Advanced Modelling and Functional Characterization of B2 Bradykinin Receptor

Muhammad Saad Khan*, Sahar Fazal

Department of Bioinformatics
Mohammad Ali Jinnah University
Islamabad Expressway, Near Kaakpul, Islamabad, Pakistan
E-mails: sadihere_007@hotmail.com, sahar@jinnah.edu.pk

**Corresponding author*

Received: November 25, 2014 **Accepted: June 18, 2015**

Published: June 30, 2015

Abstract: *Hereditary angioedema (giant hives) is an autosomal dominant malady characterized by repetitive episodes of probably life-threatening angioedema due to a partial deficiency of C1 inhibitor. B2 Bradykinin Receptor's (BKRB2) amino acid sequence is deposited within UniProt under accession number P30411. The Physicochemical properties of BKRB2 sequence are determined by using ProtParam. BKRB2's secondary structure was predicted through PROTEUS. Pfam domain was used for functional characterization of BKRB2. PSI-BLAST was used to find homologs of known structure. Modelling by satisfaction of spatial restraints, either uses distance geometry or optimization techniques to satisfy spatial restraints performed by MODELLER. The quality of the generated model was evaluated with PROCHECK by Ramachandran plot analysis. Validation of the generated models was further performed by WHAT IF. ProSA was used for the analysis of Z-scores and energy plots. The 3D structures of the modeled proteins were analyzed using UCSF Chimera. Clustal Omega is used for multiple sequence alignment that uses seeded guide trees and HMM profile-profile techniques to generate alignments.*

Keywords: *B2 Bradykinin Receptor, C1 inhibitor, Hereditary angioedema, MODELLER, Ramachandran plot, ProSA.*

Introduction

Hereditary angioedema (HAE) is an autosomal dominant disease characterized by repetitive episodes of potentially life-threatening angioedema due to a partial deficiency of C1 inhibitor (C1-INH). HAE caused by deficiency of C1-INH was discovered separately by Landerman et al. [34] and Donaldson and Evans [15] who described the molecular mechanisms underlying attacks of swelling in HAE, which have been gradually dissected in the last 50 years [13-15]. Attacks of angioedema in HAE can be severe and long lasting, with typical duration of 3-5 days before the patient is well again. Abdominal attacks may become a reason for hospitalization and often lead to inappropriate intra-abdominal surgery and oro-pharyngeal-laryngeal attacks can be life-threatening [17, 20, 40]. The key factors in the pathogenesis of HAE are factor XII (Hageman factor), kallikrein, high-molecular-weight kininogen (HK), C1-INH, and abradykinin. Evidence suggests that bradykinin is the mediator of angioedema. HAE includes the presence of kallikrein in induced blister fluids in patients with HAE, decreased prekallikrein and HK levels during HAE may increase the severity, and high level of local and circulating bradykinin in acute HAE [31]. Bradykinin is present in abnormally high quantities in patients with HAE, which increases the permeability of the vascular wall for fluid. Icatibant blocks the B2 Bradykinin Receptor (BKRB2) on the cells of the blood vessel walls, which ultimately prevents the swelling caused by bradykinin.

Amino acid sequence of BKRB2 is deposited within UniProt database under accession number P30411 in FASTA format. ProtParam [21] is most commonly used to calculate the physicochemical properties of sequence. Physicochemical properties play an important role in determining the function of a protein.

Tertiary structure of BKRB2 was not known. As we know structure is more evolutionary conserved than sequence; therefore, analysis of three-dimensional (3D) structures holds great potential. Structure elucidation which requires extensive expertise is an expensive and time consuming process. Currently used techniques to reveal 3D structures are X-ray crystallography and Nuclear Magnetic Resonance (NMR) Imaging. There has been an increasing gap of information between DNA/protein sequence information and structure information because these techniques are expensive and time consuming. It has long been argued that, if the segments of secondary structure could be accurately predicted, the tertiary structure could be predicted by simply trying different arrangements of the segments in space [11].

Computational methods and molecular dynamic simulations are good alternatives to overcome these problems in protein structure prediction [36]. Comparative modeling is a computational technique for Tertiary structure prediction of proteins using known structures as templates [49]. Templates can be found using the target sequence as a query for searching structure databases such as the Protein Data Bank [62], SCOP [12], DALI [27], and CATH [41]. There is about 20% to 70% probability of finding a related protein of known structure for a sequence picked [18, 29, 45, 48, 50]. PSI-BLAST generally finds homologs of known structure for approximately twice as many sequences [43, 56]. The use of multiple templates approximately equidistant from the target sequence generally increases the model accuracy [50, 55].

Use of multiple structures and sequence information is frequently beneficial [4]. First, the alignment of the potential templates is prepared by superposing their structures. Next, the sequences that are clearly related to the templates and are easily aligned with them are added to the alignment. Next the target sequence is aligned in the same way. Taking structural information into account as much as possible, the two profiles are aligned with each other [32, 37, 58]. Modelling by satisfaction of spatial restraints, either uses distance geometry or optimization techniques to satisfy spatial restraints obtained from the alignment [49].

The first step in model evaluation is to determine if the model has the correct fold [51]. Sequence identity above 30% is a relatively good predictor of the expected accuracy of model. The reasons are the well known relationship between structural and sequence similarities of two proteins [10]. A basic requirement for a model is to have good stereochemistry. A useful program for evaluating stereochemistry is WHAT IF [60]. The features of a model that are checked by this program include bond lengths, bond angles, Coarse Packing Quality Control. ProSA [54] was used for the analysis of Z-scores and energy plots. Ramachandran plot produced by PROCHECK [35] was used to evaluate the quality of model.

Multiple Sequence Alignment (MSA) is generally the alignment of three or more biological sequences. Homology can be inferred and the evolutionary relationships between the sequences studied. Clustal Omega [38, 52] is used for MSA.

The aim of this study is to use bioinformatics approach such as secondary and tertiary structure prediction from a sequence and calculation of different parameters related to its function [53].

Materials and methods

Sequence retrieval of BKRB2

The amino acid sequences of BKRB2 were obtained in FASTA format from the Protein sequence database UniProt [3] under accession number P30411. Amino acid sequence length is 391.

Physiochemical characterization of BKRB2

Physiochemical properties were predicted using ProtParam [21]. Various physical and chemical parameters: theoretical pI, total number of negatively (Asp + Glu) and positively (Arg + Lys) charged residues, extinction coefficients [23], instability index [25], aliphatic index [30] and grand average of hydropathicity (GRAVY) [33] were computed using ProtParam, a proteomics server [21].

Secondary structures of BKRB2

Prediction of protein structure and its features is therefore an important area of computational and structural biology. A secondary structure of BKRB2 was predicted through PROTEUS [39]. PROTEUS is a web server and a standalone application that increases the recent advancements in data mining and machine learning program to perform very accurate protein secondary structure predictions. PROTEUS consists of three high-performing de novo structure prediction methods (PSIPRED, JNET and TRANSSEC), a jury of expert's consensus tool and robust PDB-based structure alignment methods to generate all of its highly accurate secondary structure predictions.

Functional domains identification of BKRB2

To perform functional characterization of BKRB2 domains were identified by Pfam [5]. Conserve domains has been directly involved in sequence/structure/function relationships.

Model building, refinement and evaluation of BKRB2

MODELLER is a well-known tool in homology modeling. The tool is used for homology or comparative modeling of protein three-dimensional structures [16, 37]. The user provides an alignment of a sequence to be modeled with known related structures and the tool automatically calculates a model containing all non-hydrogen atoms. MODELLER performs comparative protein structure modeling by satisfaction of spatial restraints [19, 49].

Identification of the best template structure is one of the critical steps in homology modeling. A template search is done by a web based tool Position Specific Iterative Basic Local Alignment Search Tool (PSI-BLAST) [1]. Multiple templates with the E-values better than threshold are considered for modeling. The structures of selected templates are taken from Protein Data Bank (PDB) [6].

The model was generated with multiple templates using MODELLER 9.12 advanced modeling.

Table 1. The selected PDB template structures with E-value better than threshold

Templates	Resolution	PDB ID
A structure of the CXCR4 chemokine receptor in complex with small molecule antagonist IT1t	2.50 Å	3ODU
Crystal structure of the CXCR4 chemokine receptor in complex with a cyclic peptide antagonist CVX15	2.90 Å	3OE0
Crystal structure of the CXCR4 chemokine receptor in complex with a small molecule antagonist IT1t in I222 spacegroup	3.20 Å	3OE6
Crystal structure of human protease-activated receptor 1 (PAR1) bound with antagonist vorapaxar	2.20 Å	3VW7
Crystal structure of the CXCR4 chemokine receptor in complex with a small molecule antagonist IT1t in P1 spacegroup	3.10 Å	3OE8
Crystal Structure of the CCR5 chemokine receptor	2.71 Å	4MBS

The quality of generated model was evaluated with PROCHECK [3] by Ramachandran plot analysis [46]. Ramachandran plot is a two-dimensional geometrical plot of φ-ψ angles for the assessment of protein backbone structure, depicts information of the protein structure and 3D conformation and also provides information about the residues lying in favored, allowed or outlier region [26]. Validation of generated models was further performed by WHAT IF [60]. ProSA [54] was used for the analysis of Z-scores and energy plots. The 3D structures of modeled proteins were analyzed using UCSF Chimera [44]. Root Mean Square Deviation (RMSD) values were calculated between the set of targets and template protein to see how much modeled protein deviates from the template protein structure.

Multiple sequence alignment

MSA is generally the alignment of more than two sequences. From the output, homology can be inferred and the evolutionary relationships between the sequences studied. Clustal Omega [34, 35] is used for multiple sequence alignment that uses seeded guide trees and HMM profile-profile techniques to generate alignments. Percent Identity Matrix [8] was created and phylogenetic tree [28] was generated by Clustal Omega.

Results and discussion

Physiochemical characterization of BKRB2

For physiochemical characterization, theoretical pI (isoelectric point), molecular weight, -R and +R (total number of positive and negative residues), EI (extinction coefficient), II (instability index), AI (aliphatic index) and GRAVY (grand average hydropathy) were computed using the ExPASy's ProtParam server for E proteins. The computed pI value (pI > 7) indicated their basic nature. The computed pI will be beneficial for developing buffer system for purification by well known isoelectric focusing method. Extinction coefficient value 79410 for E proteins is indicating the presence of higher concentration of Tyr and Trp. On the basis of II results classified that proteins is probably not stable (II > 40). The AI which is defined as the relative volume of a protein occupied by aliphatic side chain is regarded as the positive factor for the increase of thermal stability of globular proteins. The very high AI of all E proteins infers that these proteins may be stable for a wide range of temperature. The low GRAVY index of protein infers that these proteins could result in a better interaction with water shown in Table 2.

Table 2. Physiochemical properties of BKRB2

Mw	T-pI	-R	+R	ECp	ECr	II	AI	GAH
44460.5	8.50	26	31	80285	79410	42.44	109.62	0.457

Mw – molecular weight, T-pI – theoretical pI, -R – total number of negatively charged residues (Asp + Glu), +R – total number of positively charged residues (Arg + Lys), ECp – extinction coefficient (all pairs of Cys residues from cystines), ECr – extinction coefficient (assuming all Cys residues are reduced), II – instability index, AI – aliphatic index, GAH – grand average of hydropathicity.

Secondary structures of BKRB2

Protein secondary structure predictions plays a valuable role for molecular biologists in deciding the place where need to sub clone the protein fragments for expression of gene, where to join or insert gene fragments, or choosing where to add affinity tags for protein purification [24, 57]. Secondary structure predictions can also be used to calibrate Circular Dichroism and Fourier Transform Infrared Spectroscopy measurements when observing and checking the quality of folding or unfolding proteins with unknown tertiary structure [22, 59]. Secondary structure predictions may also be helpful to understand structural parameters given by NMR spectra (of known or novel proteins) and helps to determine protein flexibility and stability [61, 63]. Secondary structure prediction of BKRB2 protein showed that the protein is having the composition of Helix = 61%, Strand = 4%, and Coil = 35%. 238 residues contain helix formation, 15 residues are confined to the beta sheet and 138 residues consist of coils in secondary structure representation of envelope protein structure. Overall confidence value of predicted secondary structure against each protein residue is very good as shown in Table 3.

Table 3. Secondary structure summary of all residues of BKRB2

Name	AA	OC value	Predicted %HC	Predicted %BSC	Predicted % CC
BKRB2	391	81.0%	61% (238 AA)	4% (15 AA)	35% (138 AA)

AA – number of amino acids, OC – overall confidence, HC – helix content, BSC – beta sheet content, CC – coil content.

Functional domains identification of BKRB2

A protein functional domain can exist independently of the rest of the protein chain having three-dimensional structure. Conserve domains act as building blocks in molecular evolution. The BKRB2 protein revealed that it has only one domain of G protein-coupled receptor [42] showed in the Table 4. Rhodopsin-like GPCRs themselves represent a widespread protein family that includes hormones, neurotransmitters, and light receptors, all of which transduce extracellular signals through interaction with guanine nucleotide-binding (G) proteins. Although their activating ligands vary widely in structure and character, the amino acid sequences of the receptors are very similar and are believed to adopt a common structural framework comprising 7 transmembrane (TM) helices [2, 7, 9].

Table 4. Functional domains identification of BKRB2

Name	Domains	Start	End
BKRB2	7 transmembrane receptor (rhodopsin family)	74	332

Model building of BKRB2

Proteins are very important organic compounds in our cells. They are involved in nearly all cell functions. Each protein in the body has a specific function. Proteins do not exist as linear conformation; they mostly exist in compact and folded structures. Rodriguez et al. [47] described that the knowledge of the 3D structure is an initial step for understanding great importance for the design of drugs and function prediction. Protein functions are determined by their overall three-dimensional conformation. For model prediction homology modeling approach [19] was used in order to derive their structures. Receptor bradykinin is associated with G proteins that activate a phosphatidylinositol-calcium second messenger system.

Templates had been selected on the basis of maximum similarity between target and template sequence. Templates were selected from PDB databank. Templates alignment is the next step after template selection in homology modeling as model building depends entirely on alignment of sequence and structures. Multiple sequence alignment of sequence with multiple templates was performed by Clustal Omega. The generated percent identity matrix is shown in Table 5.

Table 5. Multiple sequence alignment of BKRB2 sequence with templates

1.	sp|P30411|BKRB2_HUMAN	100	24.13	23.60	26.95	26.71	26.71	26.82
2.	4MBS_A|PDBID|CHAIN	24.13	100	20.11	32.43	32.43	32.43	31.91
3.	3VW7_A|PDBID|CHAIN	23.60	20.11	100	48.28	48.18	48.18	47.57
4.	3OE0_A|PDBID|CHAIN	26.95	32.43	48.28	100	99.20	99.20	99.20
5.	3ODU_A|PDBID|CHAIN	26.71	32.43	48.18	99.20	100	100	100
6.	3OE8_A|PDBID|CHAIN	26.71	32.43	48.18	99.20	100	100	100
7.	3OE6_A|PDBID|CHAIN	26.82	31.91	47.57	99.20	100	100	100

Phylogenetic tree generated by Clustal Omega is based on similarity among sequences showing the inferred evolutionary relationship among biological species as shown in Fig. 1.

Fig. 1 Phylogenetic tree generated based on similarity among sequences

Bioinformatics software tools are then used to predict the 3D structure of the target based on the known 3D structures of the templates. MODELLER is used for homology or comparative modeling of protein 3D structures [16, 37].

From each selected template an ensemble of multiple models may be generated by MODELLER 9.12 [49] by applying default model building routine 'model'. The model with lowest objective function is considered to be the best of all [49]. The final structure illustrated in Fig. 2 was modeled with MODELLER with the use of multiple template.

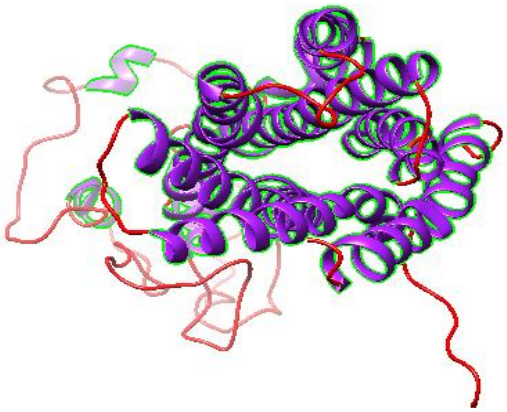

Fig. 2 Final validated model using advanced modeling technique BKRB2

Model validation and evaluation of BKRB2

Predicted model was visualized through UCSF Chimera. Ramachandran plot of BKRB2 was obtained by PROCHECK. There were 325 residues in residues in the core region (red) favorable regions; 25 residues in allowed regions (brown); 5 residues in generously allowed regions (yellow); 3 residues in disallowed regions (off white); overall G-factor: -0.08. Validation of predicted model of BKRB2 by PROCHECK is illustrated in Fig. 3.

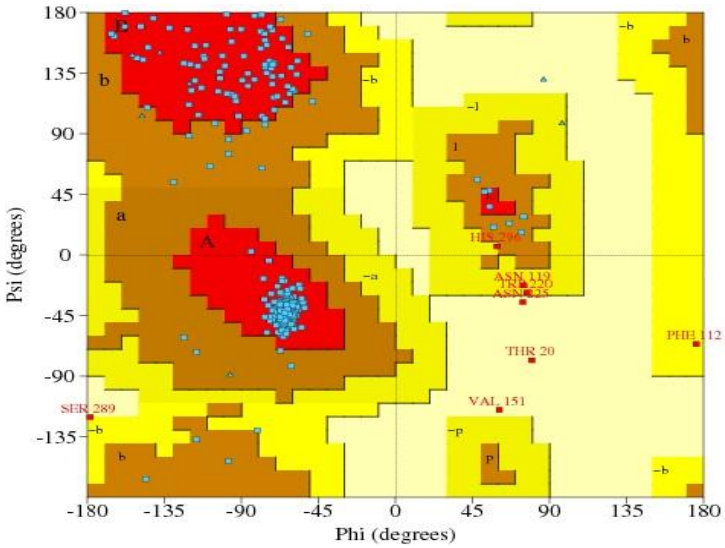

Fig. 3 Validation of predicted model of BKRB2 by PROCHECK

Altogether 97.8 % of the residues were found to be in favored and allowed regions, 1.4% in generously allowed regions, which validate the quality of homology model. The overall G-factor for Bradykinin B2 receptor was -0.08, shown in Table 6. As the value is greater than the acceptable value -0.50, this suggests that the modelled structure is acceptable with resolution of at least 2.0 Å and a good quality model would be expected to have over 90% in the most favored regions. This implies that the predicted model is compatible with its sequence.

Table 6. Summary of validation of predicted model of BKRB2

Ramachandran plot results (%)					WHAT IF report results					ProSA
MF	AAR	GAR	DR	G score	Coarse Packing Quality Control	RMS Z value for BL	RMSD value for BL	RMS Z score for BA	RMSD value for BA	Z-score
90.8	7.0	1.4	0.8	-0.08	0.702	0.969	0.015	1.263	2.283	-0.49

MF – most favored, AR – allowed region, GAR – generously allowed region, DR – disallowed region, OQF – overall quality factor, BL – bond length, BA – bond angle.

ProSA was used for the analysis of Z-scores and energy plots. ProSA program calculates an overall quality score for a specific input structure. The Z-score calculated is -0.49 and a representation of the overall quality of the model is illustrated in Fig. 4. It checks if the Z-score of the input structure is within the range of scores found for native proteins of similar size. The Z-score measures the deviation of the total energy of the structure with respect to an energy distribution derived from random conformations. Positive values correspond to problematic or erroneous parts of the input structure.

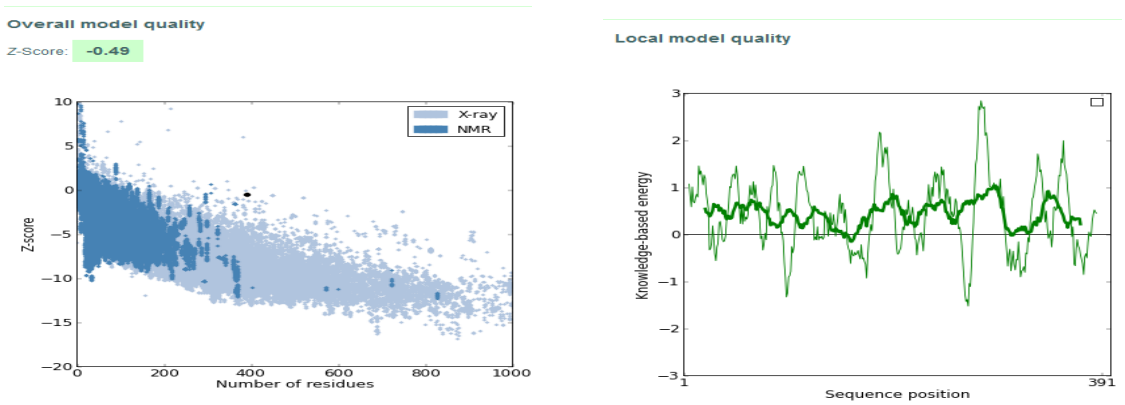

Fig. 4 Analysis of Z-scores and energy plots by ProSA program

Structure evaluation results are in favor of new predicted model with the optimal values of RMS with very small deviation of bond angle and bond length. Z-scores and overall best quality factor along with highest scores for allowed favorable residues in Ramachandran plot. Nice peaks generated by energy graphs show structure compatibility.

Acknowledgements

Authors are very thankful to their fellows, especially Muhammad Fakhar, Qurat-Ul-Ain, Sadaf Nasreen, Tehreem Zainab, Absar Mazhar, and Obaid Sheikh, for their invaluable support.

References

1. Altschul S.F., Madden T.L., Schaffer A.A., Zhang J., Zhang Z., Miller W., Lipman D.J. (1997). Gapped BLAST and PSI-BLAST: A new generation of protein database search programs, Nucleic Acids, 25, 3389-3402.
2. Attwood T.K., Findlay J.B.C. (1993). Design of a discriminating fingerprint for G-protein-coupled receptors, Protein Engineering, 6(2), 167-176.

3. Bairoch A., Apweiler R., Wu C.H., Barker W.C., Boeckmann B., Ferro S., Gasteiger E., Huang H., Lopez R., Magrane M., Martin M.J., Natale D.A., O'Donovan C., Redaschi N., Yeh L.S. (2005). The Universal Protein Resource (UniProt), Nucleic Acids Res, 33(1), D154-D159.
4. Barton G.J., Sternberg M.J.E. (1987). A strategy for the rapid multiple alignment of protein sequences: Confidence levels from tertiary structure comparisons, J Mol Biol, 198(2), 327-337.
5. Bateman A., Birney E., Cerruti L., Durbin R., Etwiller L., Eddy S.R., Griffiths-Jones S., Howe K.L., Marshall M., Sonnhammer E.L. (2002). The Pfam protein families database, Nucleic Acids Res, 30(1), 276-280.
6. Bernstein F.C., Koetzle T.F., Williams G.J.B., Meyer Jr E.F., Brice M.D., Rodgers J.R., Kennard O., Shimanouchi T., Tasumi M. (1977). The Protein Data Bank: A computer-based archival file for macromolecular structures, J Mol Biol, 112(3), 535-542.
7. Birnbaumer L. (1990). G proteins in signal transduction, Annual Review of Pharmacology and Toxicology, 30(1), 675-705.
8. Campanella J.J., Bitincka L., Smalley J. (2003). MatGAT: An application that generates similarity/identity matrices using protein or DNA sequences, BMC Bioinformatics, 4:29.
9. Casey P.J., Gilman A.G. (1988). G protein involvement in receptor-effector coupling, J Biol Chem, 263(6), 2577-2580.
10. Chothia C., Lesk A. M. (1986). The relation between the divergence of sequence and structure in proteins, The EMBO J, 5(4), 823-826.
11. Cohen F.E., Presnell S.R. (1996). The Combinatorial Approach, In "Protein Structure Prediction" (Ed. Sternberg M.J.E.), Oxford University Press, Oxford, 207-228.
12. Conte L.L., Brenner S.E., Hubbard T.J.P., Chothia C., Murzin A.G. (2002). SCOP database in refinements accommodate structural genomics, Nucleic Acids Res, 30, 264-267.
13. Davis A.E. 3rd (2006). Mechanism of angioedema in first complement component inhibitor deficiency, Immunology and Allergy Clinics of North America, 26(4), 633-651.
14. Davis A.E. 3rd, Bissler J.J., Aulak K.S. (1993). Genetic defects in the C1 inhibitor gene, Complement Today, 1, 133-150.
15. Donaldson V.H., Evans R.R. (1963). A biochemical abnormality in hereditary angioneurotic edema: Absence of serum inhibitor of C′1-esterase, Amer J Med, 35(1), 37-44.
16. Eswar N., Marti-Renom M.A., Webb B., Madhusudhan M.S., Eramian D., Shen M., Pieper U., Sali A. (2006). Comparative protein structure modeling with MODELLER, Current Protocols in Bioinformatics, S15, 5.6.1-5.6.30.
17. Fay A., Abinun M. (2002). Current management of hereditary angio-oedema (C'1 esterase inhibitor deficiency), J Clin Path, 55(4), 266-270.
18. Fischer D., Eisenberg D. (1997). Assigning folds to the proteins encoded by the genome of Mycoplasma genitalium, Proceedings of the National Academy of Sciences, 94(22), 11929-11934.
19. Fiser A., Do R.K.G., Sali A. (2000). Modeling of loops in protein structures, Protein Science, 9, 1753-1773.
20. Frank M.M., Jeffrey A.G., John P.A. (1976). Hereditary angioedema: The clinical syndrome and its management, Annals of Internal Medicine, 84(5), 580-593.
21. Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A. (2005). Protein identification and analysis tools on the ExPASy server, In "The Proteomics Protocols Handbook" (Ed. Walker J.M.), Humana Press, 571-607.

22. Gibbs A.C., Bjorndahl T.C., Hodges R.S., Wishart D.S. (2002). Probing the structural determinants of type II' beta-turn formation in peptides and proteins, J Am Chem Soc, 124, 1203-1213.
23. Gill S.C., Hippel P.H.V. (1989). Calculation of protein extinction coefficients from amino acid sequence data, Anal Biochem, 182(2), 319-326.
24. Grasselli E., Noviello G., Rando C., Nicolini C., Vergani L. (2003). Expression, purification and characterisation of a novel mutant of the human protein kinase CK2, Mol Biol Rep, 30, 97-106.
25. Guruprasad K., Reddy B.V.B., Pandit M.W. (1990). Correlation between stability of a protein and its dipeptide composition: A novel approach for predicting *in vivo* stability of a protein from its primary sequence, Protein Eng, 4(2), 155-161.
26. Ho B.K., Brasseur R. (2005). The Ramachandran plots of glycine and pre-proline, BMC Structural Biology, 5:14.
27. Holm L., Sander C. (1999). Protein folds and families: Sequence and structure alignments, Nucleic Acids Res, 27, 244-247.
28. Howe K., Bateman A., Durbin R. (2002). QuickTree: Building huge neighbour-joining trees of protein sequences, Bioinformatics, 18(11), 1546-1547.
29. Huynen M., Doerks T., Eisenhaber F., Orengo C., Sunyaev S., Yuan Y., Bork P. (1998). Homology-based fold predictions for Mycoplasma genitalium proteins, J Mol Biol, 280, 323-326.
30. Ikai A. (1980). Thermostability and aliphatic index of globular proteins, J Biochem, 88(6), 1895-1898.
31. Kaplan A.P., Joseph K. (2010). The bradykinin-forming cascade and its role in hereditary angioedema, Annals of Allergy, Asthma & Immunology, 104(3), 193-204.
32. Koretke K.K., Luthey-Schulten Z., Wolynes P.G. (1996). Self-consistently optimized statistical mechanical energy functions for sequence structure alignment, Protein Sci, 5, 1043-1059.
33. Kyte J., Doolittle R.F. (1982). A simple method for displaying the hydropathic character of a protein, J Mol Biol, 157(1), 105-132.
34. Landerman N.S.,Webster M.E., Becker E.L., Ratcliffe H.E. (1962). Hereditary angioneurotic edema. II. Deficiency of inhibitor for serum globulin permeability factor and/or plasma kallikrein, J Allergy, 33, 330-341.
35. Laskowski R.A., MacArthur M.W., Moss D.S., Thornton J.M. (1993). PROCHECK: A program to check the stereochemical quality of protein structures, J Appl Crystal, 26(2), 283-291.
36. Liu H.L., Hsu J.P. (2005). Recent developments in structural proteomics for protein structure determination, Proteomics, 5(8), 2056-2068.
37. Martí-Renom M.A., Stuart A.C., Fiser A., Roberto S., Melo F., Sali A. (2000). Comparative protein structure modeling of genes and genomes, Annu Rev Biophys Biomol Struct, 29, 291-325.
38. McWilliam H., Li W., Uludag M., Squizzato S., Park Y.M., Buso N., Cowley A.P., Lopez R. (2013). Analysis tool web services from the EMBL-EBI, Nucleic Acids Res, 41(W1), W597-W600.
39. Montgomer S., Sundararaj S., Gallin W.J., Wishart D.S. (2006). Improving the accuracy of protein secondary structure prediction using structural alignment, BMC Bioinformatics, 7:301.
40. Nzeako U.C., Frigas E., Tremaine W.J. (2001). Hereditary angioedema: A broad review for clinicians, Archives of Internal Medicine, 161(20), 2417-2429.

41. Orengo C.A., Bray J.E., Buchan D.W.A., Harrison A., Lee D., Pearl F.M.G., Sillitoe I., Todd A.E., Thornton J.M. (2002). The CATH protein family database: A resource for structural and functional annotation of genomes, Proteomics, 2, 11-21.
42. Palczewski K., Kumasaka T., Hori T., Behnke C.A., Motoshima H., Fox B.A., Trong I.L., Teller D.C., Okada T., Stenkamp R.E., Yamamoto M., Miyano M. (2000). Crystal structure of rhodopsin: A G protein-coupled receptor, Science, 289(5480), 739-745.
43. Park J., Karplus K., Barrett C., Hughey R., Haussler D., Hubbard T., Chothia C. (1998). Sequence comparisons using multiple sequences detect three times as many remote homologues as pairwise methods, J Mol Biol, 284, 1201-1210.
44. Pettersen E.F., Goddard T.D., Huang C.C., Couch G.S., Greenblatt D.M., Meng E.C., Ferrin T.E. (2004). UCSF Chimera – A visualization system for exploratory research and analysis, J Comp Chem, 25(13), 1605-1612.
45. Pieper U., Eswar N., Stuart A.C., Ilyin V.A., Sali A. (2002). ModBase, a database of annotated comparative protein structure models, Nucleic Acids Res, 30, 255-259.
46. Ramachandran G.N., Ramakrishnan C.T., Sasisekharan V. (1963). Stereochemistry of polypeptide chain configurations, J Mol Biol, 7(1), 95-99.
47. Rodriguez R., Chinea G., Lopez N., Pons T., Vriend G. (1998). Homology modeling, model and software evaluation: Three related resources, Bioinformatics, 14(6), 523-528.
48. Rychlewski L., Zhang B., Godzik A. (1998). Fold and function predictions for Mycoplasma genitalium proteins, Fold Des, 3, 229-238.
49. Šali A., Blundell T.L. (1993). Comparative protein modelling by satisfaction of spatial restraints, J Mol Biol, 234(3), 779-815.
50. Sánchez R., Sali A. (1997). Advances in comparative protein-structure modelling, Curr Opin Struct Biol, 7, 206-214.
51. Sánchez R., Sali A. (1998). Large-scale protein structure modeling of the *Saccharomyces cerevisiae* genome, Proceedings of the National Academy of Sciences, 95(23), 13597-13602.
52. Sievers F., Wilm A., Dineen D., Gibson T.J., Karplus K., Li W., Lopez R., McWilliam H., Remmert M., Söding J., Thompson J.D., Higgins D.G. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega, Molecular Systems Biology, 7:539.
53. Singh A., Pal T.K. (2014). *In silico* sequence analysis, structure prediction and function annotation of human Bcl-X beta protein, Int J Bioautomation, 18(1), 23-30.
54. Sippl M.J. (1993). Recognition of errors in three-dimensional structures of proteins, Proteins: Structure, Function, and Bioinformatics, 17(4), 355-362.
55. Srinivasan N., Blundell T.L. (1993). An evaluation of the performance of an automated procedure for comparative modelling of protein tertiary structure, Protein Eng, 6, 501-512.
56. Sternberg M.J.E., Bates P.A., Kelley L.A., MacCallum R.M. (1999). Progress in protein structure prediction: Assessment of CASP3, Curr Opin Struct Biol, 9, 368-373.
57. Szafron D., Lu P., Greiner R., Wishart D.S., Poulin B., Eisner R., Lu Z., Anvik J., Macdonell C., Fyshe A., Meeuwis D. (2004). Proteome Analyst: Custom predictions with explanations in a web-based tool for high-throughput proteome annotations, Nucleic Acids Res, W365-W371.
58. Thompson J.D., Higgins D.G., Gibson T.J. (1994). CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, Nucleic Acids Res, 22(22), 4673-4680.

59. Ullman C.G., Haris P.I., Smith K.F., Sim R.B., Emery V.C., Perkins S.J. (1995). Betasheet secondary structure of an LDL receptor domain from complement factor I by consensus structure predictions and spectroscopy, FEBS Lett, 371, 199-203.
60. Vriend G. (1990). WHAT IF: A molecular modeling and drug design program, J Mol Graph, 8(1), 52-56.
61. Wang Y., Wishart D.S. (2005). A simple method to adjust inconsistently referenced 13C and 15N chemical shift assignments of proteins, J Biomol NMR, 31, 143-148.
62. Westbrook J., Feng Z., Jain S., Bhat T.N., Thanki N., Ravichandran V., Gilliland G.L., Bluhm W., Weissig H., Greer D.S., Bourne P.E., Berman H.M. (2002). The Protein Data Bank: Unifying the archive, Nucleic Acids Res, 30, 245-248.
63. Wishart D.S., Case D.A. (2000). Use of chemical shifts in macromolecular structure determination, Methods Enzymol, 338, 3-34.

Muhammad Saad Khan, B.Sc.
E-mail: sadihere_007@hotmail.com

Muhammad Saad Khan completed B.Sc. in Bioinformatics from Mohammad Ali Jinnah University, Islamabad. His fields of interest are homology modelling and computer-aided drug design. He has command over many bioinformatics tools. He is doing M.Sc. in Bioinformatics at Mohammad Ali Jinnah University, Islamabad.

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

Dr. Sahar Fazal got her Ph.D. from South China Agricultural University, China. Presently she is working as an Associate Professor in the Department of Bioinformatics, Mohammad Ali Jinnah University, Islamabad. Her field of interests is molecular modelling.