Structure Prediction of Outer Membrane Protease Protein of *Salmonella typhimurium* Using Computational Techniques

Rozina Tabassum*, Muhammad Haseeb, Sahar Fazal

Department of Bioinformatics and Biosciences
Mohammad Ali Jinnah University
Islamabad, Pakistan
E-mails: rozina_tabassum58@yahoo.com, haseeb3389@hotmail.com, sahar@jinnah.edu.pk

Corresponding author

Received: April 04, 2015 **Accepted: November 16, 2015**

Published: March 31, 2016

Abstract: *Salmonella typhimurium, a facultative gram-negative intracellular pathogen belonging to family Enterobacteriaceae, is the most frequent cause of human gastroenteritis worldwide. PgtE gene product,outer membrane protease emerges important in the intracellular phases of salmonellosis. The pgtE gene product of S. typhimurium was predicted to be capable of proteolyzing T7 RNA polymerase and localize in the outer membrane of these gram negative bacteria. PgtE product of S. enterica and OmpT of E. coli, having high sequence similarity have been revealed to degrade macrophages, causing salmonellosis and other diseases. The three-dimensional structure of the protein was not available through Protein Data Bank (PDB) creating lack of structural information about E protein. In our study, by performing Comparative model building, the three dimensional structure of outer membrane protease protein was generated using the backbone of the crystal structure of Pla of Yersinia pestis, retrieved from PDB, with MODELLER (9v8). Quality of the model was assessed by validation tool PROCHECK, web servers like ERRAT and ProSA are used to certify the reliability of the predicted model. This information might offer clues for better understanding of E protein and consequently for developmet of better therapeutic treatment against pathogenic role of this protein in salmonellosis and other diseases.*

Keywords: *Outer membrane protease E, Structure prediction, Salmonella typhimurium, Homology modeling, MODELLER.*

Introduction

Salmonella typhimurium (*Salmonella enteric* serovar *typhimurium*) is a facultative anaerobic intracellular pathogen; gram-negative bacterium [5, 24] belongs to genus *Salmonella*, subspecie enterica, member of the Enterobacteriaceae family. This bacterium causes the highest incidence of salmonellosis in both human and nonhuman hosts and also one of the frequent reasons of food-borne gastroenteritis in humans [15, 27]. *S. typhimurium* is the most common isolated serotype, in total of the 2,400 serotypes in the genus *Salmonella* of human, which transmission mainly occur by ingestion of contaminated food, primarily meats, or by the faecal-oral route from an infected individual [3, 11, 16]. The genome of *S. typhimurium* consist of numerous prominent genes that codes for virulence factors which are absent from non-pathogen strains. These regions of the genome are recognized as pathogenecity islands, the two most important islands are, *Salmonella* Pathogenecity Island 1 and *Salmonella* Pathogenecity Island 2, Proteins secreted by the SPI2 TTSS permit *Salmonella* to grow, thus overcoming a critical element of the innate immune response, ultimately kill host macrophage [4, 7]. The *S. typhimurium* produced outer membrane

proteins, intervene the process of adhesion and localize inside the host intestinal epithelium, thus manipulating the succession of disease. As the outer membrane proteins are exposed to surface, providing effective targets for the development of better vaccines and antimicrobial drugs [6]. PgtE gene product, E protein appears significant in the intracellular phases of salmonellosis. The pgtE gene product of *S. typhimurium* and OmpT, *E. coli outer membrane protease* amino acid sequence are extremely homologous, so it was predicted that pgtE codes for a protein, is functionally homologous to OmpT and has ability to proteolyze T7 RNA polymerase and to localize in the outer membrane of *S. typhimurium*. PgtE of *S. enterica* and OmpT of *E. coli* have also been discovered to degrade macrophages,causing gastroenteritis to severe enteric fever and several other diseases [13]. *S. typhimurium* E protein comprises of 312 amino acids, belongs to a family of highly homologous outer membrane proteases, known as omptins includes other members like *E. coli outer membrane protease* OmpT, *Yersinia pestis Pla* etc. found in several gram negative bacterias [30]. The three-dimensional structure of the protein was not accessible through Protein Data Bank (PDB) until now [34], creating gap of structural information about E protein, which could be useful for better understanding of its role in pathogenecity and drugs development.

The aim of the current study was to predict the 3D model of *outer membrane protease* E of *S. typhimurium* by comparative or homology modeling. The study was carried out with the special reference to the active site and cofactor binding sites. It helps in better understanding of various aspects of structural features of the outer membrane protease and in future, it will be helpful in development of novel and improved therapeutic agents.

Materials and methods

Sequences extraction

Primary sequences of *outer membrane protease* E, i.e. E protein (SWISS-PROT AC: P06185) was retrieved from the SWISS-PROT [10] (http://www.us.expasy.org) and PIR (http://www.georgetown.edu) database [32].

Secondary structure prediction

In our study the Secondary structure prediction of target sequences was affirmed by submitting the sequences to the Consensus Secondary Structure Prediction Sever http://pbil.ibcp.fr/NPSA/npsa_npsa [25] and PDB SUM [19].

Multiple sequence alignment

Multiple sequence alignment was carried out via the program CLUSTAL X [14] to identify the homologous and functionally important regions, the Sequences which are homologous to the target Sequences, were retrieved from SWISS-PROT [10].

Phylogenetic analysis

Phylogenetic analysis is used to establish the evolutionary relationships among organisms, if a entire genome cannot obtained ,then gene content can used for taxonomical study by means of marker genes. The results of this analysis can be obtained in form of Draw gram and Draw tree [1, 14].

Comparative modeling

Comparative modeling is the most reliable *in silco* prediction method, based on principal to build a protein structure from the known protein. Homology modeling can be divided into

four steps: template identification, alignment, model construction, enhancement and validation, all steps performed through various computational tools [8].

Searching template

Template searching was carried out through BLAST (Psi-BLAST) [23] algorithm against PDB [2]. The sequence which shows high homology to the target sequence has chosen as template.

Model construction and enhancement

3D comparative model of *outer membrane protease* E was build using the crystal coordinates of template on the basis of alignment between target and template sequences. All steps of homology modeling and refinement were done by the program MODELLER (Version 9 (9v8)) [8].

Graphics screen using three dimensional visualization programs, Ds-Viewer [26] were used in order to check out the reliability of the predicted model and modeling of variable surface loops, structural investigations. Effectiveness of the predicted model was analyzed by the program PROCHECK [18], the Energy Command of the MODELLER [9] is use to investigate the geometry, chemistry, and energy distributions of the model. The ProSA (Protein Structure Analysis) web server [37] is used to determine the energy graphs structural design of protein folds to verify the protein structure quality and ERRAT [28] was used to find out the statistics of non-bonded interactions between different atom types. To detect the deviation of the modeled target from template protein structure, Root Mean Square Deviation (RMSD) [22] values were calculated between the set of targets and template proteins.

Protein-ligand interactions

Protein-ligand interactions studied using Program Ligand Explorer (http://www.kukool.com/ligand) [21] to check the affinity between target protein and ligands.

LigPlot

The program LigPlot [36] was utilized to automatically generate schematic 2-D representations of protein-ligand complexes stand utilizing and PDB file input. The output is a color, or black-and-white Post script file, which provides hydrogen bond and hydrophobic interactions among molecules and their distances in easy and informative representation.

ERRAT

ERRAT [28] is a protein structure verification algorithm that is particularly used to evaluate the progress of crystallographic model structure and refinement. It can be carried out through online web server http://nihserver.mbi.ucla.edu/ERRATv2/.

Results and discussion

Sequence analysis

After the extraction of query sequence from Swissprot under accession number (SWISS-PROT AC: P06185) [10], the template search for *outer membrane protease* (*E*) was carried out by using BLAST algorithm against PDB [2]. Outer membrane protease of *S. typhimurium* contains 312 residues and *Pla* of *Yersinia pestis* comprise of 291 residues. The blast Sequence similarity searches of E protein (1XPO) (SWISS-PROT AC: P06185) showed best 75% identity and 85% positives with *Yersinia plasminogen* activator, *Pla* of *Yersinia pestis* (SWISS-PROT AC: P17811). So the crystal structure of *Yersinia plasminogen* activator *Pla*

of *Yersinia pestis* (PDB ID: 2X4M) was chosen as a template on the basis of highest sequence similarity score and lowest E-value, for constructing the prediction of 3D structure of E protein. The BLAST result and the align 2D command of MODELLER [8], which realigned the target and template sequences, shifting the gaps at the end of the target sequence, as a result reduced the gaps in outer membrane protease sequence, are presented in Figs. 1 and 2, respectively.

```
>pdb|2X4M|A S Chain A, Yersinia Pestis Plasminogen Activator Pla
 pdb|2X4M|B S Chain B, Yersinia Pestis Plasminogen Activator Pla
 pdb|2X4M|C S Chain C, Yersinia Pestis Plasminogen Activator Pla
 pdb|2X4M|D S Chain D, Yersinia Pestis Plasminogen Activator Pla
Length=298

 Score =  463 bits (1191),  Expect = 5e-166, Method: Compositional matrix adjust.
 Identities = 217/291 (75%), Positives = 248/291 (85%), Gaps = 0/291 (0%)

Query  22   SALFIPDVSPDSVTTSLSVGVLNGKSRELVYDTDTGRKLSQLDWKIKNVATLQGDLSWEP  81
            S+  IP++SPDS T + S G+L+GKS E++YD +TGRK+SQLDWKIKNVA L+GD+SW+P
Sbjct  2    SSQLIPNISPDSFTVAASTGMLSGKSHEMLYDAETGRKISQLDWKIKNVAILKGDISWDP  61

Query  82   YSFMTLDARGWTSLASGSGHMVDHDWMSSEQPGWTDRSIHPDTSVNYANEYDLNVKGWLL  141
            YSF+TL+ARGWTSLASGSG+M D+ WM+  Q  WTD S HP T+VN+ANEYDLNVKGWLL
Sbjct  62   YSFLTLNARGWTSLASGSGNMDDYAWMNENQSEWTDHSSHPATNVNHANEYDLNVKGWLL  121

Query  142  QGDNYKAGVTAGYQETRFSWTARGGSYIYDNGRYIGNFPHGVRGIGYSQRFEMPYIGLAG  201
            Q +NYKAG+TAGYQETRFSWTA GGSY Y+NG Y GNFP GVR IGY+QRF MPYIGLAG
Sbjct  122  QDENYKAGITAGYQETRFSWTATGGSYSYNNGAYTGNFPKGVRVIGYNQRFSMPYIGLAG  181

Query  202  DYRINDFECNVLFKYSDWVNAHDNDEHYMRKLTFREKTENSRYYGASIDAGYYITSNAKI  261
             YRINDFE N LFK+SDWV AHDNDEHYMR LTFREKT  SRYYG  I+AGYY+T NAK+
Sbjct  182  QYRINDFELNALFKFSDWVRAHDNDEHYMRDLTFREKTSGSRYYGTVINAGYYVTPNAKV  241
```

Fig. 1 Blast result of the selected template (2X4M) with target (1XPO)

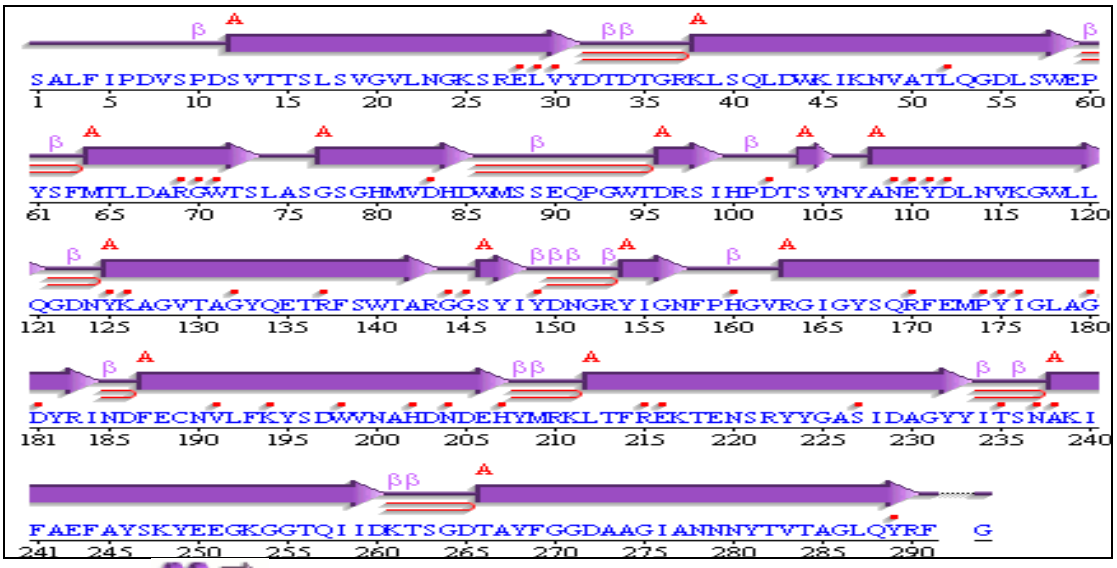
```
Length=298
Identities = 217/291 (75%), Positives = 248/291 (85%), Gaps = 0/291 (0%)

2X4M      SSQLIPNISPDSFTVAASTGMLSGKSHEMLYDAETGRKISQLDWKIKNVAILKGDISWDPYSFLTLNA
1XPO      SALFIPDVSPDSVTTSLSVGVLNGKSRELVYDTDTGRKLSQLDWKIKNVATLQGDLSWEPYSFMTLDA
          *   **  **** *   * * * *** *  **  **** *********** * ** ** **** ** *
2X4M      RGWTSLASGSGNMDDYAWMNENQSEWTDHSSHPATNVNHANEYDLNVKGWLLQDENYKAGITAGYQET
1XPO      RGWTSLASGSGHMVDHDWMSSEQPGWTDRSIHPDTSVNYANEYDLNVKGWLLQGDNYKAGVTAGYQET
          ********** * *  **   *  *** * ** * ** ************** ***** *******
2X4M      RFSWTATGGSYSYNNGAYTGNFPKGVRVIGYNQRFSMPYIGLAGQYRINDFELNALFKFSDWVRAHDN
1XPO      RFSWTARGGSYIYDNGRYIGNFPHGVRGIGYSQRFEMPYIGLAGDYRINDFECNVLFKYSDWVNAHDN
          ****** **** * ** * **** *** *** *** ******** ******* * *** **** ****
2X4M      DEHYMRDLTFREKTSGSRYYGTVINAGYYVTPNAKVFAEFTYSKYDEGKGGTQTIDKNSGDSVSIGGD
1XPO      DEHYMRKLTFREKTENSRYYGASIDAGYYITSNAKIFAEFAYSKYEEGKGGTQIIDKTSGDTAYFGGD
          ****** *******  *****  * **** * *** **** **** ******* *** ***    ***
2X4M      AAGISNKNYTVTAGLQYRFG
1XPO      AAGIANNNYTVTAGLQYRF-
          **** * ************
```

Fig. 2 Target (outer membrane protease)-template (2X4M) alignment by using ALIGN2D command of MODELLER (9v8): conserved active site – red; conserved residues – blue.

Secondary structure prediction

Protein secondary structure is playing a key role in the study of protein folds, side chain and their interactions. There are several statistical methods used for the prediction of protein secondary structure [38]. The secondary structure of outer membrane protease E (Fig. 3) contains 292 total residues. The secondary structure results obtained from PDBsum [19] shows the strands, which means it is mostly composed of β strands 75%, there is 0% α helix in the total amino acids sequence and about 24% constitutes coils and loops as shown in

Fig. 3 and Table 1. Secondary structure of E protein consist of 1 β sheet which is anti parallel type and barrel, 9 β hairpins 15 β strands.



Sec. struc: ; Helices labeled H1, H2, ... and strands by their sheets A, B, ...;
Motifs: beta turn beta hairpin; Residue contacts: to ligand

Fig. 3 The overall secondary structure information of E protein which is almost consists of β sheets

Table 1. Secondary structure summary

Strand	α helix	3-10 helix	Other	Total residues
221 (75.7%)	0 (0.0%)	0 (0.0%)	71 (24.3%)	292

Multiple sequence alignment and phylogenetic analysis

Examination of these several sequences (9 sequences) shows outstanding sequence resemblance among primary structures of all identified natural proteases of gram negative Enterobacteriaceae family. The conservation rate of active site and the metal binding residues are high among all the family members. A Phylogenetic tree and gram results show that both target outer membrane protease protein (1XPO) (pgtE SALTY) from *S. typhimurium* and template *Pla* of *Yersinia pestis* (2X4M) (COLY YERPE) are derived from common ancestors, as tree clearly shows the divergence of these OMPTIN family members [17]. The pgtE of *S. typhimurium*, OmpT of *E. coli*, *Pla* of *Yersinia pestis* and *sopA* of *shigilla flexenri* etc. from a common ancestral gene as indicated by the mid-point root, are highly homologous to one another as per evolutionary perspective [12]. It also demonstrates that they have a placed within the same OmpT in super family as shown in Fig. 4.

Topology of predicted E protein

The general structure of *outer membrane protease E* highly resembles to that of its template *Pla* of *Yersinia pestis*, OmpT from *E. coli*, and other members of omptin family as both target and template have, β barrel topology. The tertiary structure mainly consists of anti-parallel β-sheets that projected distant from the lipid bilayer into the extracellular space [35]. This can be clearly seen in topology and tertiary structures presented in Figs. 5 and 6. About 5 exposed loops (L1-L5) are present on the surface and ten transmembrane β strands comprises the two structures, which can be also observed usually within the other members of omptins family [17, 35].

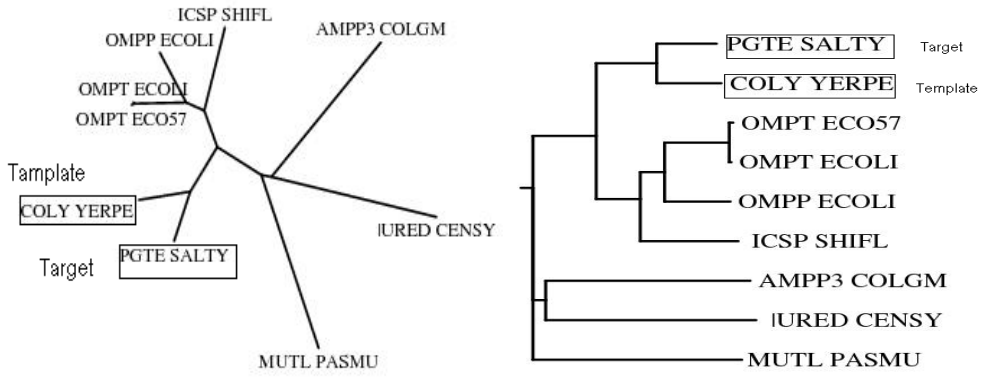


Fig. 4 The results of draw gram and draw tree

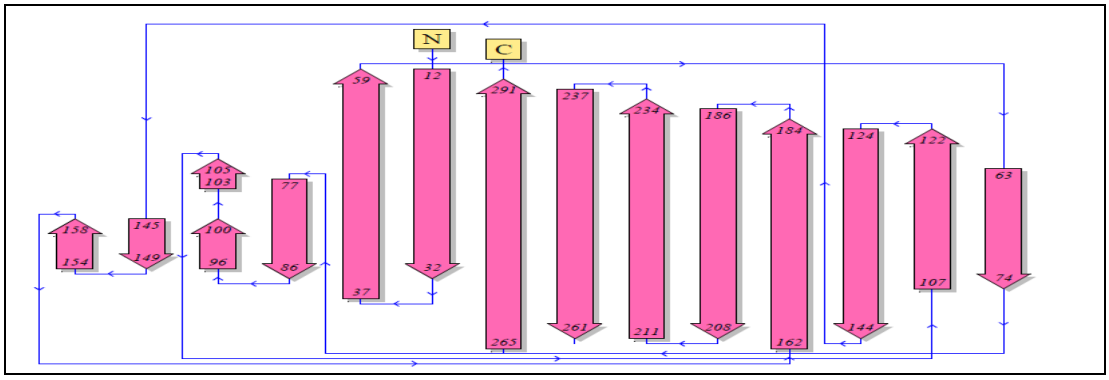


Fig. 5 An anti-parallel β topology of secondary structure of outer membrane protease 2. The N terminal and C terminal are indicted as N and C, pink arrows are β sheets and blue lines shows the coils n turns and there is no single α helices.

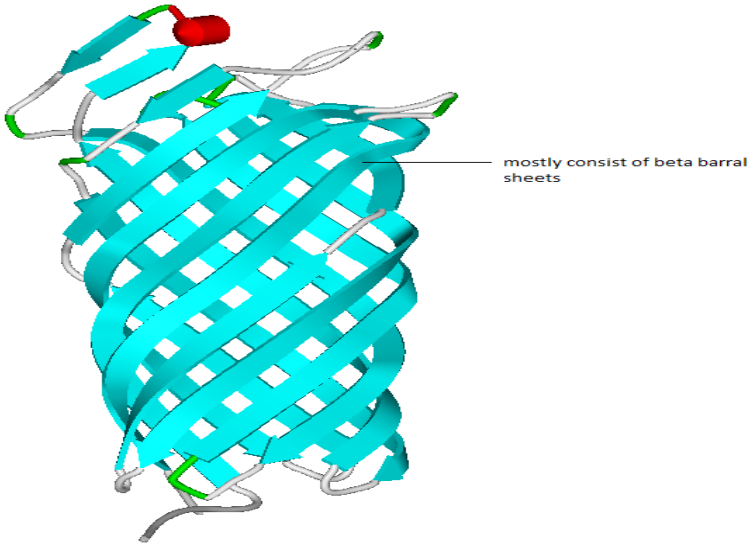


Fig. 6 The schematic representation of homology model of outer membrane protease demonstrating arrangement of tertiary structure of E protein consisting chiefly of anti-parallel β-sheets

Superposition

General depiction of similarities and differences between the backbones of target (1XPO) and template (2X4M) were found by superposing the two models on each other. The RMSD value of superposed model (Fig. 7) was found to be 0.1110 Å, as apparent from the lower RMSD value and sequence similarity, the two structures are well superimposed and have alike folding topology but have different conformations at certain regions, e.g. in region ser205-ser207 as shown in Fig. 10. As the structure is well superimposed, the active sites of the two proteins are identical with similar geometry as most active site and metal binding residues are similar between two models.

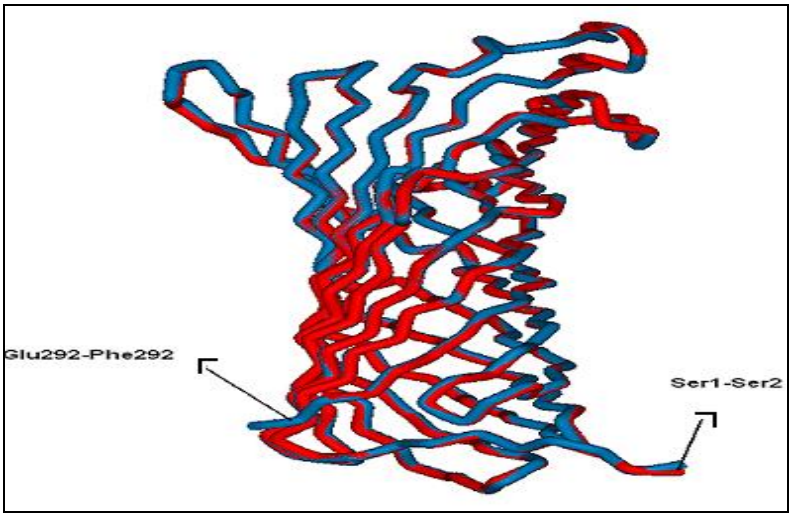


Fig. 7 Structure superposition of atoms of outer membrane protease (blue) on the crystal structure of 2X4M (red). The structural difference at the beginning of the superposed structure due to missing first residue of template.

Model assessments

Evaluation of the target model involved analysis of geometry, stereochemistry and energy distribution in the models, which is carried out though different validation tools such as PROCHEK [18] and online source ProSA [37], which are used for the geometric examination of the model. The predicted model of Target E protein (1XPO) (Fig. 8) fulfills all criteria executed in PROCHECK as the crystal structure of target refined at 2.5 Å resolution equivalent as template (2X4M), provide output in the form of Ramachandran plot which demonstrates that the majority 95.6% of the main-chain dihedral angles are found in the most favored regions (> 90% for a good model) [29], 4.0% in the additionally allowed, 0.4% in the generously allowed while no residue is found in the disallowed regions. In general, the model show relatively good protein geometry like the homolog outer membrane protease of *E. coli* [35], as most of the quality parameters are better than or in the range of lenience. Therefore it can be considered as structurally practical model.

The energetic structural design of protein folds of newly constructed structure was calculated by utilizing the online program ProSA, which gives the general quality score in the form of plot. Its value is shown in a plot that holds the *Z*-scores of all experimentally determined protein chains in current PDB. *Z*-score was used as the measure of this energy, which demonstrated the general quality of protein structures. In general, positive values correspond to problematic or erroneous parts of the input structure, while zero and negative score is indication of stabilized structure [37]. In this plot, groups of structures from diverse sources

(X-ray, NMR) are controlled by distinctive colors (light blue, dark blue) [20]. *Z*-score of the template molecule (2X4M) is -3.46, while the *Z*-score (-2.85) shows general model quality of the predicted target (1XPO) (Figs. 9 and 10). From this negative score of target established the model is structurally reliable. As is evident from the result, the energy graphs corresponded to highly stable structures and the values of energy, for the target molecule, mostly well below from Zero point on x-axis, which was in consistence with the values of experimentally found Template molecule. That is why, from the energy stabilizations point of view, there seemed to be no problems in the modeled structure and on the basis of this graph, the modeled 3D structure can be approved.

```
 +----------<<<  P  R  O  C  H  E  C  K     S  U  M  M  A  R  Y  >>>----------+
 |                                                                             |
 |  1XPO5   2.5                                                  292 residues  |
 |                                                                             |
+|  Ramachandran plot:    95.6% core     4.0% allow      .4% gener      .0% disall |
 |                                                                             |
 |  Gly & Pro Ramach:       0 labelled residues (out of  38)                   |
+|  Chi1-chi2 plots:        2 labelled residues (out of 173)                   |
 |                                                                             |
 |  Main-chain params:      6 better      0 inside       0 worse               |
 |  Side-chain params:      5 better      0 inside       0 worse               |
 |                                                                             |
*|  Residue properties:  Max.deviation:      2.9              Bad contacts:     9 |
*|                       Bond len/angle:     6.1     Morris et al class:  1  1  2 |
 |                                                                             |
 |  G-factors            Dihedrals:    -.08  Covalent:    -.21     Overall:    -.12 |
 |                                                                             |
 |  M/c bond lengths: 98.1% within limits   1.9% highlighted                   |
 |  M/c bond angles:  75.3% within limits  24.7% highlighted                   |
 |  Planar groups:   100.0% within limits    .0% highlighted                   |
 |                                                                             |
 +-----------------------------------------------------------------------------+
   + May be worth investigating further.   * Worth investigating further.
```

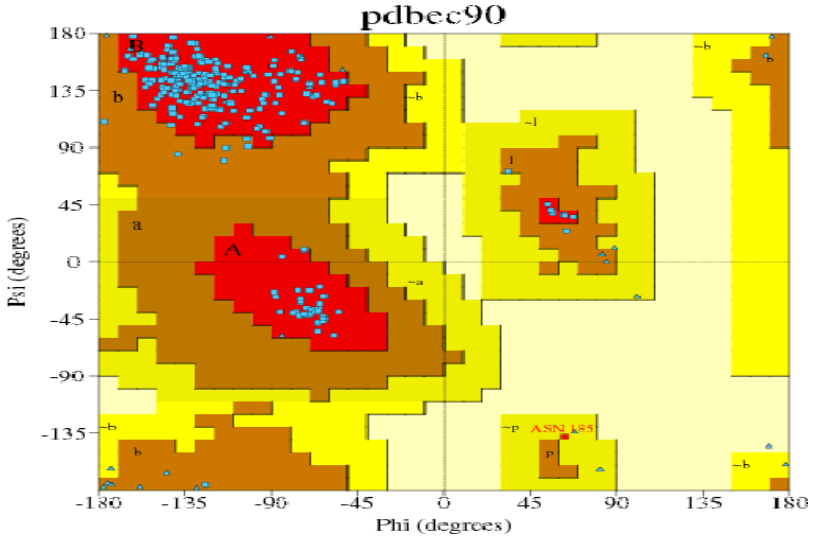


Fig. 8 Ramachandran plot statistics of the 1XPO models obtained by the software PROCHECK

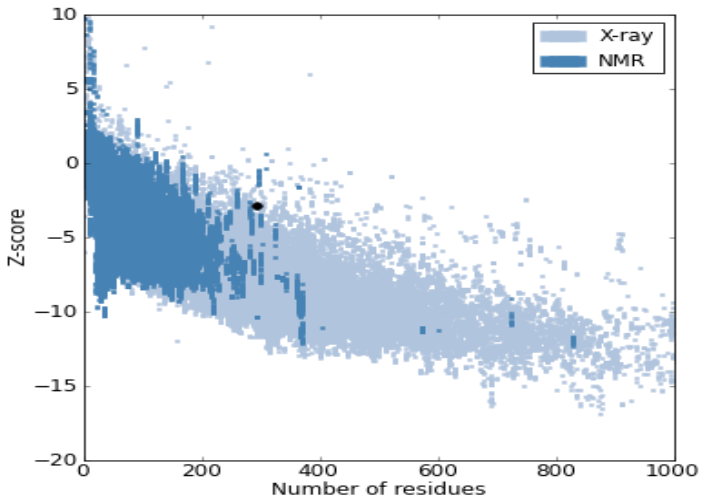


Fig. 9 Screenshot of ProSA *Z*-score plot of E protein showing the Z value < 0 negative value -2.85

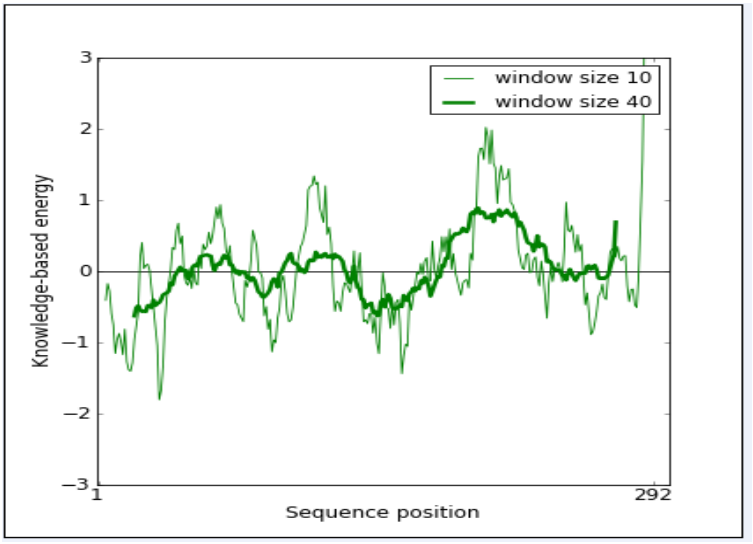


Fig. 10 Screenshot of ProSA plot of E protein showing energy graph of residue score of a predicted protein structure

Active site prediction

The template *Pla* and the target E protein are outer membrane proteases, two ligands SO4 and C8E that is vital for their catalytic activity, which clearly shown by program LigPlot from Pdbsum (Figs. 11 and 12). The catalytic site residues conserved in all omptins family members. A highly conserved feature between the two classes of a highly conserved feature between these two classes of proteases is conservation of sequence surrounding the catalytic histidine and serine residues (Thr-Ala-Gly-His-Cys and Gly-Asp-Ser-Gly-Gly) [35].

The Tyr 289 form bond with SO4 (293) having distance of 2.91. Asn 237 and Thr 235 form bond with oxygen atom of SO4 having 2.89 distances showing better interaction. Asn 204 form bond with oxygen of SO4 294 with distance of 2.79. The LigPlot result of interactions of SO4 295 showing that not interact with any amino acid residue or other ligand.

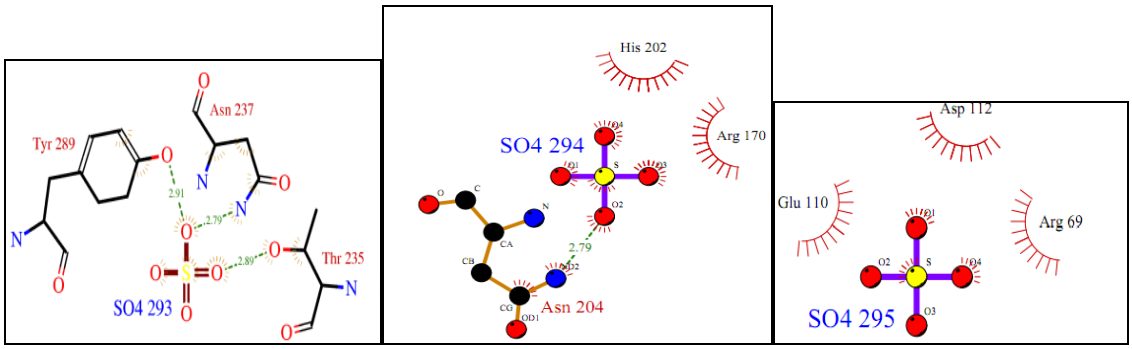


Fig. 11 LigPlot result shows the interactions of metal ion (SO4) with the amino acid residues of the protein

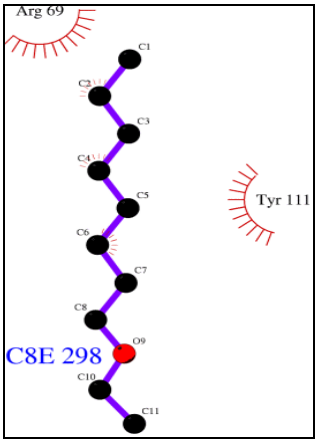


Fig. 12 LigPlot result of interactions of C8E 298 shows that not interact with any amino acid residue or other ligand

ERRAT model

ERRAT is a sensitive and good method for the recognition of incorrect folded regions in initial protein models, which evaluate the arrangement of diverse types of atoms with respect to one another in the protein model [31]. The predicted ERRAT plot (Fig. 13) shows that the overall quality factor of target model is 74.29%, which confirmed the reliability of predicted model.

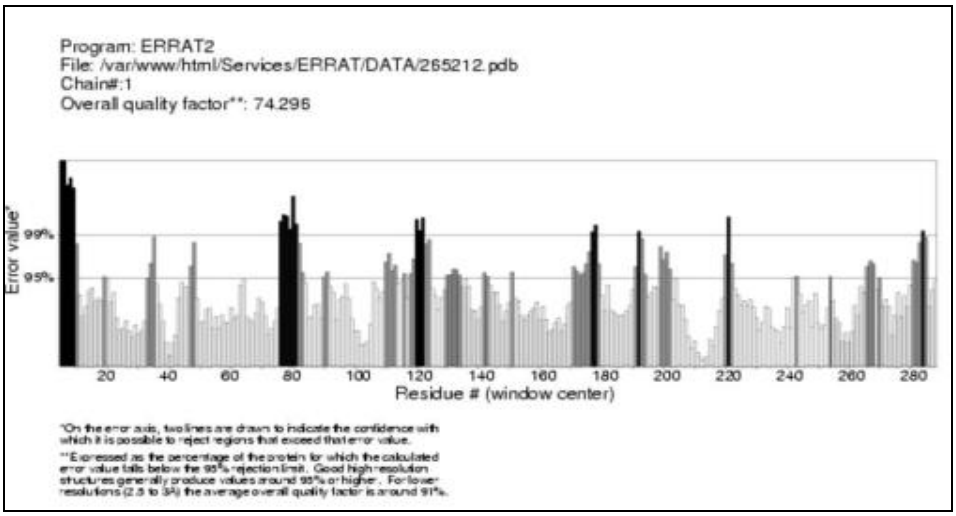


Fig. 13 Overall quality of the target model

Conclusion

The three dimensional structure of the query sequence was modeled using MODELLER, and validation of this model was carried out through variety of tools. The Ramachandran plot, -2.85 *Z*-score of ProSA, overall quality of the model 74.29 through ERRAT validate the reliability of the predicted model. As the primary structure of target E protein and template proteins *Pla* shows best amino acid homology, the predicted 3d structure of outer membrane protease also has *β* barrel topology, showing high resemblance to its template *Pla* of *Yersinia pestis* and other members of omptin family. A Phylogenetic tree results show that both target protein *outer membrane protease* from *S. typhimurium* and plasmogen activator *Pla* from *Yersinia pestis* and other omptin family members are highly homologous to each other as these derived from common ancestor. The catalytic site residues conserved in all omptins family members. A highly conserved feature between the two classes of proteases is conservation of sequence surrounding the catalytic histidine and serine residues (Thr-Ala-Gly-His-Cys and Gly-Asp- Ser-Gly-Gly), E protein also have same active sites. The predicted model of aspartic outer membrane E protein will be useful in the better understanding the active site and ligand binding sites in favor of control of diseases caused from this protein for improved drug in future with more efficacy and lesser side effects.

Acknowledgements

I owe special thanks to Dr. Sahar Malik, Faculty Mohammad Ali Jinnah University, Islamabad, for their support and supervision in this research work.

References

1. Al-khayyat M. Z. S. (2013). Taxonomic Study of Campylobacter Species Using hsp60 Protein, Int J Bioautomation, 17(1), 17-24.
2. Berman H., K. Henrick, H. Nakamura, J. L. Markley (2007). The Worldwide Protein Data Bank (wwPDB): Ensuring a Single, Uniform Archive of PDB Data, Nucleic Acids Research, 35(Suppl. 1), D301-D303.
3. Brenner F. W., R. G. Villar, F. J. Angulo, R.T. Axue, B. Swaminathan (2000). Salmonella Nomenclature, Journal of Clinical Microbiology, 38, 2465-2467.
4. Burns-Keliher L. L., A. Protteus, R. Curtiss (1997). Specific Detection of *Salmonella typhimurium* Proteins Synthesized Intracellularly, Journal of Bacteriology, 179, 3604-3612.
5. Cabral J. P. (2010). Water Microbiology – Bacterial Pathogen and Water, International Journal of Environmental Research and Public Health, 7, 3657-3703.
6. Chooneea D., R. Karlsson, V. Encheva, C. Arnold, H. Appleton, H. Shah (2010). Elucidation of the Outer Membrane Proteome of *Salmonella enterica* Serovar *typhimurium* Utilising a Lipid-based Protein Immobilization Technique, BMC Microbiology, 10(1), 44.
7. Dieye Y., K. Ameiss, M. Mellata, R. 3rd Curtiss (2009). The Salmonella Pathogenecity Island(SPI)1 Contributes More Than SPI 2 to the Colonization of the Chicken by *Salmonella enterica* Sarovar *typhimurium*, BioMed Central Microbiology, 9, 3-9.
8. Eswar N., B. Webb, M. A. Marti-Renom, M. S. Madhusudhan, D. Eramian, M. Y. Shen, U. Pieper, A. Sali (2006). Comparative Protein Structure Modeling Using Modeller, Current Protocols in Bioinformatics, Chapter 5, Unit 5.6, doi: 10.1002/0471250953. bi0506s15.
9. Eswar N., B. John, N. Mirkovic, A. Fiser, V. A. Ilyin, U. Pieper, A. C. Stuart, M. A. Marti-Renom, M. S. Madhusudhan, B. Yerkovich, A. Sali (2003). Tool for Comparative Protein Structure Modeling and Analysis, Nucleic Acids Research, 31, 3375-3380.

10. Gasteiger E., E. Jung, A. Bairoch (2001). SWISS-PROT: Connecting Biomolecular Knowledge via a Protein Database, Current Issues in Molecular Biology, 3, 47-55.
11. Gray V. L., O. M. Reilly, C. T. Muller, I. D. Watkins, D. Lloyd (2006). Low Tyrosine Content of Growth Media Yields Aflagellate *Salmonella enterica* serovar *typhimurium*, 152, 8-23.
12. Haiko J., M. Suomalainen, T. Ojala, K. Lähteenmäki, T. K. Korhonen (2009). Invited Review: Breaking B Arriers-attack on Innate Immune Defences by OmpT in Surface Proteases of Enterobacterial Pathogens, Innate Immunity, 15(2), 67-80.
13. Hardt W. D., L. M. Chen, K. E. Schuebel, X. R. Bustelo, J. E. Galán (1998). *S. typhimurium* encodes an Activator of Rho GTPases that Induces Membrane Ruffling and Nuclear Responces in Host Cells, Cell, 93(5), 815-826.
14. Huson D. H., D. C. Richter, C. Rausch, T. Dezulian, M. Franz, R. Rupp (2007). Dendroscope: An Interactive Viewer for Large Phylogenetic Trees, Biomedical Center Bioinformatics, 8, 460, doi: 10.1186/1471-2105-8-460.
15. Ibarra J. A., O. S Mortimer (2009). The Ultimate Insider: Salmonella Virulence Factors that Modulate Intracellular Survival, Cellular Microbiology, 11, 1579-1586.
16. Jarvik T., C. Smillie, E. A. Groisman, H. Ochman (2010). Short-term Signatures of Evolutionary Change in the *Salmonella enteric* Serovar *typhimurium* 14028 Genome, Journal of Bacteriology, 192, 560-567.
17. Kukkonen M., T. K. Korhonen (2004). The Omptin Family of Enterobacterial Surface Proteases/adhesins: From Housekeeping in *Escherichia coli* to Systemic Spread of *Yersinia pestis*, International Journal of Medical Microbiology, 294(1), 7-14.
18. Laskowski R. A., J. A. Rullmannn, M.W. MacArthur, R. Kaptein, J. M. Thornton (1996). AQUA and PROCHECK-NMR: Programs for Checking the Quality of Protein Structures Solved by NMR, Journal of Biomolecular NMR, 8(4), 477-486.
19. Laskowski R. A., V. V. Chistyakov, J. M. Thornton (2005). PDBsum More: New Summariesand Analyses of the Known 3D Structures of Proteins and Nucleic Acids, Nucleic Acids Res, 33, 266-268.
20. Liu H., L. Chen, Q. Li, M. Zheng, J. Liu (2014). Computational Study on Substrate Specificity of a Novel Cysteine Protease 1 Precursor from Zea mays, International Journal of Molecular Sciences, 15(6), 10459-10478.
21. Lutfullah G., F. Amin, Z. Khan, N. Azhar, M. K. Azim, S. Noor, K. Shoukat (2008). Homology Modeling of Hemagglutinin/protease [HA/P (vibriolysin)] from *Vibrio cholerae*: Sequence Comparision, Residue Interactions and Molecular Mechanism, The Protein Journal, 27(2), 105-114.
22. Maiorov V. N., G. M. Crippen (1994). Significance of Root-mean-square Deviation in Comparing Three-dimensional Structures of Globular Proteins, Journal of Molecular Biology, 235(2), 625-634.
23. McGinnis S., T. L. Madden (2004). BLAST: At the Core of a Powerful and Diverse Set of Sequence Analysis Tools, Nucleic Acids Research, 32, W20-W25.
24. Niedergang F., J. C. Sirard, C. T. Blanc, J. P. Kraehenbhuhl (2000). Entry and Survival of *Salmonella typhimurium* in Dendritic Cells and Presentation of Recombinant Antigens do not Require Macrophage Specific Virulence Factors, Proceedings of the National Academy of the Sciences of America, 97, 14650-14655.
25. Park J., S. Dietmann, A. Heger, L. Holm (2000). Estimating the Significance of Sequence Order in Protein Secondary Structure and Prediction, Bioinformatics, 16(11), 978-987.
26. Pazel D. P. (1989). DS-Viewer – Interactive Graphical Data Structure Presentation Facility, IBM Systems Journal, 28(2), 307-323.
27. Que J. U., D. J. Hentges (1985). Effect of Streptomycin Administration on Colonization Resistance to *Salmonella typhimurium* in Mice, Infect Immun, 48(1), 169-1174.

28. Shen M. Y., A. Sali (2006). Stastical Potential for Assessment and Prediction of Protein Structures, Protein Science, 15, 2507-2524.
29. Singh A., T. K. Pal (2014). *In silico* Sequence Analysis, Structure Prediction and Function Annotation of Human Bcl-X Beta Protein, International Journal Bioautomation, 18(1), 23-30.
30. Sodeinde O. A., J. D. Goguen (1989). Nucleotide Sequence of the Plasminogen Activator Gene of *Yersinia pestis*: Relationship to ompT of *Escherichia coli* and gene E of *Salmonella typhimurium*, Infection and Immunity, 57(5), 1517-1523.
31. Swain S. S. (2013). *In silico* Approach in the Prediction and Analysis of the Three-dimensional Structure of *Maleylacetate reductase*: A Biodegrading Protein, International Journal Bioautomation, 17(4), 217-226.
32. Tárraga J., I. Medina, L. Arbiza, J. Huerta-Cepas, T. Gabaldón, J. Dopazo, H. Dopazo (2007). Phylemon: A Suite of Web Tools for Molecular Evolution, Phylogenetics and Phylogenomics, Nucleic Acids Research, 35, W38-W42.
33. Thompson J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, D. G. Higgins (1997). The CLUSTAL_X Windows Interface: Flexible Strategies for Multiple Sequence Alignment Aided by Quality Analysis Tools, Nucleic Acids Research, 25, 4876-4882.
34. UniProt Consortium (2008). The Universal Protein Resource (UniProt), Nucleic Acids Research, 36,190-195.
35. Vandeputte-Rutten L., R. A. Kramer, J. Kroon, N. Dekker, M. R. Egmond, P. Gros (2001). Crystal Structure of the Outer Membrane Protease OmpT from *Escherichia coli* Suggests a Novel Catalytic Site, The EMBO Journal, 20(18), 5033-5039.
36. Wallace A. C., R. A. Laskowski, J. M. Thornton (1995). LIGPLOT: A Program to Generate Schematic Diagrams of Protein-ligand Interactions, Protein Eng, 8(2), 127-134.
37. Wiederstein M., M. J. Sippl (2007). ProSA-Web: Interactive Web Service for the Recognition of Errors in Three-dimentional Structures of Proteins, Nucleic Acids Research, 35, 10-40.
38. Wu Z. Y., R. P. Han (2013). SAAS: Short Amino Acid Sequence – A Promising Protein Secondary Structure Prediction Method of Single Sequence, International Journal Bioautomation, 17(2), 65-72.

Rozina Tabassum, M.Sc. Student
E-mail: rozina_tabassum58@yahoo.com

Rozina Tabassum is presently M.Sc. Scholar of Bioinformatics in Muhammad Ali Jinnah University, Islamabad, Pakistan. She has interests in variety of bioinformatics research areas such as computer aided drug designing, data integration and analysis, phylogenetic analysis and mathematical modeling of biological pathways and physiology based pharmacokinetics modeling.

Muhammad Haseeb, Ph.D. Student
E-mail: haseeb3389@hotmail.com

Muhammad Haseeb is currently a Ph.D. Scholar in Bioinformatics in Mohammad Ali Jinnah University, Islamabad, Pakistan. He has interests in many areas of bioinformatics such as mathematical modeling of biological pathways, data analysis, protein structure prediction and prediction tools for drug designing and phylogenetic analysis.

Assoc. Prof. Sahar Fazal, Ph.D.
E-mail: sahar@jinnah.edu.pk

Dr. Sahar Fazal has done her Ph.D. in Applied Entomology from South China Agriculture University Gunangzhou, People Republic of China. Currently she is working as an Associate Professor in Department of Bioinformatics at Mohammad Ali Jinnah University, Islamabad, Pakistan. Her interest research areas are bioinformatics (phylogenetics, protein interactions, modeling, pathways), resistance management, molecular entomology, microbiology and genetics.