Journal of Engineering and Applied Sciences 12 (5): 1152-1158, 2017
ISSN: 1816-949X

Identification of Functionally and Structurally Important Residues of HutA Protein in *Vibrio cholera*

Hamed Karamizade, Fateme Sefid, Nazgol Emamian, Vahide Saeidjavan and Sepide Akhgari
Departeman of Biology, Science and Art University, Yazd, Iran

Abstract: Iron is essential for the growth of most bacteria but in nature the element is highly insoluble in an aerobic environment and therefore unavailable to most organisms. Inside the human body, most iron is in the cell in the form of hemoglobin or other iron-containing proteins or is stored as ferritin. Trace amounts of iron are found outside the cell complexed to high-affinity iron-binding proteins such as lactoferrin or transferrin. Inside the human body, most iron is in the cell in the form of hemoglobin or other iron-containing proteins or is stored as ferritin. *Vibrio cholera*, the intestinal pathogen that causes the disease cholera can acquire iron in two ways. Under low-iron conditions, the organism synthesizes and secretes the siderophore vibriobactin which binds ferric iron. To survive in such conditions, *Vibrio cholera* expresses HutA in its outer membrane. Evidence suggests that HutA is a useful antigen for inclusion in an effective vaccine hence, the identification of its structure and functionally and structurally important residues is very important.

Key words: *Vibrio cholera*, HutA, bioinformatic, secrets, bacteria

INTRODUCTION

Vibrio cholerae, a rod-shaped, highly motile, gram-negative bacterium is the causative agent of the life threatening diarrheal disease cholera (Reidl and Klose, 2002). Iron is essential for the growth of most bacteria but in nature the element is highly insoluble in an aerobic environment and therefore unavailable to most organisms (Sandkvist *et al*., 1997). Inside the human body, most iron is in the cell in the form of hemoglobin or other iron-containing proteins or is stored as ferritin. Trace amounts of iron are found outside the cell complexed to high-affinity iron-binding proteins such as lactoferrin or transferrin (Cianciotto, 2005).

The scarcity of free iron inside the body makes it difficult for pathogens to obtain sufficient amounts of the element to survive and multiply in the host. To acquire iron from their environment, bacteria have evolved specialized systems that permit them to scavenge iron from their environment or to directly utilize host iron-containing proteins as iron sources. For example, many bacteria synthesize and secrete siderophores, low-molecular-weight compounds which bind ferric iron (Fe3) (Fullner and Mekalanos, 1999). Certain siderophores bind Fe3" with sufficient affinity to remove the element from host proteins such as transferrin (Marsh and Taylor, 1998). A number of bacteria directly utilize transferrin and lactoferrin or hemoglobin as iron sources. Many iron transport systems characterized to date involve iron-regulated outer membrane receptors which bind a specific iron-containing compound and facilitate transport of the iron into the cell. The expression of many iron-regulated receptors is controlled at the transcriptional level by an iron-binding repressor protein called Fur (ferric uptake regulation) (Sandkvist *et al*., 1997).

Under conditions of iron sufficiency, Fur binds to a highly conserved region called the Fur box upstream of iron-regulated genes and blocks transcription (Filloux, 2004). Under low-iron conditions, repression by Fur is relieved and the genes are transcribed. Another common feature of many iron transport systems in *Escherichia coli* and other gram-negative bacteria is their dependence on the TonB inner membrane protein which provides energy for transport of the ligand across the outer membrane (Hirst and Holmgren, 1987). TonB can span the periplasmic space and physically interact with the outer membrane receptor in a highly conserved region called the TonB box (Howard *et al*., 1993). This interaction is thought to lead to a conformational change in the receptor protein, permitting transport of the ligand across the outer membrane into the periplasmic space. *Vibrio cholera*, the intestinal pathogen that causes the disease cholera, can acquire iron in two ways. Under low-iron conditions, the organism synthesizes and secretes the siderophore vibriobactin (Jiang and Howard, 1992; Sandkvist *et al*., 1997; Shi *et al*., 2008) which binds ferric iron. Ferric vibriobactin then binds to the vibriobactin receptor, ViuA, a 74 kDa iron-regulated protein (Sandkvist *et al*., 1997) to allow transport of iron into the cell. Transcription of viuA is controlled by the *V. cholerae* Fur protein which exhibits 76% amino acid homology with the *E. coli* Fur protein (Sikora *et al*., 2007). ViuA contains two predicted Fur boxes in its promoter operator region (Howard *et al*., 1993). Analysis of the

Corresponding Author: Fateme Sefid, Departeman of Biology, Science and Art University, Yazd, Iran

predicted amino acid sequence of ViuA indicates that the protein has limited homology with TonB dependent outer membrane proteins in *E. coli* and other bacteria (Shi *et al*., 2008). It should be mentioned that *V. cholerae* IrgA, aniron-regulated outer membrane protein of unknown function is homologous to a number of TonB-dependent receptors, suggesting that transport systems related to the TonB-dependent systems in *E. coli* are present in *E. cholerae*, *V. cholerae* can also acquire iron from heme or hemoglobin in a siderophore-independent fashion (Sikora *et al*., 2007).

New genome analysis tools based on bioinformatics and immunoinformatics approaches help us select suitable antigens or epitopes directly from the genomes of pathogens in order to design a vaccine. These tools could be employed for epitope selection and vaccine design. Moreover, prediction of protein structures is one of their wide applications (Kafee and Sefid, 2016a, b; Masoumi and Sefid, 2016; Payandeh *et al*., 2015; Sefid *et al*., 2013, 2015, 2016).

Evidence suggests that HutA is a useful antigen for inclusion in an effective vaccine, hence the identification of its structure and functionally and structurally important residues is very important. In order to explore the potential application of HutA as a vaccine candidate, the present study deals with the bioinformatics analyses of HutA structute and identification of functionally and structurally important residues in *Vibrio cholera*.

MATERIALS AND METHODS

Sequence availability: The HutA protein sequence with Accession number CSA70573.1 and GI 903538441 obtained from NCBI at http://www.ncbi.nlm.nih.gov/protein was saved in FASTA format for further analyses.

Homology search: The HutA sequence served as a query for BLAST at http://blast.ncbi.nlm.nih.gov/Blast.cgi against non-redundant protein database. Probable putative conserved domains of the query protein were also searched for at the above address.

Template search: The query protein sequence was used as an input data for the PSI-BLAST against Protein Data Bank (PDB) at http://blast.ncbi.nlm.nih.gov/Blast.cgi to identify its homologous structures.

Secondary structure prediction: Secondary structure of the protein was predicted by Phyre2 at http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index

D modeling: Phyre2 at http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index uses the alignment of hidden Markov models via HHsearch to significantly improve accuracy of alignment and detection rate. Phyre2 also incorporates a new ab initio folding simulation called poing to model regions of proteins with no detectable homology to known structures.

Phyre2 is a major update to the original Phyre server. A range of new features have been included, accuracy has been substantially improved and the interface has been redesigned to be more intuitive and powerful. Poing is also used to combine multiple templates. Distance constraints from individual models are treated as linear elastic springs. Poing then synthesises your entire protein in the presence of these springs and at the same time models unconstrained regions using its physics simulation.

Ligand binding site predictions: Cofactor at http://zhanglab.ccmb.med.umich.edu/COFACTOR/ is a structure-based method for biological function annotation of protein molecules. Important amino acid involved in ligand binding site is predicted by this server.

Identification of functionally and structurally important residues: InterProSurf at http://curie.utmb.edu/pattest9.html predicting functional sites on protein surface using patch analysis was employed. HutA 3D structure determined in a previous study, served as an input file for this server.

Single-scale amino acid properties assay: IEDB at http://tools.immuneepitope.org/tools/bcell/iedb_input parameters such as hydrophilicity, flexibility, accessibility, turns and antigenic propensity of polypeptide have been correlated with the location of B cell epitopes. This has led to a search for empirical rules that would allow the position of B cell epitopes to be predicted from certain features of the protein sequence.

RESULTS AND DISCUSSION

Sequence availability and homology search: The protein sequence with 693 residue obtained from NCBI and saved in FASTA format. Protein sequence serving as query for BLAST produced a set of sequences as the highest similar sequence. BLAST search revealed numerous hits to the HutA subunit sequence. All hits were of vibrio. Putative conserved domains were detected within this sequence and are shown in Fig. 1.

Template search: PSI-BLAST Against Protein Data Bank (PDB) results displayed several hits as homologous structures. The first hit possessing the highest score was selected as a template for modelling. The first hit (Description: Chain A, Crystal structure of the

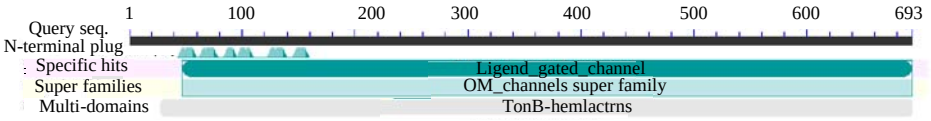

Fig. 1: Putative conserved domains have been detected. Most of the sequences belong to OM-channels super family

Fig. 2: Phyre2 secondary prediction

HEMOGLOBIN OUTER membrane Transporter Shua From Shigella Dysenteriae, Accession: 3FHH-A, Max score: 133, Total score: 133 Query coverage: 90%, e value: 3e-32, Max ident: 25%) possessing the highest score was selected as a template.

Secondary structure prediction: Secondary structure of the proteins was predicted by phyre2 server. Coil , helix and strands are components constituting secondary structure of the proteins. The secondary structure could be used to validate the tertiary structures. Phyre2 secondary prediction result is shown in Fig. 2.

3D structure prediction: Building a homology model comprises four main steps: identification of structural template(s) alignment of target sequence and template structure(s) model building and model quality evaluation. These steps can be repeated until a satisfying modelling result is achieved. Each of the four steps requires specialized software and access to up-to-date protein sequence and structure databases. Swiss model software recruited for homology modeling introduced 1 model. Predicted model is shown in Fig. 3.

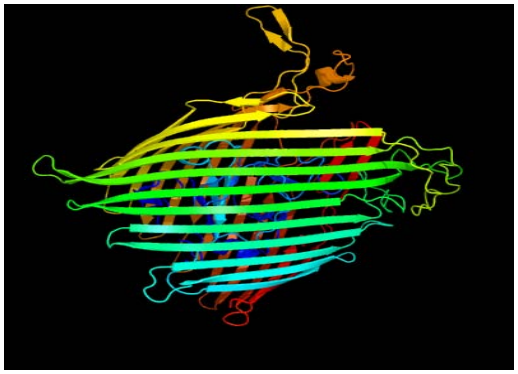

Fig. 3: 3D structure of HutA

Identification of functionally and structurally important residues: Interprosurf annotated functional residues on the 3D structure of HutA. Residues Predicted by Auto Patch Analysis are: 279, 280, 281, 24, 561, 562, 600, 601, 602, 606, 607, 506, 556, 557, 558, 559, 564, 598, 299, 363, 366, 367, 386, 387, 388, 389, 365, 368, 369, 242, 244, 246, 261, 263, 297 (Fig. 4).

Ligand binding site predictions: Ligand binding sites determined using Cofactor software, indicate involvement of conserved residues include 61, 66 and 68 in binding site with the highest CscoreLB (the confidence score of predicted binding site). The calculated BS-score for this predicted binding site was 1.63.BS-score is a measure of local similarity (sequence and structure) between template binding site and predicted binding site in the query structure. Based on large scale benchmarking analysis, observed that a BS-score >1 reflects a significant local match between the predicted and template binding site (Fig. 5, Table 1).

Single-scale amino acid properties assay: IEDB server predict several properties such as hydrophilicity, accessibility, antigenicity, flexibility and beta turn secondary structure in the protein sequence. Propensity scale methods assign a propensity value to each amino acid which measures the tendency of an amino acid to be part of a B-cell epitope (as compared to the background). To reduce fluctuations, the score for each target amino acid residue in a query sequence is computed as the average of the propensity values of the amino acids in a sliding window centered at the target residue. hydrophilicity, accessibility, antigenicity, flexibility and secondary structure properties have fundamental role in B cell epitope prediction. Relying on just one of these properties, reliable results could not be achieved (Fig. 6).

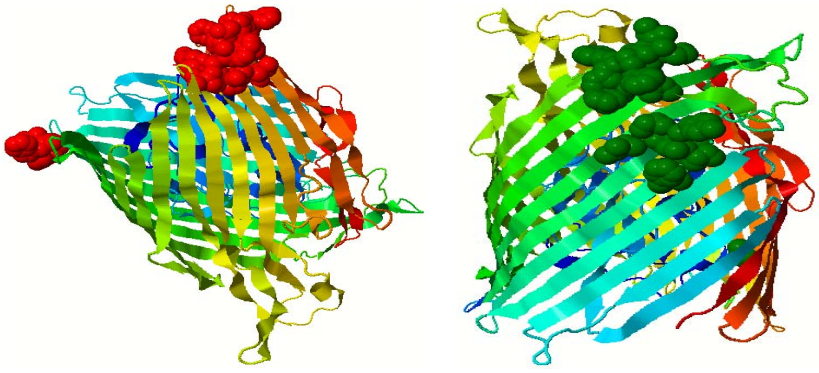

Fig. 4: Continue

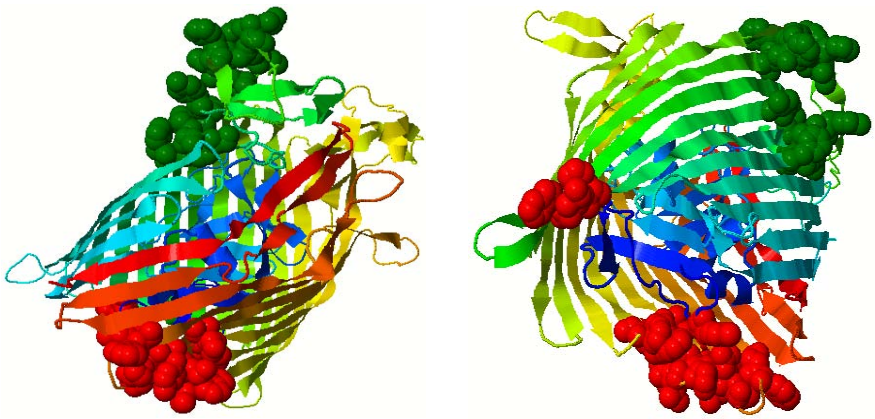

Fig. 4: Functional residues on HutA 3D structure

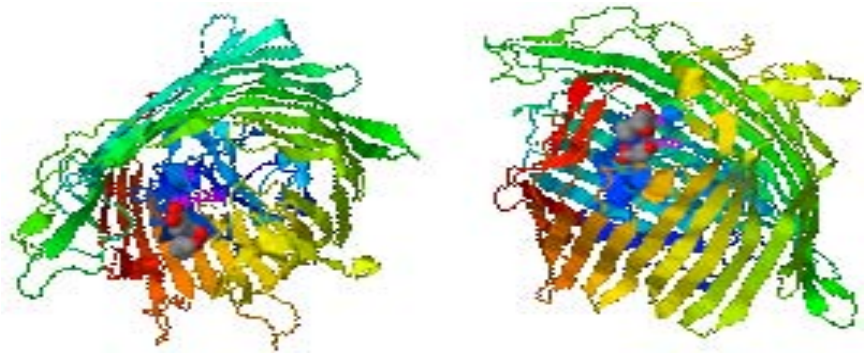

Fig. 5: Cofactor ligand binding site prediction involvement of conserved residues in binding site

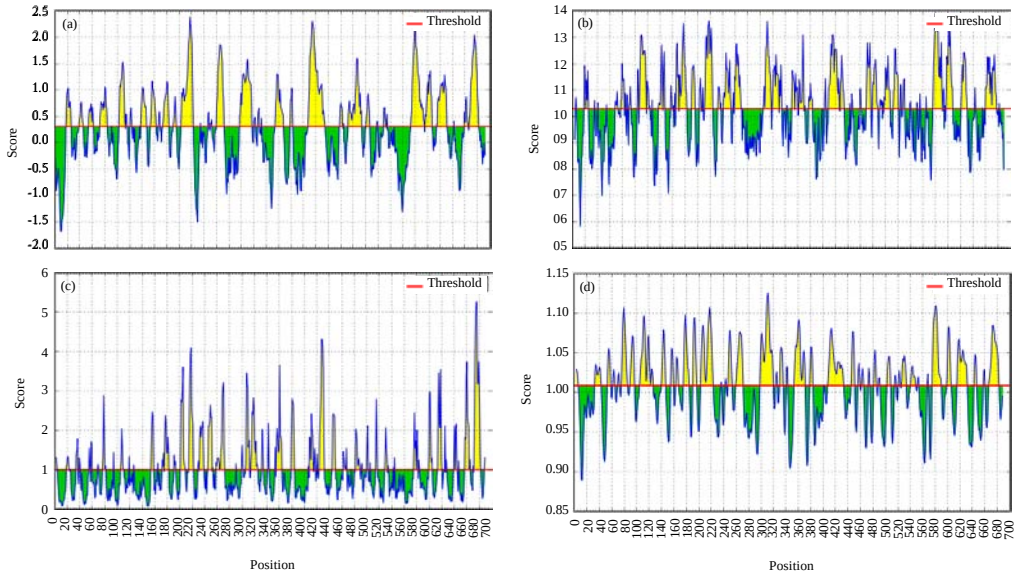

Fig. 6: Continue

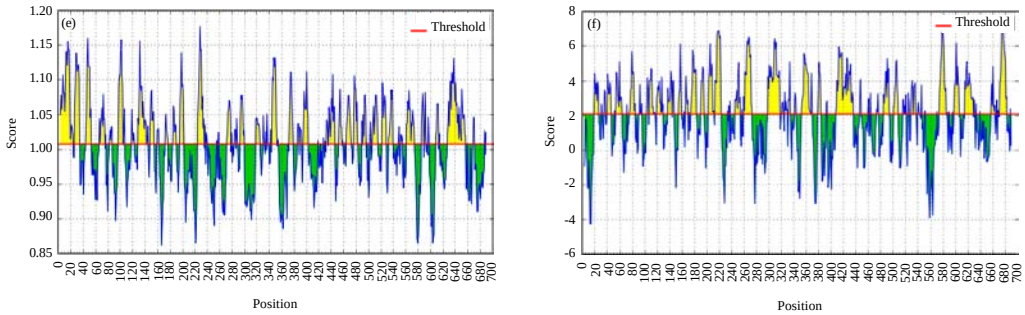

Fig. 6: IEDB linear B cell epitope prediction results for HutA protein: a) hydrophilicity; b) accessibility; c) antigenicity; d) flexibility; e) beta turn and f) Bepipred linear epitope

Table 1: Co factor results for huta 3D structure

Rank	$Cscore^{LB}$	PDB hit	TM-score	$RMSD^{a}$	$IDEN^{a}$	Cov.	BS-score	Lig. name	Predicted binding site residues
1	0.04	2w78B	0.949	0.95	0.144	0.959	1.63	PVE	61, 66, 68
2	0.04	3rgmA	0.738	2.76	0.177	0.791	1.01	MTN	438, 439, 440, 442, 477
3	0.04	3qlbA	0.879	2.38	0.137	0.930	0.53	Mul.part	66, 67, 91, 102, 303, 305, 340
4	0.04	1ujwA	0.758	3.06	0.177	0.821	0.55	Mul.part	424, 436, 477, 479
5	0.01	1a0tP	0.380	5.49	0.076	0.481	0.67	CA	122, 125, 128, 130, 441

$Cscore^{LB}$ is the confidence score of predicted binding site. $Cscore^{LB}$ values range in between [0-1]; where a higher score indicates a more reliable ligand-binding site prediction; BS-score is a measure of local similarity (sequence and structure) between template binding site and predicted binding site in the query structure. Based on large scale benchmarking analysis, we have observed that a BS-score >1 reflects a significant local match between the predicted and template binding site; TM-score is a measure of global structural similarity between query and template protein; $RMSD^{a}$ the RMSD between residues that are structurally aligned by TM-align; $IDEN^{a}$ is the percentage sequence identity in the structurally aligned region; Cov. represents the coverage of global structural alignment and is equal to the number of structurally aligned residues divided by length of the query protein

CONCLUSION

To survive in such conditions, *Vibrio cholerae* expresses HutA in its outer membrane. Evidence suggests that HutA is a useful antigen for inclusion in an effective vaccine hence, the identification of its structure and functionally and structurally important residues is very important.

REFERENCES

Cianciotto, N., 2005. Type II secretion: A protein secretion system for all seasons. Trends Microbiol., 13: 581-588.

Filloux, A., 2004. The underlying mechanisms of type II protein secretion. Biochim. Biophys. Acta Mol. Cell Res., 1694: 163-179.

Fullner, K.J. and J.J. Mekalanos, 1999. Genetic characterization of a new type IV-A pilus gene cluster found in both classical and El tor biotypes of Vibrio cholera. Infect. Immun., 67: 1393-1404.

Hirst, T.R. and J. Holmgren, 1987. Conformation of protein secreted across bacterial outer membranes: A study of enterotoxin translocation from Vibrio cholera. Proc. National Acad. Sci., 84: 7418-7422.

Howard, S.P., J.E.F.F. Critch and A.N.I.L. Bedi, 1993. Isolation and analysis of eight exe genes and their involvement in extracellular protein secretion and outer membrane assembly in Aeromonas hydrophila. J. Bacteriol., 175: 6695-6703.

Jiang, B. and S.P. Howard, 1992. The Aeromonas hydrophila exeE gene, required both for protein secretion and normal outer membrane biogenesis is a member of a general secretion pathway. Mol. Microbiol., 6: 1351-1361.

Kafee, N. and F. Sefid, 2016. Functional exposed amino acids of OSPA as a candidate for lyme disease vaccine. Intl. J. Adv. Biotechnol. Res., 7: 957-966.

Kafee, N. and F. Sefid, 2016. In silico identification and validation of OSPA 3D structure and its topology as a candidate for lyme disease vaccine. Intl. J. Adv. Biotechnol. Res., 7: 949-956.

Marsh, J.W. and R.K. Taylor, 1998. Identification of the Vibrio cholera type 4 prepilin peptidase required for cholera toxin secretion and pilus formation. Mol. Microbiol., 29: 1481-1492.

Masoumi, S. and F. Sefid, 2016. In silico determination and validation of FptA structure and ligand binding site as a vaccine candidate in Pseudomonas aeruginosa. Intl. J. Adv. Biotechnol. Res., 7: 1038-1045.

Payandeh, Z.A.H.R., A.S.E.F. Kofeiti and F.A.T.E. Sefid, 2015. Nanobody structure analysis and determination of the functional conserve amino acid with bioinformatic tools. Natl Agronomique Institute Ltd, France.

Reidl, J. and K.E. Klose, 2002. Vibrio cholera and cholera: Out of the water and into the host. FEMS Microbiol. Rev., 26: 125-139.

Sandkvist, M., L.O. Michel, L.P. Hough, V.M. Morales and M. Bagdasarian *et al*., 1997. General secretion pathway (eps) genes required for toxin secretion and outer membrane biogenesis in Vibrio cholera. J. Bacteriol., 179: 6994-7003.

Sefid, F., I. Rasooli and A. Jahangiri, 2013. In silico determination and validation of baumannii acinetobactin utilization a structure and ligand binding site. BioMed Res. Intl., 2013: 1-14.

Sefid, F., I. Rasooli and Z. Payandeh, 2016. Homology modeling of a Camelid antibody fragment against a conserved region of Acinetobacter baumannii biofilm associated protein (Bap). J. Theor. Biol., 397: 43-51.

Sefid, F., I. Rasooli, A. Jahangiri and H. Bazmara, 2015. Functional exposed amino acids of BauA as potential ImmunoGen against Acinetobacter baumannii. Acta Biotheor., 63: 129-149.

Shi, L., S. Deng, M.J. Marshall, Z. Wang and D.W. Kennedy *et al*., 2008. Direct involvement of type II secretion system in extracellular translocation of Shewanella oneidensis outer membrane cytochromes MtrC and OmcA. J. Bacteriol., 190: 5512-5516.

Sikora, A.E., S.R. Lybarger and M. Sandkvist, 2007. Compromised outer membrane integrity in Vibrio cholera type II secretion mutants. J. Bacteriol., 189: 8484-8495.