Data Article

Data retrieved from *in silico* evaluation of vaccine potential of ZnuD protein in *Acinetobacter baumannii*

Maryam Mobarak Qamsari [a], Iraj Rasooli [a,b,*], Shakiba Darvish Alipour Astaneh [c]

[a] *Department of Biology, Shahed University, Tehran, Iran*
[b] *Molecular Microbiology Research Center, Shahed University, Tehran, Iran*
[c] *Department of Biotechnology, Semnan University, Semnan, Iran*

A R T I C L E   I N F O

A B S T R A C T

The nosocomial *Acinetobacter baumannii* is one of the most opportunistic pathogens that resists almost a broad range of antibiotics. Some bacterial metal acquisition systems are introduced as potential virulence factors in pathogen-host interaction. Bacterial zinc uptake (Znu) systems are critical transporters that play a major role in the survival of the pathogen and the establishment of infections. In silico characterization of ZnuD shows that it is an outer membrane protein with high conservancy rates in most *Acinetobacter baumannii* strains. The in silico derived physicochemical properties, structural and conserved immunoinformatic prediction data among *Acinetobacter baumannii* strains confirm that ZnuD belongs to Ton B dependent siderophore receptor families and is considered as a potential vaccine candidate. A more detailed interpretation of the data presented in this article is provided in "Identification and immunogenic properties of recombinant ZnuD protein loops of *Acinetobacter baumannii*" [1].

Specifications table

Subject	Medicine and Dentistry
Specific subject area	Infectious Diseases - In silico evaluation of ZnuD as a vaccine candidate against *Acinetobacter baumannii*
Type of data	Table
	Images
How data were acquired	All data were retrieved from various bioinformatics servers.
Data format	Raw and analyzed
Parameters for data collection	Data related to the secondary structure prediction were obtained from various bioinformatics tools and arranged in an appropriate framework for better presentation and comparison of structural predictions. The multiple sequence alignment between ZnuD and other *A.baumannii* strains was provided by PRALINE server to illustrate the conservancy rates.
Description of data collection	The physicochemical properties were retrieved from Protparam server at (https://web.expasy.org/protparam/). The secondary structure prediction data was assessed by Porter 5 at (http://distilldeep.ucd.ie/porter/), s2D at (http://www-mvsoftware.ch.cam.ac.uk/index.php/s2D), SPIDER2 at (http://sparks-lab.org) and The PRALINE server at (http://ibivu.cs.vu.nl/programs/pralinewww/). The conservancy rates between various *A.baumannii* strains were also done by PRALINE server.
Data source location	Molecular Microbiology Research Center, Shahed University, Tehran, Iran
Data accessibility	With the article
Related research article	Maryam Mobarak Qamsari, Iraj Rasooli, Shakiba Darvish Alipour Astaneh Identification and immunogenic properties of recombinant ZnuD protein loops of *Acinetobacter baumannii*, Informatics in Medicine Unlocked. "In Press"

Value of the data

- The data here presented are likely to be valuable in introducing basic properties of a potential vaccine candidate such as physicochemical, secondary structure and conservancy rates of a vaccine candidate among various *A.baumannii* strains.
- This study emphasizes the importance of the conserved outer membrane receptors in *A. baumannii* as appropriate targets for a protective subunit vaccine for their exposure, accessibility, and expression during the infection. These information can be used by vaccine oriented projects to develop new generations of vaccines particularly against multi drug resistant *A.baumannii*.
- The stable and conserved regions of ZnuD in the predicted protein model as a vaccine candidate, could be useful for researches in experimental vaccinology.
- There is no crystallographic structure or physicochemical analyses of this receptor in *A. baumannii*. These data can be employed as basis in the prediction of a 3D structure of ZnuD that could specify the competence of this protein as a vaccine candidate.

1. Data description

The high aliphatic and hydrophobicity index of ZnuD protein causes a robust resistance of the protein against high temperatures and denaturalizing solutions such as urea [2]. They are considered as more favored options in protein fusion technology [3]. The physicochemical prop-

Table 1
Physicochemical properties of ZnuD predicted by the Protparam server.

Number of amino acids	684
Molecular weight	76.147 kDa
Theoretical pI	4.94
Total number of negatively charged residues (Asp+ Glu)	101
Total number of positively charged residues (Arg + Lys)	59
Formula	C 3387H5112N926O1070S7
Total number of atoms	10,502
Extinction coefficients	96,735
Instability index	23.27 (stable)
Aliphatic index	74.83
The estimated half-life	
	30 h (mammalian reticulocytes, *in vitro*).
	>20 h (yeast, *in vivo*).
	>10 h (*Escherichia coli*, *in vivo*).
Atomic	Carbon C
composition	Hydrogen H
	Nitrogen N
	Oxygen O
	Sulfur S

erties of ZnuD protein sequence were predicted by ProtParam server (Table 1). General sequence specifications of the protein sequence such as the number of amino acids, molecular weight, theoretical isoelectric point, amino acids, atomic composition, extinction coefficient, instability index, aliphatic index, and grand average of hydropathicity were calculated as presented in Table 1. The prediction of the secondary structure of ZnuD is shown in Fig. 1. The secondary structure prediction was carried out by 4 servers *viz.*, Porter5, S2D, SPIDER2, and PRALINE. Datasets of all the servers were arranged and the position of three classes of secondary structures (Helix, Coil, and Extended strands) was compared to each other as illustrated in Fig. 1. The conservancy rates between ZnuD protein sequence with 202 high blast score with identity ≥95%, query coverage:100%, and Evalue:0, from various *A.baumannii* strains, were done by multiple sequence alignment in PRALINE server and illustrated in Fig. 2. The consistency rates were shown from low conservative residues in blue to high conservative residues in red against various *A.baumannii* strains. Table 1 shows the physicochemical properties of ZnuD predicted by the Protparam server.

Materials and methods

The physicochemical properties, secondary structure prediction, and conserved residues of ZnuD were evaluated by in silico approaches. The ProtParam server at (https://web.expasy.org/protparam/) was employed to identify the general properties of the ZnuD sequence such as the number of amino acids, molecular weight, theoretical isoelectric point, amino acid, atomic composition, extinction coefficient, instability index, aliphatic index, and grand average of hydropathicity. The secondary structure prediction approach was done by employing 4 servers to predict the position of 3 class of Helix, Extended strands, and coils in ZnuD Structures. Porter 5 at (http://distilldeep.ucd.ie/porter/), s2D at (http://www-mvsoftware.ch.cam.ac.uk/index.php/s2D), SPIDER2 at (http://sparks-lab.org) and PRALINE at (http://www.ibi.vu.nl/programs/pralinewww/) server was used to provide the secondary structure data. The PRALINE server was also used for multiple sequence alignment in the detection of the conservancy rate between the ZnuD sequence and 202 high score *A.baumannii* strains retrieved from the blast. The strains with identity ≥95%, query

```
Ruler      .........11........21........31........41........51........61........71........
Sequence   MLFYKNILTLSILAVISIPVFAAENENVEKLETIRIKAHPLEQTSKDFAVADTVVDQKHLTEGAATIGDALNSEVGIYAN
Porter5    CCCHHHHHHHHHHHHHHHHHHCCCCCCCCCEEEEECCCCCCCCCCCCCCEEEECHHHHHCCCCHHHHHHCCCCEEEE
s2D        CCCCCHHHHHHHHHHCCHHHHCCCCCCCCCEEEEEEECCCCCCCCCCCCCCEEEECHHHHHCCCHHHHHHHHHCCCCCEEC
SPIDER2    ------HHHHHHHHH--HHHH------EE---EEE-------HHH----EEE-HHHHH-----HHHH H----EEE--
PRALINE    ---HHHHHHHHHHHHHHH----------EEEE-------HHH----EEEE-HHHHHH----HHHH HH----EEEE

Ruler      .........91........101.......111.......121.......131.......141.......151.......
Sequence   QFGAGSSRPVIRGQDGPRVKVLQNSSENVDVSTLSPDHAVTVDPVLAKQVEVIRGPSTLLFGAGTVGGLVNVIDNKIPTQ
Porter5    CCCCCCCCEEECCCCCCEEEEECCCCCCCCCCCCCCCCCCCHHHCEEEEEEECCCHHHHCCCCCCEEEEEEECCCCCC
s2D        CCCCCCCCEEECCCCCEEEEECCCCCCCCCCCCCCCCCCCCCHCHHEEEEEECCCCHHHHCCCCCEEEEEEECCCCCC
SPIDER2    -------EEE------EEEEEE---------H-----E----HHH-EEEEEEE-HHHH-------EEEEE-----E---
PRALINE    -------EEEE------EEEEE------------------HHH-EEEEEEE--HHH--------EEEEEEE------

Ruler      .........171.......181.......191.......201.......211.......221.......231.......
Sequence   MPENGYEGQVGLRYNTGSDEKLASAGVTVGLGSQVALRVEGLTRDANNYIAPNYIYEGEKERRVDNTFAQGDSVNVGLSW
Porter5    CCCCCEEEEEEEEECCCCCEEEEEEEEECCCEEEEEEEEEEECCCCECCCCEECCCEEEEECCCCCCCCEEEEEEE
s2D        CCCCCEEEEEEEEECCCCCEEEEEEEEECCCEEEEEEEEEEECCCCCECCCCCCCCCCCCCCCCCCCCCEEEEEEE
SPIDER2    ------EEEEEEEEEE-----EEEEEEEE-----EEEEEEEEE-----E-----------E---------EEEEEEE
PRALINE    -----EEEEEEEEEE----EEEEEEEEE----EEEEEEEEEE---EE----------------------EEEEEEE

Ruler      .........251.......261.......271.......281.......291.......301.......311.......
Sequence   IYDRGYTGVSYSNRRDQYGLPGHSHEYETCHIHGLSLHCGDHEEGEEGHDHEHEEHEHGGPWIDLKSERYDFKTELNDPF
Porter5    EECCCEEEEEEEEECCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCEEECCCCEEEEEEEEEEEEEEEECCC
s2D        EECCCEEEEEEECCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCEEEEHHHHEEEEEEEEECCC
SPIDER2    EE--EEEEEEEE-------E------------------------H------EEEEE-EEEEEEEE---
PRALINE    EE---EEEEEEE-------------------------------EE-----EEEEEEEEEEEEE---

Ruler      .........331.......341.......351.......361.......371.......381.......391.......
Sequence   AGFQKLRAQASYTDYQHDEIEEGTIATRFQNKGYDGRIELVHNPIASWEGVIGAQLGQQKLDLTGEEAFMAPTTTKKWSV
Porter5    CCCEEEEEEEEEEEECCEEECCEEEEEEEEEEEEEEEEEECCCCCEEEEEEEEEEEEEECCCCCCCCCCEEEEEEEE
s2D        CCHCHCEHEECCCCCCCCCCCEEEEEECHEEHHEEEEECCCCCEEEEEEEECCCCCCCCCCCCCCCCCEEEEHE
SPIDER2    -E-EEEEEE-EE-E-EEEEE--EEEEEEEEEEEEEEE----EEEEEEEEEEEE---------EEEEEEE
PRALINE    -HHEEEEEE--------EEEE-----EEEEEE---EEEEEE------EEEEEEEEEEEEEEE-----------EEEEEE

Ruler      .........411.......421.......431.......441.......451.......461.......471.......
Sequence   FALEHKQWKDVHFELSARADQQEIDVDDNSKQDFDGSAFSYAGAANWEFAPNYKLSFVASHQERLPLAQELYANGGHFAT
Porter5    EEEEEEEECCEEEEEEEEEEEEECCCCCCCCCCEEEEEEEEEECCEEEEEEEEEECCHHHHHCCCCCCCC
s2D        EHEEEECCCCEEEEEEEEEECEEEEECCCCCCCCCCCEEEEEEEEECCCEEEEEEEEEECCCHHHHHCCCCCCC
SPIDER2    EEEEEE---EEEEEEEEEEEEEE-------------EEEEEEEEEE---EEEEEEEEE-E---HHHHH--------
PRALINE    EEEEEEE--EEEEEEEEEEEE-----------------EEEEEEE---EEEEEEEEEE----HHHH---------

Ruler      .........491.......501.......511.......521.......531.......541.......551.......
Sequence   NTYELGNDQLGKEKSNNVELGLHFDNDKLDYHLHVYHNWFDDYIYAQTLDRYKDFRLVQYTQDKARFYGAEGEIGYQITP
Porter5    CCEECCCCCCCCCEEEEEEEEEECCCEEEEEEEEEEECCCEEEEEEEECCCCCEEEEEEECCEEEEEEEEEEEEEECC
s2D        EEEEECCCCCCCCEEEEEEEEEEECCCEEEEEEHEEHECCCCCEECCCCCCCEEEEEEEEEEEEEEEEEEEEEEHHH
SPIDER2    --EE---------EEEEEEEEEE---EEEEEEEEEE--EEEEEE---E--E-EEEE---EEEEEEEEEEEEEE
PRALINE    --EE-------EEEEEEEEEE---EEEEEEEEE---EEEEEE-----EEEEEE---EEEEEEEEEEEEE--

Ruler      .........571.......581.......591.......601.......611.......621.......631.......
Sequence   MYKISAFGDYVRGKIDAEGNAPRIPAGRLGTKVDADFGDGFSGSAEYYHVFNQDKIAAYETETEGYNMLNLGVAYSGQYG
Porter5    CEEEEEEEEEEEEEECCCCCCCCCCCEEEEEEEEEECCCEEEEEEEEEEEECCCCCCCEEECCEEEEEEEEEEEEEECC
s2D        HEEEEEEEECCCCCCCCCCCCCCCEEEEEEEECCCEEEEEEHEEHECCCCCCCEEECCCCCCCEEEEEEEEEEEEEEC
SPIDER2    -EEEEEEEEEEEEE-------E---EEEEEEEEE---EEEEEEEEEE------------EEEEEEEEEE--
PRALINE    -EEEEEEEEEEEEE-----------EEEEEEEEE---EEEEEEEEEEE------EE---EEEEEEEEEEEE---

Ruler      .........651.......661.......671.......684
Sequence   AKTDYRVYLKANNLLDDTVYQHASFLSNIPQVGRNFTVGVDFS
Porter5    CCCCEEEEEEEECCCCCEEEECCCCCCCCCCCCNFTVGVDFS
s2D        CCCEEEEEEEEHCCCCCHCECCCHHHHCCCCCCCEEEEEEEEC
SPIDER2    ----EEEEEEEE---HHHHHH------------EEEEEEEE-
PRALINE    ----EEEEEEEE----HHHHH------------EEEEEEEE-
```
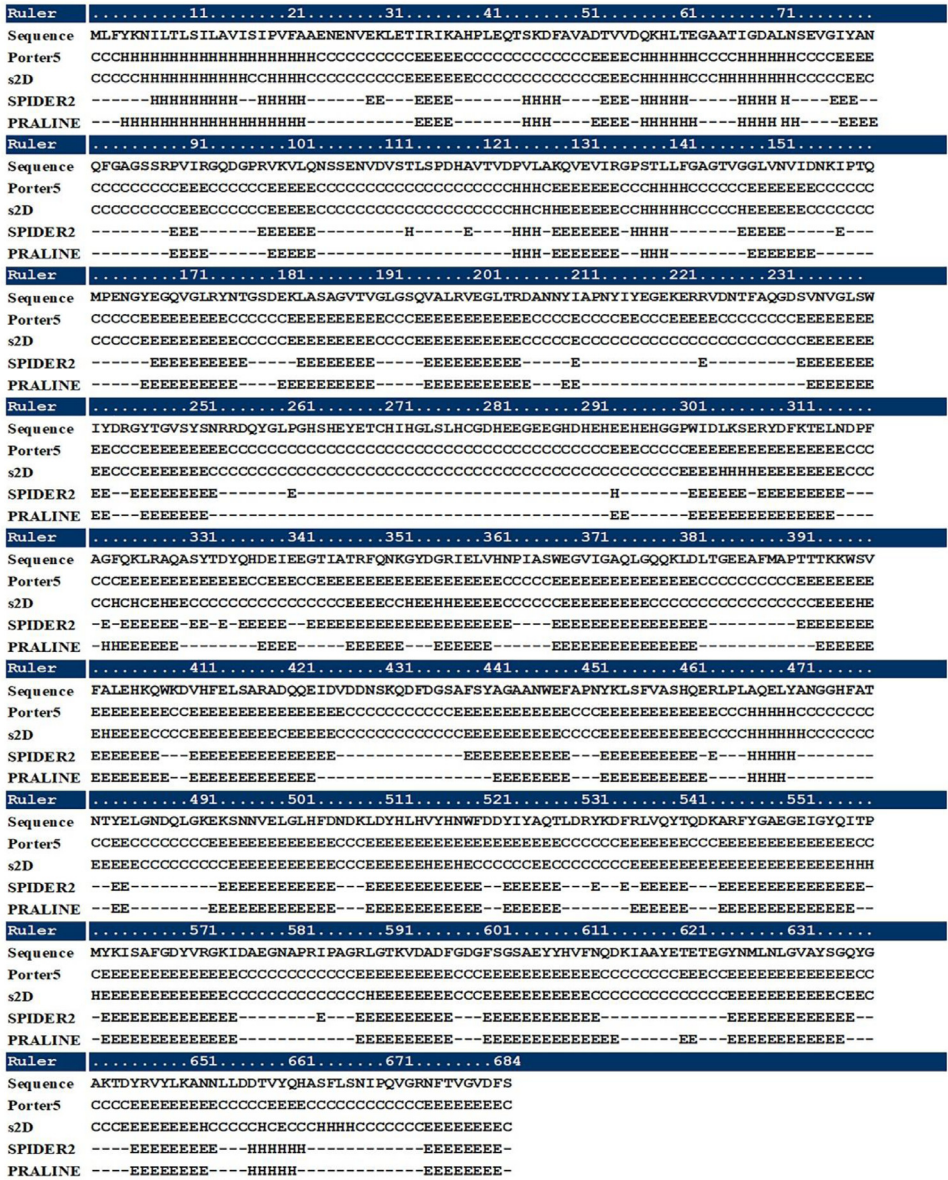
Fig. 1. Secondary structure prediction of ZnuD by Porter5, S2D, SPIDER2, and PRALINE servers. H: helix; E: extended strand; C (dash or blank): Coil.
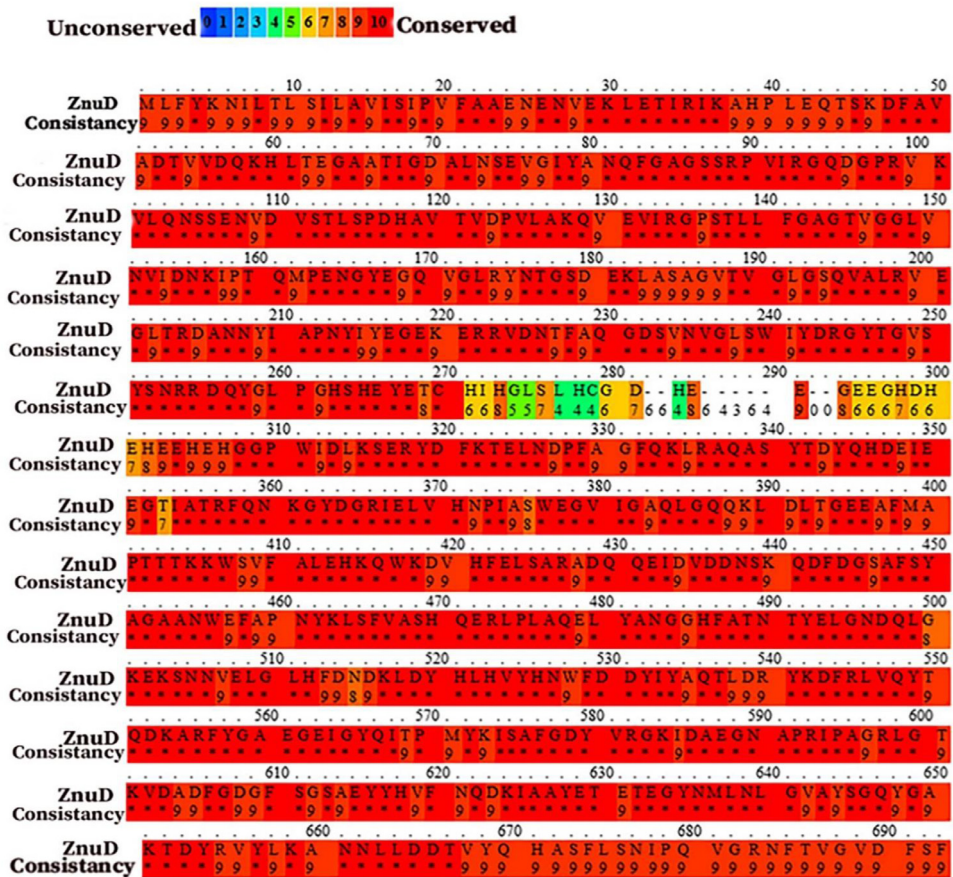
Fig. 2. Multiple sequence alignments of ZnuD with 202 reference sequences from different *A.baumannii* strains were made by the PRALINE server and adjust manually. These sequences contain the highest score (identity ≥95%, query coverage:100%, and Evalue:0) in ZnuD Blast and are selected for estimation of consistency rates. Residues conservancy is depicted from unconserved (blue) to conserved residues (red).

Declaration of Competing Interest

The authors declare no conflict of interest.

Acknowledgement

References

[1] M. Mobarak Qamsari, I. Rasooli, S.D.A. Astaneh, Identification and Immunogenic Properties of Recombinant ZnuD Protein Loops of Acinetobacter baumannii, Inform. Med. Unlocked, 2020.

[2] Ikai Atsushi, Thermostability and aliphatic index of globular proteins, J. Biochem. 88 (1980) 1895–1898 https://doi.org/, doi:10.1093/oxfordjournals.jbchem.a133168.

[3] A. Thapa, M. Shahnawaz, P. Karki, G. Raj Dahal, M. Golam Sharoar, S. Yub Shin, J. Sup Lee, B. Cho, I.-.S. Park, Purification of inclusion body forming peptides and proteins in soluble form by fusion to Escherichia coli thermostable proteins, BioTechniques 44 (2008) 787–796 https://doi.org/, doi:10.2144/000112728.