ARTICLE

STRUCTURE ENGINEERING OF FHUA AS A VACCINE CANDIDATE IN *ESCHERICHIA COLI*

Navid Farahmandian[1], Mohammad Seifipour[2*], Fateme Sefid[1]

[1]*Departeman of biology, Science and Art University, Yazd, IRAN*

[2] *Department of Biotechnology, Faculty of Biotechnology, Semnan University, Semnan, IRAN*

ABSTRACT

FhuA which is an integral membrane protein of Escherichia coli has the following functions: a) as the ferric hydroxamate uptake receptor which actively transports the siderophores ferrichrome and ferricrocin, the cyclic peptide antibiotic microcin J25, the siderophore-antibiotic conjugate albomycin, and the bacterial toxin colicin M across the outer membrane. FhuA also transports the semi-synthetic rifamycin derivative CGP 4832, although the chemical structure of this antibiotic differs markedly from that of ferric hydroxamates. b) as the primary receptor for bacteriophages T1, T5, φ80, and UC-1. The transport is coupled to the proton motive force, which energizes FhuA through the inner-membrane protein TonB Here we describe the topology and 3D structure of a novel antigen which was discovered by mining the bacterial genome and that is very effective in inducing bactericidal antibodies. This antigen is a very good candidate for inclusion in universal vaccines against Escherichia coli.

KEY WORDS
Escherichia coli, FhuA, Topology and 3D structure

Published: 16 October 2016

INTRODUCTION

The uptake of antimicrobial agents across the outer membrane of gram-negative bacteria is mediated by a family of transport proteins employing a variety of mechanisms [1].

FhuA which is an integral membrane protein of *Escherichia c*oli has the following functions: a) as the ferric hydroxamate uptake receptor which actively transports the siderophores ferrichrome and ferricrocin, the cyclic peptide antibiotic microcin J25, the siderophore-antibiotic conjugate albomycin, and the bacterial toxin colicin M across the outer membrane [2]. FhuA also transports the semi-synthetic rifamycin derivative CGP 4832, although the chemical structure of this antibiotic differs markedly from that of ferric hydroxamates. b) as the primary receptor for bacteriophages T1, T5, φ80, and UC-1.

The transport is coupled to the proton motive force, which energizes FhuA through the inner-membrane protein TonB [3, 4, 5]. The energy-transducing TonB-ExbB-ExbD complex couples the proton motive force of the cytoplasmic membrane to a family of diverse outer-membrane proteins, the TonB-dependent transporters [2].

Structural studies represented that albomycin is actively transported across both the outer and cytoplasmic membranes, indeed it is one of the most potent antibiotics against *E. coli* (minimal inhibitory concentration [MIC] of 0.005 μg/ml).The three-dimensional struture of FhuA in complex with albomycin also confirmed that this antibiotic occupies the same ligand binding site as ferrichrome, ferricrocin, and phenylferricrocin. This similarity extends to the set of residues that are involved in ligand binding and which are essentially conserved among these hydroxamate-type siderophores, and it thereby provides a structural explanation for high-affinity binding. Moreover, this structural information provides a basis for the rational design of synthetic antibiotics that are actively transported by this receptor or by its homologs [5].

Here we describe the topology and 3D structure of a novel antigen which was discovered by mining the bacterial genome and that is very effective in inducing bactericidal antibodies. This antigen is a very good candidate for inclusion in universal vaccines against Escherichia coli.

METHODS

Sequence availability and homology search

The FhuA reference sequence with accession No NP_414692.1acquired from NCBI at http://www.ncbi.nlm.nih.gov/protein was saved in FASTA format for further analyses. The sequence served as a query for protein 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.

***Corresponding Author**
Email: sefid@shahed.ac.ir

Primary sequence analysis

Protparam online software at http://expasy.org/tools/protparam.html was employed for estimation and determination of properties such as molecular weight, theoretical pI, amino acid composition, total number of negatively and positively charged residues, instability index and aliphatic index.

Subcellular localization

Subcellular localization of protein was predicted by CELLO at http://cello.life.nctu.edu.tw/

Homology modeling

In the process of modeling, default restraint settings were applied, and a rigorous relaxation protocol involved 2000 simulated annealing relaxation cycles (4.4 ps stepwise warming from 0–1000 K, followed by 19.2 PS stepwise cooling back down to300 K, all done through Charm force field and charges). The loop regions geometry was corrected using MODELER/Refine Loop command. The SWISS-MODEL Workspace at http://swissmodel.expasy.org/ is a web-based integrated service dedicated to protein structure homology modelling. Secondary structure of the protein was predicted by SWISS-MODEL too. It assists and guides the user in building protein homology models at different levels of complexity.
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.

Models evaluations

All 3D models of the proteins built, were qualititatively estimated by GMQE and QMEAN4 scores.

Topology and signal peptide prediction

SPOCTOPUS at http://spoctopus.cbr.su.se/ was employed to determine protein topology and signal peptide. Prediction of the hydrophobic transmembrane regions in a protein sequence forming probable β-barrel could help determination of the 3D protein structure. Full-length protein served as input in topology prediction. PRED-TMBB at http://biophysics.biol.uoa.gr/PRED-TMBB/is a sever that predicts transmembrane β-strands in protein sequences of Gram-negative bacteria. The web-server could find the topology of the loops in addition to localize the transmembrane strands

RESULTS

Sequence availability and homology search

The FhuA protein sequence with 747 residues 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 FhuA sequence. All hits were of *Escherichia coli.* Putative conserved domains were detected within this sequence. Most of the sequences belong to OM_channels super family.

Porin superfamily. These outer membrane channels share a beta-barrel structure that differ in strand and shear number. Classical (gram-negative) porins are non-specific channels for small hydrophillic molecules and form 16 beta-stranded barrels, which associate as trimers. Maltoporin-like channels have specificities for various sugars and form 18 beta-stranded barrels, which associate as trimers. Ligand-gated protein channels cooperate with a TonB associated inner membrane complex to actively transport ligands via the proton motive force and they form monomeric, barrels. The 150-200 N-terminal residues form a plug that blocks the channel from the periplasmic end.

TonB-dependent siderophore receptor. This subfamily model encompasses a wide variety of TonB-dependent outer membrane siderophore receptors. It has no overlap with TonB receptors known to transport other substances, but is likely incomplete due to lack of characterizations. It is likely that genuine siderophore receptors will be identified which score below the noise cutoff to this model at which point the model should be updated. [Transport and binding proteins, Cations and iron carrying compounds, Transport and binding proteins, Porins] Putative conserved domains have been detected within the sequence are shown in [Fig.1].

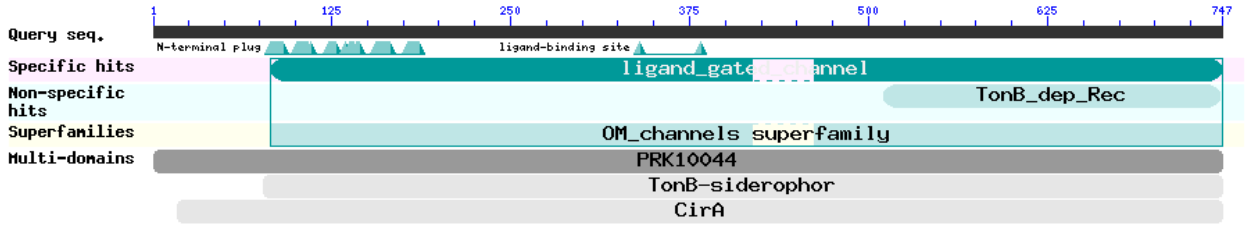

Fig.1: Putative conserved domains have been detected.

Template search

PSI-BLAST against protein data bank (PDB) result displayed several hits as homologous structures. The first hit possessing the highest score was selected as a template for homology modelling. This top hit for FhuA sequence blast results was a protein with PDB code 1BY3-A (99%identity, 95%query coverage, 1476 Max score and 1476 Total score, Chain A, Fhua From E. Coli).
Our BLAST results showed that FhuA exists in all pathogenic strains of *E.coli*. This protein antibodies cross-react with a range of *E.coli* isolates for high similarity reason.
In this regard, FhuA sequence served as a query for BLAST search against Protein Database (PDB) to find the best template for 3D structure prediction. In addition to E value, query coverage and Max. identity are also involved in Max. score definition. Lower E-value and higher query coverage and max. identity are appropriate criteria for the selection. Thus, a hit with the highest total score could be the most reliable template. The use of some sequence alignment methods to identify a relationship between the target sequence and one or more possible templates is the first step in structure prediction. Based on BLAST search and alignments generations, the predicted 3D model of the FhuA could be applied to all FhuA proteins in *E.coli*.

Primary sequence analysis

The protein sequence served as input for the computation of various physical and chemical parameters. The computed parameters included the molecular weight, theoretical pI, instability index, aliphatic index and grand average of hydropathicity (indicates the solubility of the proteins: positive GRAVY (hydrophobic), negative GRAVY (hydrophilic)) are listed below.

Number of amino acids: 747

Molecular weight: 82182.2

Theoretical pI: 5.47

Amino acid composition:

Ala (A) 69	9.2%	Leu (L) 46	6.2%
Arg (R) 33	4.4%	Lys (K) 40	5.4%
Asn (N) 43	5.8%	Met (M) 12	1.6%
Asp (D) 49	6.6%	Phe (F) 36	4.8%
Cys (C) 4	0.5%	Pro (P) 37	5.0%
Gln (Q) 33	4.4%	Ser (S) 54	7.2%
Glu (E) 35	4.7%	Thr (T) 58	7.8%
Gly (G) 62	8.3%	Trp (W) 9	1.2%
His (H) 8	1.1%	Tyr (Y) 42	5.6%
Ile (I) 17	2.3%	Val (V) 60	8.0%

Total number of negatively charged residues (Asp + Glu): 84

Total number of positively charged residues (Arg + Lys): 73

Atomic composition:

Carbon	C	3655
Hydrogen	H	5565
Nitrogen	N	987
Oxygen	O	1146
Sulfur	S	16

Formula: $C_{3655}H_{5565}N_{987}O_{1146}S_{16}$

Total number of atoms: 11369

Extinction coefficients:

Extinction coefficients are in units of M^{-1} cm^{-1}, at 280 nm measured in water.

Ext. coefficient 112330

Abs 0.1% (=1 g/l) 1.367, assuming all pairs of Cys residues form cystines

Ext. coefficient 112080

Abs 0.1% (=1 g/l) 1.364, assuming all Cys residues are reduced

Estimated half-life:

The N-terminal of the sequence considered is M (Met).

The estimated half-life is: 30 hours (mammalian reticulocytes, in vitro).

>20 hours (yeast, in vivo).

>10 hours (Escherichia coli, in vivo).

Instability index:

The instability index (II) is computed to be 30.04

This classifies the protein as stable.

Aliphatic index: 65.42

Grand average of hydropathicity (GRAVY): -0.481

Subcellular localization

FhuA Subcellular localization predicted by CELLO was OuterMembrane with the highest reliability (4.783).CELLO results are shown below [Table 1].

CELLO Prediction:	
OuterMembrane	4.783 *
Extracellular	0.124
Periplasmic	0.071
InnerMembrane	0.013
Cytoplasmic	0.008

3D structure prediction with homology medeling

The use of some sequence alignment methods to identify a relationship between the target sequence and one or more possible templates is the first step in structure prediction. Based on BLAST search and alignments generations, the predicted 3D model of the FhuA could be applied to all FhuA proteins. Accuracy of prediction depends on the degree of sequence similarity. If a structure template with sequence identity of >50% is found for a query protein, homology modeling could be chosen as the best *in silico* method with an accuracy equal to low-resolution X-ray predictions. When template and query sequences share 30-50% identity, more than 80% of the C-atoms can be expected to be within 3.5˚A of their true positions. Significant errors would occur in prediction when the sequences share less than 30% sequence identity. Since identity between the query and its template sequence was 35% (>30%) in our study, we assumed that homology modeling could be more powerful than threading.
Swiss modeler recruited for homology modeling introduced one model [Fig. 2]. Its model was selected for further scrutinizes and validation analyses.

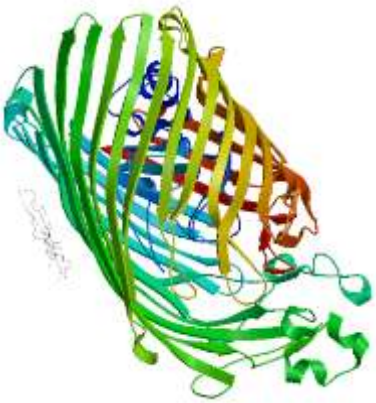

Fig.2. Swiss model 3D structure prediction.

Models evaluations

QMEAN is a composite scoring function for the estimation of the global and local model quality. QMEAN consisting of four structural descriptors: The local geometry is analyzed by a torsion angle potential over three consecutive amino acids. Two pairwise distance-dependent potentials are used to assess all-atom and C-beta interactions. A solvation potential describes the burial status of the residues. The pseudo energies returned from the four structural descriptors and the final QMEAN4 score get directly related to what we would expect from high resolution X-ray structures of similar size using a Z-score scheme.

The score of a model in also shown in relation to a set of high-resolution PDB structures (Z-score). The plot relates the obtained global QMEAN4 value to scores calculated from a set of high-resolution X-ray structures. Local estimates of the model quality based on the QMEAN scoring function are shown as per-reside plot. Each residue is assigned a reliability score between 0 and 1, describing the expected similarity to the native structure. Higher numbers indicate higher reliability of the residues.

GMQE (Global Model Quality Estimation) is a quality estimation which combines properties from the target-template alignment. The resulting GMQE score is expressed as a number between zero and one, reflecting the expected accuracy of a model built with that alignment and template.
Higher numbers indicate higher reliability. Once a model is built, the GMQE gets updated for this specific case by also taking into account the QMEAN4 score of the obtained model in order to increase reliability of the quality estimation. 3D structure validations are shown in[Fig. 3].

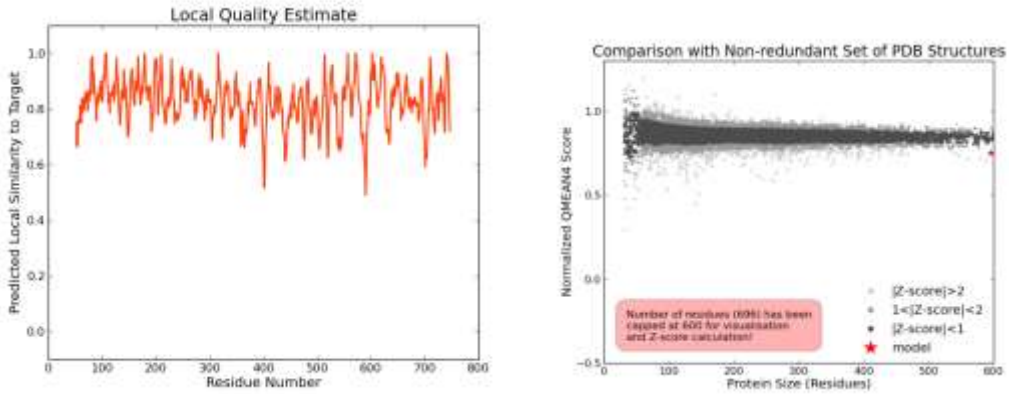

Fig.3: 3D structure validations.

Topology and signal peptide prediction

FhuA signal peptide cleavage site was predicted between positions 30 and 31 of protein sequence.
FhuA topology predicted by spoctopus server is shown in [Fig. 4].
A 2D topology model of protein was built based on predicted inside, transmembrane and outside regions of the protein [Fig. 5].

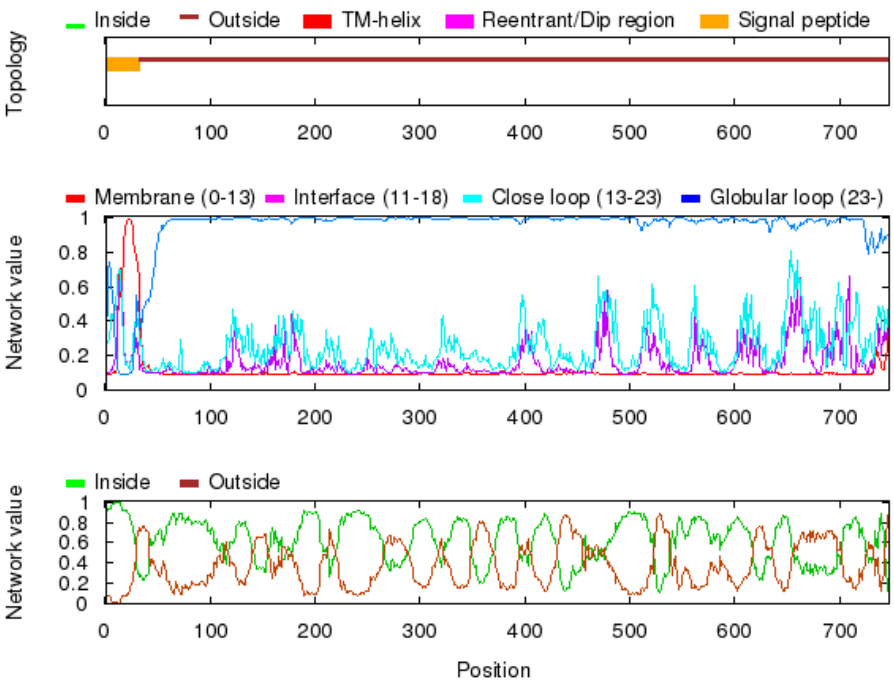

Fig. 4: FhuATopology and signal peptide cleavage site predicted by spoctopus server.

Topology output: This is a graphic representation of the most likely topology as predicted by SPOCTOPUS.
Network output: The two diagrams show the estimated preference for each residue to be located in different structural regions. The top diagram shows the preference of being either in:
- the hydrophobic part of the membrane, 0-13Å from the membrane center (M)
- the membrane water-interface, 11-18Å from the membrane center (I)
- a close loop region, 13-23Å from the membrane center (L)
- a globular region, further than 23Å from the membrane (G)

The bottom diagram shows the estimated preference of a particular residue to be located either on the inside (i) or on the outside (o) of the membrane.
The raw data underlying these two plots can be found in the SPOCTOPUS network file.

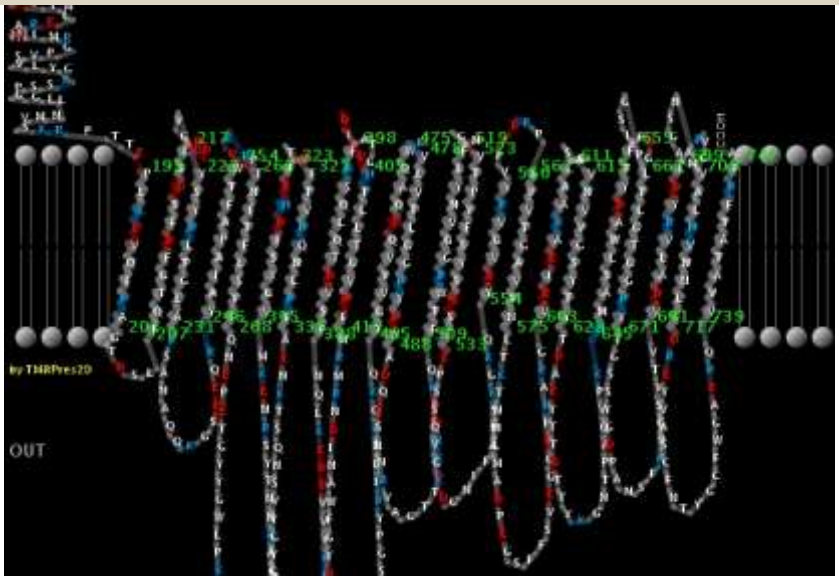

Fig.5: A 2D topology model of FhuA.

CONFLICT OF INTEREST
There is no conflict of interest.

ACKNOWLEDGEMENTS
The authors are thankful to Hon'ble Dean and Management

FINANCIAL DISCLOSURE
None

REFERENCES

[1] Welte W, Nestel U, Wacker T, Diederichs K. [1995]
[2] Structure and function of the porin channel. Kidney Int. 48: 930–940.
[3] Braun V, Hantke K, Ko ¨ ster W. [1998] Bacterial Iron Transport: Mechanisms, Genetics, and Regulation. In IronTransport and Storage in Microorganisms, Plants, and Animals, volume 35, A. Sigel and H. Sigel, eds. (New York: Marcel DekkerInc.), pp. 67–145.
[4] Ferguson AD, Hofmann E, Coulton JW, Diederichs K, Welte W. [1998] Siderophore-mediated iron transport: crystal structure of FhuA with bound lipopolysaccharide. Science 282: 2215–2220.
[5] Locher, K.P., et al., and Moras, D. (1998). Transmembrane signalling across the ligand-gated FhuA receptor: crystal structures of free and ferrichrome-bound states reveal allosteric changes. Cell 95: 771–778.
[6] Ferguson AD, Welte W. [2000] Crystal structure of the antibiotic albomycin in complex with the outermembrane transporter FhuA. Protein Sci. 9: 956–963