In silico Sequence Analysis, Structure Prediction and Function Annotation of Human Bcl-X Beta Protein

Anjali Singh, Tapan Kumar Pal*

Department of Biotechnology
Bengal Institute of Technology
Hadia, Pin-700150, West Bengal, India
E-mail: tapankpal@gmail.com

**Corresponding author*

Received: November 11, 2013 **Accepted: March 14, 2014**

Published: March 28, 2014

Abstract: *Bcl-X proteins are the one of the best categorized member of the Bcl-2 protein families which acts as primary regulators of apoptosis in mammalian cells. The Bcl-X proteins are potential anti-cancer drug targets. In this study, the tertiary structure of the beta isoform of the apoptosis regulator Bcl-X in humans (h-Bcl-X_β) has been predicted by fold-recognition (threading) approach. In silico assessment of the h-Bcl-X_β protein revealed the characteristic structural features of anti-apoptotic Bcl-2 protein family in h-Bcl-X_β protein. The predicted model was comprised of BH1-BH4 domains, seven alpha-helices and a C-terminal transmembrane domain for membrane localization and sub-cellular targeting. Quality assessment of the predict model confirmed its reliability as fairly good model. Active sites of h-Bcl-X_β protein were identified using CASTp server. The future work can be directed towards drug designing for cancer treatment by regulating the activity of h-Bcl-X_β proteins.*

Keywords: *h-Bcl-X_β, Threading, Apoptosis, Cancer.*

Introduction

The living cells divide in repeated fashion for growth, reproduction, regeneration and replacement of damaged cells. The removal of damaged cells comprises of controlled sequence of events in which cells undergoes self-termination by apoptosis [5]. An inhibition in apoptosis might result in unrestrained tumor growth, which can be cancerous [14, 18] whereas enhanced apoptosis may lead to unwanted depletion in cell population as reported in neurodegenerative disorders like Alzheimer's and Parkinson's diseases [2, 12]. The members of the Bcl-2 (B-cell lymphoma 2) family proteins play a very substantial role in promotion or inhibition of apoptotic processes [11, 23]. Bcl-2 proteins play a shielding role in cell survival by blocking apoptosis [16, 19]. All the members contain conserved Bcl-2 Homology (BH) domains. The Bcl-2 subfamily proteins contain at least BH1 and BH2 which promote cell survival by inhibiting caspases activation [22]. Bcl-X proteins are the one of the best categorized member of these Bcl-2 protein families. They are leading regulator of apoptosis in mammalian cells [3]. Alternative RNA splicing generates three isoforms of the apoptosis regulator Bcl-X. The long isoform Bcl-X_L has apoptosis repressor activity [10], whereas the short isoform Bcl-X_S promotes apoptosis [13] but the role of beta-isoform Bcl-X_β in humans is still not clear [1].

In order to gain further understanding of the physiological nature of the Bcl-X_β proteins, the current study aims at *in silico* sequence analysis, structure modeling and function annotation of human Bcl-X_β (h-Bcl-X_β) proteins.

Methodology

Sequence retrieval

The complete primary sequence of h-Bcl-X_β protein (accession No: AAB17354) was retrieved from Genbank database at NCBI. The length of the h-Bcl-X_β protein was 227 amino acids.

Sequence analysis

Physical and chemical parameters of the protein was computed by ProtParam tool (http://web.expasy.org/protparam/) which analyzed for molecular weight, theoretical pI, amino acid composition, and grand average of hydropathicity (GRAVY). HSLPred [6] and Hum-mPLoc v2.0 [17] servers were used to predict the subcellular localization of h-Bcl-X_β protein. TMPred server [8] was used to check the presence of the transmembrane domains within the protein.

Secondary structure prediction

SOPMA server [7] was used for the secondary structure prediction of the h-Bcl-X_β protein. It assessed the conformational information about positional possibility of the β-strands and turns, random coils and α-helices within the protein structure.

Tertiary structure prediction

The protein sequence of h-Bcl-X_β protein was submitted to the automated modeling server, I-TASSER [24] (http://zhanglab.ccmb.med.umich.edu/I-TASSER) for tertiary structure modeling. I-TASSER server predicts the protein structure by threading (fold recognition) approach by choosing the templates from proteins of known structures in PDB databank having same fold as target protein. The HEX [15] interface and RASMOL were used for visualization of the generated model. The model with highest C-score value was selected for model refinement and validation. A higher C-score of value signifies a model with a high confidence.

Model refinement and validation

The predicted model was submitted to automated protein structure refinement server – $3D^{refine}$ (http://sysbio.rnet.missouri.edu/3Drefine) for optimization of hydrogen bonding network and energy minimization. After refinement, the structural evaluation and stereo-chemical assessment of the model were performed using ProSA-web [21] and VERIFY3D. The Psi/Phi Ramachandran plot was generated by RAMPAGE sever.

Functional characterization of the predicted structure

The functional assessment of the predicted model was done by searching against Pfam database. The ligand binding site and active site residues of the modeled h-Bcl-X_β protein was assessed by CASTp [4] and 3DLigandSite [20] servers respectively.

Model Submission

The predicted model of h-Bcl-X_β protein was successfully submitted in Protein Model Data Base (http://mi.caspur.it/PMDB) having PMID PM0079226.

Results and discussion

Primary sequence analysis

The physiochemical properties of h-Bcl-X_β protein were assessed by ProtParam tool. The h-Bcl-X_β protein was predicted to have molecular weight of 25290.3 Daltons and the

theoretical isoelectric point (pI) of 4.73, indicating that the protein is negatively charged. The negative Grand average of hydropathicity (GRAVY) value of -0.175 for protein indicates that it is hydrophilic and soluble in nature. The subcellular localization prediction using HSLPred and Hum-mPLoc v2.0 servers predicted that the query protein is cytoplasmic and localized in mitochondria. Furthermore, TMPred server predicted that the sequence positions 137-155 and 207-225 are probable transmembrane helix regions of the protein.

Secondary structure prediction

The secondary structure prediction of h-Bcl-X_β protein using SOPMA (with default parameters) showed that the protein having the composition of Helix = 48.46%, Strand = 6.61%, and Coil = 44.93% (Fig. 1). As evident from this secondary structure prediction, h-Bcl-X_β protein is mostly comprised of alpha helix and loops with traces of beta turns and strands.

Tertiary structure prediction by threading approach

The 3D structure of the h-Bcl-X_β protein was generated using I-TASSER protein prediction server. In accordance with the secondary structure prediction, the predicted 3D structure of the h-Bcl-X_β protein (Fig. 2) mainly comprised of alpha helices and beta turns/coils.

Interestingly, the predicted 3D model of h-Bcl-X_β protein has the characteristic features of anti-apoptotic members, namely BH1-BH4 domains, seven alpha-helices and a C-terminal transmembrane (TM) domain for membrane localization and sub-cellular targeting. The presence of these features suggests that h-Bcl-X_β protein is anti-apoptotic in nature.

```
        10        20        30        40        50        60        70
         |         |         |         |         |         |         |
MSQSNRELVVDFLSYKLSQKGYSWSQFSDVEENRTEAPEGTESEMETPSAINGNPSWHLADSPAVNGATG
hccchhhhhhhhhhcccccccccccceehccccccccccccccccccccccccccccccccccccccc
HSSSLDAREVIPMAAVKQALREAGDEFELRYRRAFSDLTSQLHITPGTAYQSFEQVVNELFRDGVNWGRI
hhhccccccccchhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhheecccccchhhhhhhhhhhhhttccttee
VAFFSFGGALCVESVDKEMQVLVSRIAAWMATYLNDHLEPWIQENGGWVRTKPLVCPFSLASGQRSPTAL
eeeeetttceehhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhtttccceeecccccccchhccccchhhh
LLYLFLLCWVIVGDVDS
hhhhhhhhhhhhtchhh

Sequence length :   227

SOPMA :
   Alpha helix     (Hh) :   111 is  48.90%
   3₁₀ helix       (Gg) :     0 is   0.00%
   Pi helix        (Ii) :     0 is   0.00%
   Beta bridge     (Bb) :     0 is   0.00%
   Extended strand (Ee) :    16 is   7.05%
   Beta turn       (Tt) :    11 is   4.85%
   Bend region     (Ss) :     0 is   0.00%
   Random coil     (Cc) :    89 is  39.21%
   Ambigous states (?)  :     0 is   0.00%
   Other states         :     0 is   0.00%
```

Fig. 1 Secondary structure prediction of h-Bcl-X_β protein using

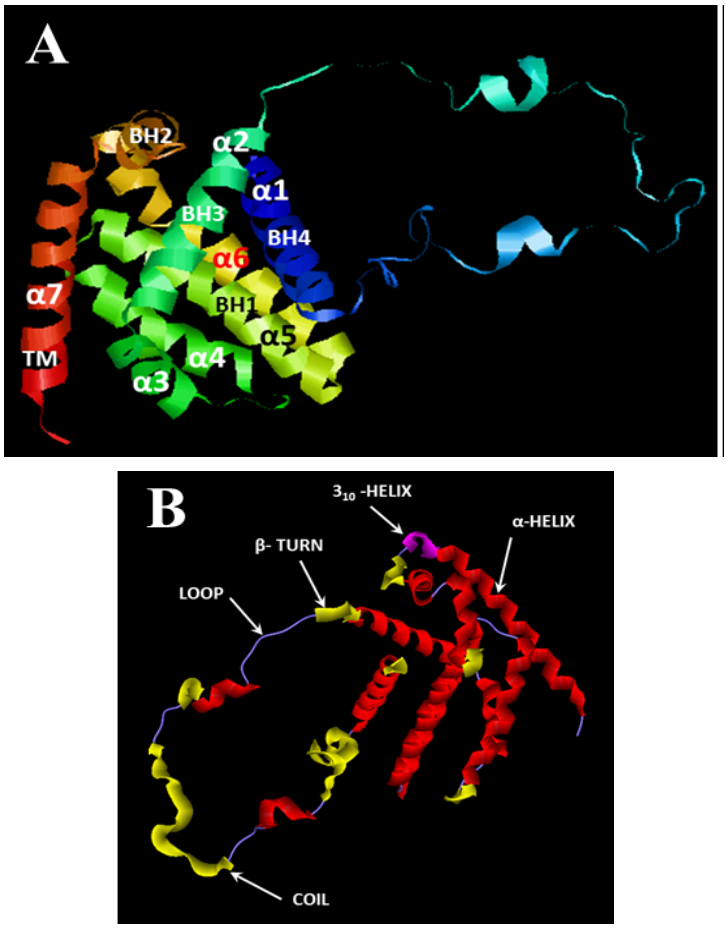


Fig. 2 Predicted tertiary structure of h-Bcl-X_β protein, viewed in (A) RasMol and (B) HEX interface

Model refinement and quality assessment

The predicted 3D model was submitted $3D^{refine}$ server for refinement and the refined structure was validated by VERIFY3D, ProSA-web and Ramachandran Plot assessments. The predicted structure passed the VERIFY3D evaluation with 80.26% of the residues having an averaged 3D-1D score > 0.2. The residues with score over 0.2 suggest that the predicted tertiary structure is of good quality. The calculated Z-score of -6.06 (Fig. 3A) by ProSA-web indicated that the overall quality of the predicted tertiary structure lies within the score range of experimentally determined protein tertiary structures by NMR and X-ray crystallography. RAMPAGE server was used to determine the Ramachandran Plot to assure the model quality. The Ramachandran Plot revealed that 91.6% (> 90% for a good model) of the overall residues are in the most favored region (Fig. 3B).

Functional annotation of the predicted structure

The predicted model was searched against the Pfam database for conserved domains using BLAST by POLYVIEW-3D. The results showed that the modeled h-Bcl-X_β protein structure belongs to the family of Bcl-2 apoptosis regulator proteins with homologous BH4 domains (Fig. 4). The ligand binding sites (pockets) of the predicted model was assessed by CASTp server. Out of 35 predicted structural pockets, the pockets with volume > 500 $Å^3$ were reported (Fig. 5). Pocket 1 (green) is located in cavity between α1, α5 and α6 helices and Pocket 2 (blue) is located in cavity between α7 helix and BH2 domain of α6 helix.

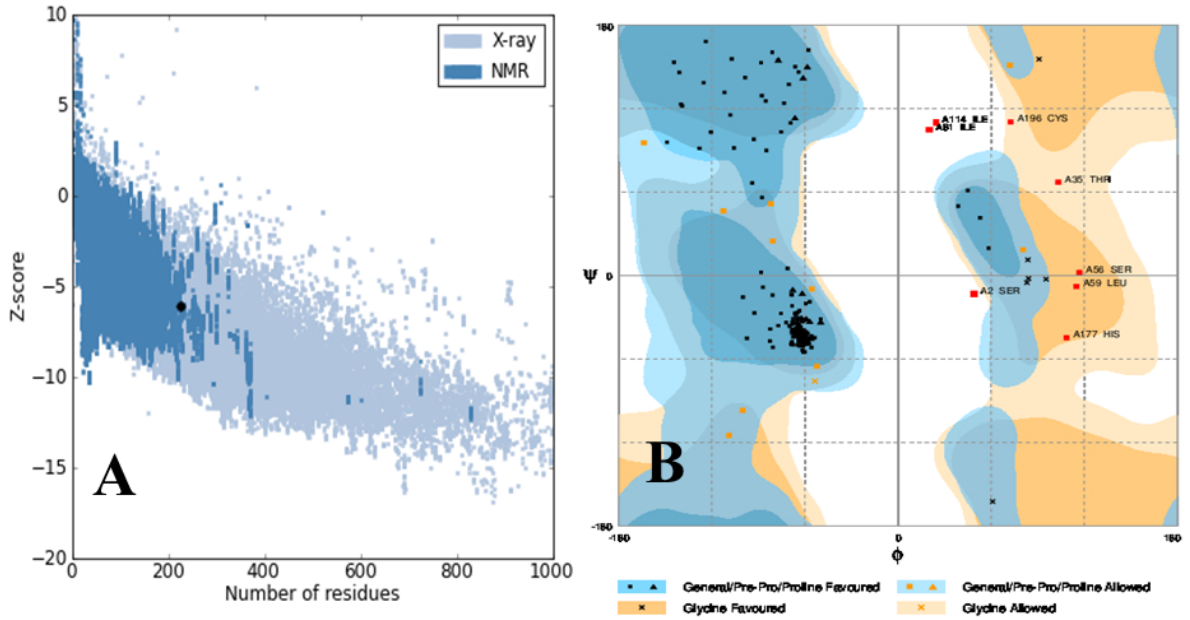
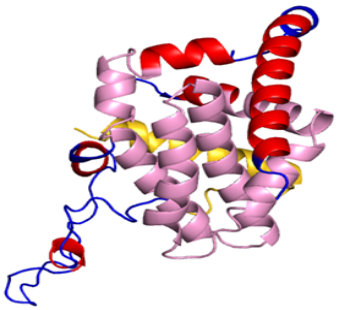
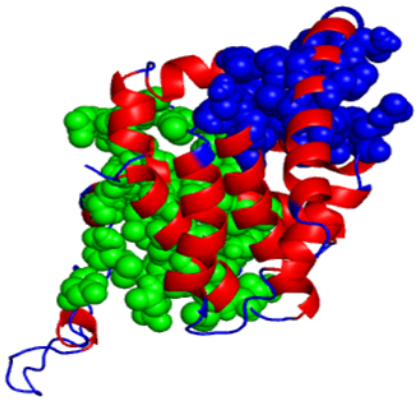


Fig. 3 (A) Z-score plot and (B) Ramachandran plot for predicted tertiary structure of h-Bcl-X_β protein



Pfam Entry	E-value	Description	Residues
Family Bcl-2 (Pink)	7e-53	Apoptosis regulator proteins, Bcl-2 family	90-188
Family BH4 (Gold)	2e-07	Bcl-2 homology region 4	1-27

Fig. 4 Identification of superfamily for predicted model by protein annotation (as viewed in POLYVIEW-3D)



Pocket	Area, Å^2	Volume, Å^3	Residues
1 (green)	701.0	2074.2	3, 4, 6, 7, 9 ,10 ,11, 13, 14, 24 ,25, 27, 28, 29, 30, 31, 33, 36, 72, 73, 74, 75, 76, 77, 78, 80, 81, 83, 87,160, 161, 164, 165, 167, 168, 171
2 (blue)	439.7	641.7	92, 93, 96, 97, 100, 136, 137, 138, 141, 194, 195, 196, 199, 204, 207, 208, 210, 211, 212, 214, 215

Fig. 5 Predicted ligand binding sites of the tertiary structure of h-Bcl-X_β protein (as viewed in POLYVIEW-3D).

Conclusion

The main objective of the study was to perform sequence analysis and predict the tertiary structure and function of h-Bcl-X_β protein. The 3D structure modeling resulted in high quality structure of h-Bcl-X_β protein [1] with all the unique characteristic features of Bcl-2 anti-apoptotic protein family. This suggested that h-Bcl-X_β protein is an anti-apoptotic protein. The predicted ligand binding sites can be used for designing anti-cancer drugs targeting h-Bcl-X_β proteins [9] and to understand the protein-protein interaction between the anti-apoptotic and pro-apoptotic proteins.

Acknowledgements

The authors thank the Head, Department of Biotechnology, Bengal Institute of Technology for providing facilities required successful completion of the project.

References

1. Ban J., L. Eckhart, W. Weninger, M. Mildner, E. Tschachler (1998). Identification of a Human cDNA Encoding a Novel Bcl-x Isoform, Biochemical and Biophysical Research Communications, 248, 147–152.
2. Behl C. (2007). Apoptosis and Alzheimer's Disease, Journal of Neural Transmission, 107, 1325-1344.
3. Boise L. H., M. González-Garcia, C. E. Postema, L. Ding, T. Lindsten, L. A. Turka, X. Mao, G. Nunez, C. B. Thompson (1993). Bcl-x, a Bcl-2-related Gene that Functions as a Dominant Regulator of Apoptotic Cell Death, Cell, 74, 597-608.
4. Dundas J., Z. Ouyang, J. Tseng, A. Binkowski, Y. Turpaz, J. Liang (2006). CASTp: Computed Datas of Surface Topography of Proteins with Structural and Topographical Mapping of Functionally Annotated Residues, Nucleic Acids Research, 34, W116-W118.
5. Elmore S. (2007). Apoptosis: A Review of Programmed Cell Death, Toxicologic Pathology, 35, 495-516.
6. Garg A., M. Bhasin, G. Raghava (2005). SVM-based Method for Subcellular Localization of Human Proteins using Amino Acid Compositions, their Order and Similarity Search, The Journal of Biological Chemistry, 280, 14427-14432.
7. Geourjon C., G. Deleage (1995). SOPMA: Significant Improvements in Protein Secondary Structure Prediction by Consensus Prediction from Multiple Alignments, Computer Applications in the Biosciences, 11, 681-684.
8. Hofmann K., W. Stoffel (1993). TMbase – A Database of Membrane Spanning Proteins Segments, Biological Chemistry Hoppe-Seyler, 374, 166.
9. Huang Z. (2000). Bcl-2 Family Proteins as Targets for Anticancer Drug Design, Oncogene, 19, 6627-6631.
10. Jeong H. S., H. Y. Choi, T. W. Choi, B. W. Kim, J. H. Kim, E. R. Lee, S. G. Cho (2008). Differential Regulation of the Anti-apoptotic Action of B-cell Lymphoma 2 (Bcl-2) and B-cell Lymphoma Extra-long (Bcl-xL) by c-Jun N-terminal Protein Kinase (JNK) 1-Involved Pathway in Neuroglioma Cells, Biological & Pharmaceutical Bulletin, 31, 1686-1690.
11. Korsmeyer S. J. (1992). Bcl-2: An Antidote to Programmed Cell Death, Cancer Surveys, 15, 105-118.
12. Lev N., E. Melamed, D. Offen (2003). Apoptosis and Parkinson's Disease, Progress in Neuro-Psychopharmacology and Biological Psychiatry, 27, 245-250.
13. Lindenboim L., J. Yuan, R. Stein (2000). Bcl-xS and Bax Induce Different Apoptotic Pathways in PC12 cells, Oncogene, 19, 1783-1793.
14. Lowe S. W., A. W. Lin (2000). Apoptosis in Cancer, Carcinogenesis, 21, 485-495.

15. Macindoe G., L. Mavridis, V. Venkatraman, M. D. Devignes, D. W. Ritchie (2010). Hex Server: an FFT-based Protein Docking Server Powered by Graphics Processors, Nucleic Acids Research, 38, W445-W449.
16. Reed J. C. (2008). Bcl-2-family Proteins and Hematologic Malignancies: History and Future Prospects, Blood, 111, 3322-3330.
17. Shen H., K. Chou (2009). A Top-down Approach to Enhance the Power of Predicting Human Protein Subcellular Localization: Hum-mPLoc 2.0, Analytical Biochemistry, 394, 269-274.
18. Taylor R. C., S. P. Cullen, S. J. Martin (2008). Apoptosis: Controlled Demolition at the Cellular Level, Nature Reviews of Molecular Cell Biology, 9, 231-241.
19. Van Delft M. F., D. C. Huang (2006). How the Bcl-2 Family of Proteins Interact to Regulate Apoptosis, Cell Research, 16, 203-213.
20. Wass M. N., L. A. Kelley, M. J. Sternberg (2010). 3DLigandSite: Predicting Ligand-binding Sites using Similar Structures, Nucleic Acids Research, 38, W469-W473.
21. Wiederstein M., M. J. Sippl (2007). ProSA-web: Interactive Web Service for the Recognition of Errors in Three-dimensional Structures of Proteins, Nucleic Acids Research, 35, W407-W410.
22. Willis S., C. L. Day, M. G. Hinds, D. C. S. Huang (2003). The Bcl-2-regulated Apoptotic Pathway, Journal of Cell Science, 116, 4053-4056.
23. Youle R. J., A. Strasser (2008). The BCL-2 Protein Family: Opposing Activities that Mediate Cell Death, Nature Reviews Molecular Cell Biology, 9, 47-59.
24. Zhang Y. (2008). I-TASSER Server for Protein 3D Structure Prediction, BMC Bioinformatics, 9, 40, doi:10.1186/1471-2105-9-40.

Tapan Kumar Pal, Ph.D.
E-mail: tapankpal@gmail.com

Dr. Tapan Kumar Pal is currently an Assistant Professor, Department of Biotechnology, Bengal Institute of Technology, Kolkata. He has completed B.Sc. in Chemistry and M.Sc. in Biochemistry from University of Calcutta. He has completed his Ph.D. in Biochemistry from University of Calcutta. With more than 11 years of teaching and research experience, he has contributed more than 10 publications in journals in national and international levels. He is also a guest faculty, Department of Home Science, University of Calcutta and Department of Food and Nutrition, Sarada Ma Girl's college, Barasat. His areas of interests are microbial biotechnology, plant protection and bioinformatics.

Anjali Singh
E-mail: singhanjali123@gmail.com

Anjali Singh is presently in pre-final year of her B.Tech degree in Biotechnology from Bengal Institute of Technology, Kolkata, India. Her research interests include bioinformatics, molecular biology and genetics.