYNATVDKTAGTWIV	540

**Fig. 3.** Multiple amino acids alignments of Bap protein with closely related sequences using UniProtKB database (<https://www.uniprot.org/help/sequence-alignments>).

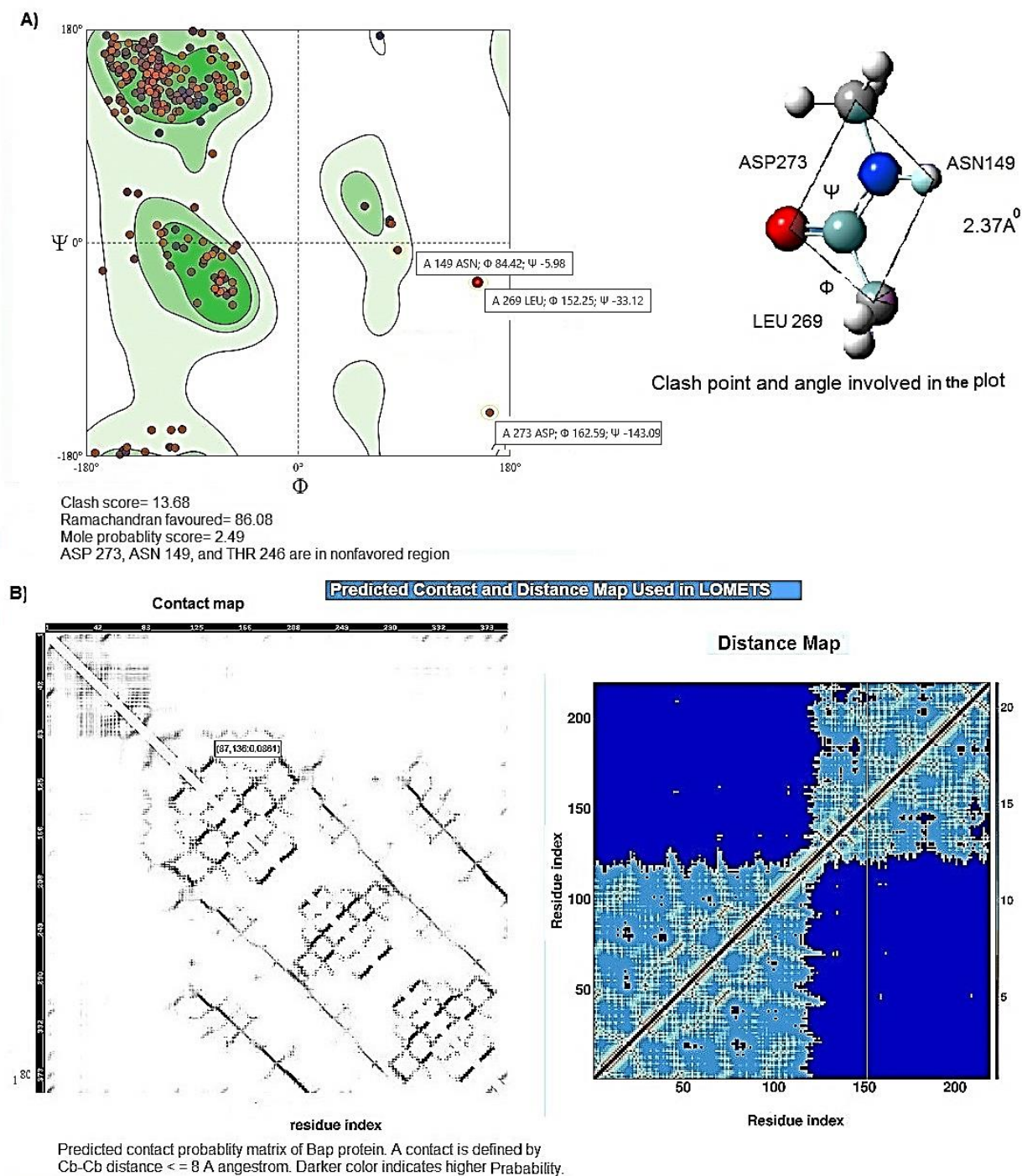


**Fig. 4.** The generated cartoon structure of Bap protein. **A)** The protein consisted of 19 antiparallel  $\beta$ -sheets with two small  $\alpha$ -helices. The amino acids in  $\alpha 15$ ,  $\alpha 16$ , and  $\alpha 18$  are located in the C-terminal and anchored to the membrane. **B)** Bap protein regions and C-TM-RMSD scores of a highly accurate model of Bap protein. The model contained 4 regions where N-terminal forms oligomers and mediates biofilm development. This sequence was conserved in many bap proteins. In contrast, the C-terminal region was membrane bound and variable in amino acid composition. **C)** Domains analysis of Bap protein.

## DISCUSSION

Bacteria can assume two modes of growth: unicellular or planktonic cells and multicellular forms or biofilms. The planktonic bacteria have better freedom of migration, are more prone to mutation but are more sensitive to the environment and antibiotics. In the multicellular life phase or biofilms, they are better protected from the harsh environment, are less subjected to mutation, and are more resistant to antibiotics and disinfectants [22]. Most of the available information on the Bap protein is related to *Staphylococcus aureus*, but the data can most likely be attributed to the other Bap orthologs. The *A. baumannii* bap gene encodes an extensive surface protein, comprising 8,621 amino acids homologous to the Bap protein first described in *S. aureus* [7].

The protein structure predicted in this study using homology modeling indicated model 1, the best Bap protein model in *A. baumannii* that shows strong biofilm. Multiple sequence alignment resulted in significant sequence identity to similar sequences that existed in the UniProtKB database. This close identity was observed mainly in conserved regions such as the N-terminal consensus sequence and surface domain through homophilic interactions between coiled-coil domains of opposing proteins in neighboring cells. The results were further supported by dendrogram analysis which suggested 99.6% homology with those Bap-like prefix domains containing protein. The protein that belonged to the superfamily Ca\_tandamer contained Ig-like domains; these domains are found in a wide variety of extracellular bacterial proteins, often in multiple copies [15].



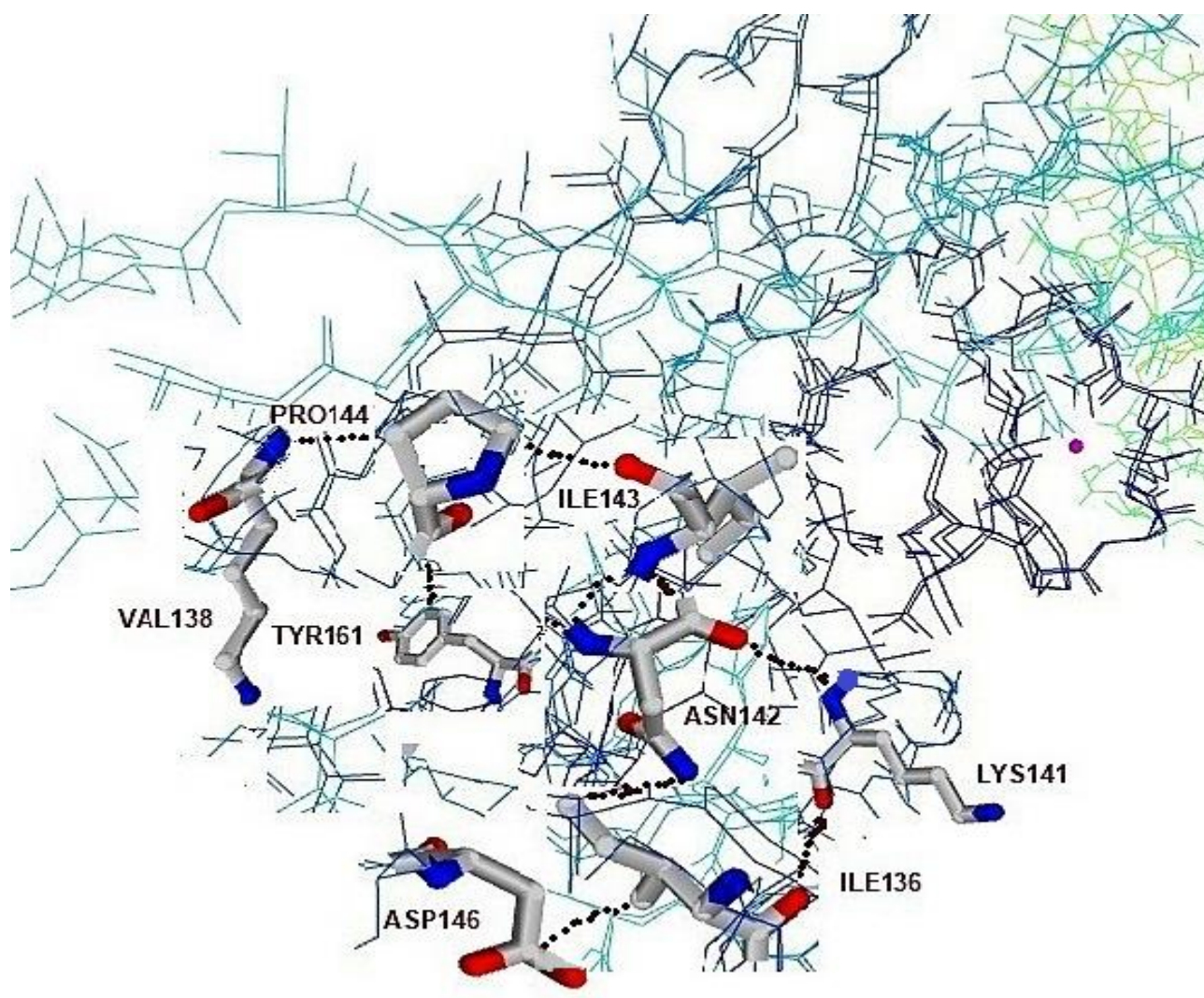
**Fig. 5. A)** Ramachandran plot for quality of the 3D structure of Bap protein. **B)** Predicted contact and distance map of Bap protein using the LOMETS program.

Stereochemical analysis of Bap protein revealed a monomeric high molecular weight protein that contained 4 regions well suited in X, Y, and Z axes. The high quality of the model structure was determined by I-TASSER software and approved by the ExPasy modeler. The amino acid residues position in segments from the logistic

representation of high-fidelity Ramachandran plot data of the model protein structure showed mainstream  $\delta'$ -region near 86% of the residues, representing the most favorable combinations of phi and psi values. The  $\Phi$ - $\Psi$  population density in the Ramachandran plot represents the total folding free

energy. Further analysis of amino acid residues of Bap protein in the present study revealed Glu with 9.2% as the most frequent amino acid, while Arg and Cys were absent. In a similar study [7], analysis of the primary

structure of the Bap protein showed absence of Cys and an abundance of Thr (1,389 residues; 16% of the total), Ala (1,176 residues; 14% of the total), and Val (1,109 residues; 13% of the total) [7].



**Fig. 6.** The schematic presentation of amino acids involved in cellular adhesion activity of Bap protein.

Previously, analysis of the physicochemical characteristics of Bap revealed that the protein was acidic (estimated isoelectric point  $[pI] = 3.37$ ) with a molecular weight of 220959.8 Da [23]. Furthermore, the core region of Bap represents 52% of the protein consisting of 13 successive nearly identical repeats, each containing 86 amino acids and two canonical calcium-binding EF-hand motifs able to control Bap functionality in response to the calcium present in the growth media [23]. Surprisingly, a large part of the protein is not required for the adhesion-accumulation biofilm phenotype, and that Bap-mediated

biofilm development occurs in response to environmental cues [24].

Despite all the studies supporting the crucial role of Bap protein in biofilm formation, a few reports have described the detailed stereochemical structure of this protein in high biofilm-producing *A. baumannii* isolates [25]. This study, therefore, paved the way for better understanding the structure, amino acids similarities, modeling prediction, and domains analysis of Bap protein in *A. baumannii* isolates.

## ACKNOWLEDGMENT

I thank the Department of Microbiology, Kerman University of Medical Sciences, Center for Informatics Research, and the Computer laboratory of Afzalipour Medical School for cooperation.

## CONFLICT OF INTEREST

The author declares there are no conflicts of interest associated with this manuscript.

## REFERENCES

- Shakibaie MR, Adeli S, Salehi MH. Antibiotic resistance patterns and extended-spectrum  $\beta$ -lactamase production among *Acinetobacter* spp. isolated from an intensive care unit of a hospital in Kerman, Iran. *Antimicrob Resist Infect Control*. 2012; 1 (1): 1.
- Peleg AY, Seifert H, Paterson DL. *Acinetobacter baumannii*: emergence of a successful pathogen. *Clin Microbiol Rev*. 2008; 21 (3): 538–82.
- Ghahraman MRK, Hosseini-Nave H, Azizi O, et al. Molecular characterization of lpxACD and pmrA/B two-component regulatory system in the colistin resistance *Acinetobacter baumannii* clinical isolates. *Gene Rep*. 2020; 21: 100952.
- Modarresi F, Azizi O, Shakibaie MR, Motamedifar M, Valibigi B, Mansouri S. Effect of iron on expression of efflux pump (adeABC) and quorum sensing (luxI, luxR) genes in *Acinetobacter baumannii*. *APMIS*. 2015; 123 (11): 959–68.
- Bodelon G, Palomino C, Fernandez LA. Immunoglobulin domains in *Escherichia coli* and other Enterobacteria: from pathogenesis to applications in antibody technologies. *FEMS Microbiol Rev*. 2013; 37 (2): 204–50.
- Lasa I, Penades JR. Bap: a family of surface proteins involved in biofilm formation. *Res Microbiol*. 2006; 157 (2): 9–107.
- Loehfelm TW, Luke NR, Campagnari AA. Identification and characterization of an *Acinetobacter baumannii* biofilm-associated protein. *J Bacteriol*. 2008; 190 (3): 1036–44.
- Taglialegna A, Navarro S, Ventura S, Garnett JA, Matthews S, Penades JR, et al. Staphylococcal Bap Proteins Build Amyloid Scaffold Biofilm Matrices in Response to Environmental Signals. *PLoS Pathogens*. 2016; 12: e1005711.
- Azizi O, Shahcheraghi F, Salimizand H, Modarresi F, Shakibaie MR, Mansouri Sh, et al. Molecular analysis and expression of *bap* gene in biofilm-forming multi-drug-resistant *Acinetobacter baumannii*. *Rep Biochem Mol Biol*. 2016; 5 (1): 62–71.
- Bouvet PJM, Grimont PAD. Identification and biotyping of clinical isolates of *Acinetobacter*. *Ann Inst Pasteur, Microbiol*. 1987; 138 (5): 569–78.
- Clinical and Laboratory Standards Institute. 2018. Performance standards for antimicrobial disk susceptibility tests. M02 standard, 13th ed Clinical and Laboratory Standards Institute, Wayne, PA.
- Gholamrezazadeh M, Shakibaie MR, Monirzadeh F, Hashemizadeh Z. Effect of nano-silver, nano-copper, deconex and benzalkonium chloride on biofilm formation and expression of transcription regulatory quorum sensing gene (rhIR) in drug-resistance *Pseudomonas aeruginosa* burn isolates. *Burns*. 2018; 44 (3): 700–8.
- Chen TL, Sin LK, Wu RC, Shaio MF, Huang Ly, Fung CP, et al. Comparison of one tube multiplex PCR, automated Ribotyping and intergenic spacer (ITS) sequencing for rapid identification of *Acinetobacter baumannii*. *Clin Microbiol Infect*. 2007; 13 (8): 801–6.
- Loehfelm, TW, Luke NR, Campagnari AA. Identification and Characterization of an *Acinetobacter baumannii* Biofilm-Associated Protein. *J Bacteriol*. 2020; 190 (3): 1036–44.
- Tamura, K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA-5: molecular evolutionary genetics analysis using maximum like hood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol*. 2011; 28 (10): 2731–9.
- Wiederstein M, Sippl MJ. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res*. 2007; 35 (Web Server issue): W407–10.
- Movahedi M, Zare-Mirakabad F, Arab SS. Evaluating the accuracy of protein design using native secondary sub-structures. *BMC Bioinform*. 2016; 17: 353.
- Pakhrin S, Shrestha B, Adhikari B, Dukka KC. Deep Learning-Based Advances in Protein Structure Prediction. *Int J Mol Sci*. 2021; 22 (11): 5553.
- Zhang C, Zheng W, Mortuza S, Li Y, Zhang Y. Deep MSA: Constructing deep multiple sequence alignment to improve contact prediction and fold-recognition for distant-homology proteins. *Bioinformatics*. 2020; 36 (7): 2105–12.
- Deng L, Zhong G, Liu C, et al. MADOKA: an ultra-fast approach for large-scale protein structure similarity searching. *BMC Bioinform*. 2019; 20: 662.
- Kelley LA, Sternberg MJ. Protein structure prediction on the Web: a case study using the Phyre server. *Nat Protoc*. 2009; 4 (3): 363–71.
- Shakibaie MR. Bacterial Biofilm and its Clinical Implications. *Ann Microbiol Res*. 2018; 2: 45–50.
- Abedulhussein TM, Saadedin MK et al. *In silico* Designing of Biofilm-associated Protein (Bap) producing of *Acinetobacter baumannii*. *Int. J Nat Eng Sci*. 2019; 13 (3): 79–82.
- Cucarella C, Solano C, Valle J, Amorena B, Lasa I, Penadés JR. Bap, a *Staphylococcus aureus* surface protein

involved in biofilm formation. J Bacteriol. 2001; 183 (9): 2888-96.

25. Valle J, Fang X, Lasa I. Revisiting Bap Multidomain Protein: More than sticking bacteria together. Front Microbiol. 2020; 11 (3): 613581.

**Cite this article:**

Shakibaie MR. *In silico* Characterization of Biofilm-Associated Protein (Bap) Identified in a Multi-drug Resistant *Acinetobacter baumannii* Clinical Isolate. J Med Microbiol Infect Dis, 2021; 9 (4): 210-220. DOI: 10.52547/JoMMID.9.4.210.