Structure prediction of Bestrophin for the induced-fit docking of anthocyanins

Sathyan Sri Lavvanya Priya1*, Ponnuswamy Renuka Devi1, Palanisami Eganathan2 & Nishith Saurav Topno3

1Department of Biotechnology, Anna University of Technology, Coimbatore- 641 047, Tamilnadu, India; 2Plant Tissue Culture and Bioprospecting Laboratory, M. S. Swaminathan Research Foundation, 3rd Cross Road, Institutional Area, Taramani, Chennai- 600 113. India; 3Centre for Bioinformatics, School of Life Sciences, Pondicherry University, Kalapet, Puducherry- 605 014, India; Sathyan Sri Lavvanya Priya – Email: lavanya19@yahoo.com; Phone: 91-9566005811; *Corresponding author

Received July 26, 2012; Accepted August 13, 2012; Published August 24, 2012

Abstract:
Bestrophin, an integral membrane protein existing in basolateral region of the retina is a propitious target for drug discovery. Mutations in the Bestrophin protein cause Best Vitelliform Macular Dystrophy (BVMD) leading to retinal damages and loss of visual acuity. Owing to the lack of three dimensional structure and related structural homologs in the protein data bank, we modeled the bestrophin protein using Robetta ab initio method. Further, no treatment is available for the disease. In this situation, anthocyanins from natural sources are reported to combat retinal damages. Hence, we identified anthocyanins from Syzygium cumini fruit skin using Electrospray Ionization tandem mass spectrometry. These compounds were docked into the predicted bestrophin model to study the interactions within the active site. The results may provide a valuable insight into the structure of bestrophin and efficacy of anthocyanins in molecular docking studies.

Abbreviations: PTP-Putative transmembrane proteins, VMD- Vitelliform macular dystrophy, BVMD- Best's vitelliform macular dystrophy, RPE- Retinal pigment epithelium, ESI-MS/MS- Electrospray Ionization Tandem Mass Spectrometry, UNIPROT-Universal Protein Resource, PSIPRED- Protein secondary structure prediction, TMH- Transmembrane Helices, SCFS- Syzygium cumini fruit skin, DP - Declustering Potential, IFD- Induced Fit Docking

Key words: Bestrophin, Ab initio, Anthocyanin, Induced Fit Docking

Background:
PTPs are attributed a cardinal role in numerous physiological functions [1]. Owing to difficulty in crystallizing, computational methods are employed to predict the PTPs as the basis for structure based drug design.

Bestrophin being a PTP is encoded in the gene VMD2. Mutations in the VMD2 gene causes BVMD, an inherited progressive macular dystrophy, first identified in 1905 by a German ophthalmologist Frederich Best. BVMD advances with juvenile onset causing loss of visual acuity due to atrophic macular changes or choroidal neovascularization associated with sub retinal hemorrhages and fibrosis [2]. Early stages of the disease are characterized by abnormal depositions of lipofuscin-like material at the level of the RPE classically resembling an egg yolk [3]. Disintegration of the central yellow lesions progressively leads to vision loss at a later stage. The elemental cause of the Best disease is the BEST1 gene (previously called VMD2, currently as hbest1), identified in 1998 [4].

Determining the three-dimensional structures of PTPs remains a challenge, although they constitute 15-30% of the entire genome. Bestrophin being a plasma membrane protein has a
BIOINFORMATION

molecular weight of 68 kDa and consists of 585 amino acids [5]. Computationally, bestrophin is predicted as a transmembrane protein with four membrane-spanning α helical domains, while presence of at least five helices are reported by few studies [6]. An ecumenical outcome suggests that Bestrophin functions as a chloride ion channel [7]. The homology modeling of Asp-rich domain of hbest1 identified two calcium binding sites, yet the complete structural details of the transmembrane protein are not available [8]. Obtaining the protein structure is crucial to understand protein function which eventually leads to drug designing [9, 10]. Hence the hbest1 protein was modeled using the Robetta web server (http://robetta.bakerlab.org) [11].

Natural compounds are tangible as therapeutic drug targets. More significantly, anthocyanins belonging to the flavonoid group are beneficial in curing visual acuity [12]. Anthocyanins are water-soluble pigments, pre-eminent in a variety of plants mainly imparting colour to flowers and fruits [13]. Innumerable reports exist for the rich resource of anthocyanins in the fruits of Syzygium cumini (L.) Skeels (Black Plum) [14, 15]. Our study also involves quantification of anthocyanins from the fruit skin by ESI – MS/MS Mass Spectrometer. Small molecular structures of the separated anthocyanins were obtained from publicly available chemical databases and used for molecular docking with hbest1. According to the literature, Resveratrol and Niflumic acid found to be useful in combating retinal damages was also used in our docking studies [16].

Figure 1: (A) Secondary structure prediction of hbest1 sequence using PSI-PRED server; (B) Graph representing propensities of TM helices. Numbers circled in red are the positions of TM helices; (C) Robetta modeled structure of hbest1
Methodology:

Retrieval of Target Sequence
The FASTA format of human Bestophin-1 (hbest1) with Accession number O76090 was obtained from the UNIPROT database [17]. The protein sequence contains 585 amino acid residues.

Secondary structure Prediction
The secondary structure of hbest1 was predicted using PSIPRED v3.0 [18, 19]. Neural nets are used to convert PsiBlast profile data to secondary structure propensities. A putative secondary structure is obtained for each residue associated with a confidence value for the prediction.

Prediction of Transmembrane regions
The predictors available are inaccurate in predicting the ends of TMHs, or TMHs of unusual length [20]. MemBrain, a predictor based on machine learning approach was used to predict the TMHs of hbest1.

Ab Initio modeling
Pertaining to lack of suitable structural template for hbest1, we resorted to Ab initio modeling of the hbest1 using Robetta, a full chain Protein prediction server. The primary sequence was submitted to the server, and the generated models were received by email.

Structure Validation
Five structures were obtained from Robetta server. The qualities of the models were evaluated using the ProQ webserver [21]. The models were validated using PROCHECK program [22].

Experimental
Fruit sample
Mature fruits of Syzygium cumini were collected from Sirumalai hills of Dindigul, Tamilnadu, India and identified by Dr P. Eganathan, M. S. Swamnathan Research Foundation, Chennai. Rotten, damaged fruits were removed; skin was peeled from fruits without pulp and shade dried under room temperature (31°C) until complete disappearance of moisture. Dried skin was stored for further analysis.

Solvents and Reagents
Analytical grade chemicals and solvents purchased from Sigma Aldrich (Saint Louis, MO) were used in the present investigations.

Preparation of extract for ESI- MS/MS analysis
SCFS was ground to a fine powder using a Thomas Wiley Machine (Model 5 USA) at room temperature. Subsequently, 50g of powdered plant material was extracted with 1.5 L methanol using soxhlet apparatus at 65°C for 4 hours consecutively. The solvent was removed in vacuo using a Buchi Heating Bath (B-490) rotavapor to yield dried methanol (purple color) extract. The extract was filtered and evaporated to dryness in a vacuum at 40°C with a rotary evaporator.

ESI - MS/MS Analysis
Crude methanol extract of SCFS was diluted with methanol and filtered with 0.22µm nylon membrane filter and subjected to ESI-MS/MS analysis. Anthocyanin identification was performed on a 3200 QTRAP instrument (ABSciex Instruments, Singapore) equipped with Linear Ion Trap Quadrupole Mass Spectrometer and electrospray ionization (ESI) source. Mass parameters were adjusted to get the maximum sensitivity. Data was generated by Analyst 1.4.1 software. The MS-MS conditions are : positive ion mode; gas (N2), curtain gas was set to 15 psi, heater gas and nebulizer gas were set to 10 psi and source temperature maintained at ambient. The positively identified compounds (cyanidin, petunidin, malvidin) were subjected to MS-MS analysis to study the fragment patterns and were found to match with that of the earlier reported compounds.

Docking
The active site residues are Cys69, Cys42, Cys23, Phe80, Phe84, Val86, Pro77, Leu82, Tyr85, Gly83, Arg92, Trp94, Trp93 [7, 23]. Cyanidin 3, 5 Diglucoside, Malvidin 3, 5 diglucoside Petunidin 3, 7 diglucoside were docked with hbest1. In addition, other compounds included in docking were peonidin, delphinidin (other anthocyanins). Niflumic acid [24] and resveratrol [25] were included as positive controls in our docking studies for standardization.

Preparation of protein and ligands
Glide (Grid-based Ligand Docking with Energetics) software, developed by Schrödinger running on Red Hat Enterprise Linux 5 (RHEL 5) workstation, was used for Protein, Ligand preparation and Induced Fit Docking.

The hbest1 structure was prepared with Protein Preparation module Wizard of Glide software. Water molecules were removed, the hydrogens atoms were added to the protein and all atom force filed (OPLS -2001) charges and atom types were assigned.

The CID files of the ligands Cyanidin 3,5 Diglucoside (CID 441688), Delphinidin 3,5 glucoside (CID 10100906), Malvidin 3, 5 diglucoside (CID 441765), Petunidin 3,7-diglucoside (CID 44256973), Peonidin 3,5-diglucoside (CID 44256843) Niflumic acid (CID 4488), Resveratrol (CID 445154) were acquired from the NCBI Pubchem database [26]. Using the Impact module of glide the ligands were minimized with 1000 cycles of steepest gradient and 5000 cycles of conjugate gradient.

IFD protocol
Induced fit docking combines Glide and Prime modules to arrive at accurate prediction of ligand binding modes and concomitant structural changes in the receptor [27]. Systematic and Simulation methods are adopted by glide for searching poses and ligand flexibility. Incremental construction for searching is employed by the systematic method, with Gscore being the empirical scoring function [28]. The calculation of GScore in Kcal/mol is: G-Score = H bond + Lipo+ Metal + Site + 0.130 Coul + 0.065vdW + Bury P - RotB. Where Hbond= Hydrogen bonds, Lipo = hydrophobic interactions, Metal – metal binding term, Site = Polar interactions in the binding site, vdW = Vander Waals forces, Coul = coulombic forces, Bury P= penalty for buried polar group, RotB= freezinf rotatable bonds.

The prepared protein was docked with the minimized ligands. The active sites in the protein hbest1 were selected to be docked with the ligand. IFD was performed and best conformations were selected based on Glide Score, Glide energy, and Glide e-model scores.
Hydrophobic interactions
Ligplot was used for analyzing the hydrophobic contacts of the protein and ligand [29]. The best pose for each ligand was submitted to Ligplot server and results obtained.

Figure 2: Protein-ligand interaction 2D map of hbest1 and inhibitors (A-G) using Ligplot diagram. Green dotted lines are Hydrogen bond interactions and red semicircles are amino acid residues of the protein showing hydrophobic interactions. (A) cyanidin 3,5 diglucoside; (B) Malvidin 3,5 diglucoside; (C) Petunidin 3, 7 diglucoside; (D) Delphidin 3, 5 diglucoside; (E) Peonidin 3, 5 diglucoside; (F) Niflumic acid; (G) Resveratrol.
Results & Discussion: The percentage of secondary structures predicted by PSIPRED showed that 35.90% of the total structures were alpha helices, 3.07% were beta sheets and 59.48% were random coils within the target sequence. The protein was predicted to have 16 helices, 18 coils and 4 beta sheets (Figure 1A). The percentage of coils is more than the helices in the sequence. The C-terminal region of the hbest1 is composed of about 150 residues indicating the presence of a randomly coiled region. The N-terminal region, predicted to have signal peptide region has also a coiled region. Secondary structure prediction provides valuable information about the content of the protein which provides insights into the tertiary structure of hbest1.

Prediction of TMHs in helical membrane proteins provides valuable information about the protein topology when the high resolution structures are not available. The predictors available are inaccurate in predicting the ends of TMHs, or TMHs of unusual length [20]. The prediction accuracy of MemBrain is 97.9%. N-terminal signal peptides were also detected. Five Transmembrane helices were predicted by MemBrain (Figure 1B).

The models obtained for hbest1 from Robetta server were evaluated by the ProQ program with the following results: model-1 (ProQ-LG=1.267, ProQ-MX=0.080), model-2 (ProQ-LG=1.088, ProQ-MX=0.075), model-3 (ProQ-LG=1.171, ProQ-MX=0.061), model-4 (ProQ-LG=1.852, ProQ-MX=0.062), model-5 (ProQ-LG=2.310, ProQ-MX=0.104). The cutoffs to finalize the best model are ProQ-LG > 1.5 or ProQ-MX > 0.1 Model-5 is in complete agreement with the cutoffs providing a good modeled structure (Figure 1C).

Validation of the model by PROCHECK presented a Ramachandran plot analysis rendering 98.1% residues in the most favoured regions, 1.9% in the additionally allowed regions indicating the efficacy of Ab initio modeling. Secondary structural elements of the predicted hbest1 were found to be almost similar to the PSIPRED server results of the primary sequence.

Docking studies suggest the ability of the three anthocyanins to bind to active and additional active sites around hbest1. The binding poses and interactions were analyzed by Glide Table 1 (see supplementary material) & (Figure 2). The glide energy, glide score and glide emodel score for Cyanidin 3, 5 glucoside were -64.34 Kcal/mol and -11.80 Kcal/mol -71.60 Kcal/mol respectively. The compound interacts with Leu 82, Gln59, Arg105, and Arg47 that are the active sites within 20 A° distance of the literature cited residues. Hydrophobic contacts are with Phe82, Phe48, and Thr66. Malvidin 3, 5 diglucoside generated glide energy of -71.61 kcal/mol, glide score of -5.15 Kcal/mol and glide emodel score of -92.42 Kcal/mol. Interacting residues are Arg92, Arg461, Arg534, and Thr 87. Arg92 is the active site in interaction with Malvidin. Residues Ile129, Thr 91, Pro152, Asn 136 are in the hydrophobic pockets of docked structure. Petunidin 3, 7 diglucoside provided glide energy of -68.39 Kcal/mol, a glide score of -7.98 Kcal/mol and glide emodel score of -94.47 Kcal/mol. Asn136, Thr87, Arg461, are the residues in interaction besides Arg92 and Trp94 that are the two active sites involved in the interactions. Hydrophobic interactions comprise of Pro152 and Val90.peonidin 3, 5 diglucoside resulted in a glide energy of -58.82 Kcal/mol, glide score of -8.47 Kcal/mol and glide emodel score of -85.17 Kcal/mol. Hydrogen bond interactions were exhibited with Asn133, Arg 534, and Gln460. Trp94, Val90, Pro406 residues were found in hydrophobic contacts. A glide energy of -77.17 Kcal/mol, glide score of -9.18 and a glide emodel score of -107.57 Kcal/mol was generated by delphinidin 3,5 diglucoside. Hydrogen bond interactions include Gly159, Glu382, Ile49, Arg461, and Gln 96, while hydrophobic interactions involve active sites at Arg92, and Trp93. All anthocyanins have
favourable binding potential with the involvement of the active site or neighboring residues.

Resveratrol rendered glide energy of -30.89 Kcal/mol, glide score of -8.25 Kcal/mol and glide emodel score of -48.71 with one interaction at Tyr131. Similarly the glide energy for Niflumic acid was -38.19 Kcal/mol, glide score of -8.80 and glide emodel score being -49.34 Kcal/mol. Niflumic acid has three interactions including an active site interaction at Arg92.

ESI-MS/MS, an accomplished method [30] for determining anthocyanins shows the presence of cyanidin, petunidin, malvidin in the methanolic extract of SCFS (Figure 3). In accordance to earlier studies [31, 32], the chromatographic profile displays the domination of anthocyanins in the sample. The identified compounds were Malvidin diglucoside at peak 655, Cyanidin 3, 5 diglucoside at peak 611 and, peak 641 corresponds to Petunidin 3, 7 diglucoside. Anthocyanins from other berries, amenable for oral intake, are reported to cure visual acuity [33]. Our study indicates the potential of the three anthocyanins as inhibitors for BVMD, whereas, pronounced results are not obtained with other compounds. The inherent capability of anthocyanins to protect the retinal pigment epithelium from age related macular degeneration has been reported [34]. The tangibility of anthocyanins as antioxidants has been substantiated by many authors [35, 36]. In conclusion, we imply that anthocyanins are safe [37] to use for BVMD and similar retinopathies.

Conclusion:
Mutations in bestrophin causes age related macular degeneration eventually leading to loss of central vision. Currently no specific drug is available. Anthocyanin compounds are known to exert a positive effect in the treatment of visual acuity. Our present work provides an understanding into the structure of hbest1, identification of anthocyanins in the Syzygium cumini fruit skin, and docking of hbest1 with the compounds. The efficacy of anthocyanins demonstrates that natural compounds may serve as ideal therapies for diseases affecting the retinal pigment epithelium.

Acknowledgement:
The authors thank Bioinformatics Infrastructure Facility (DBT-BIF), University of Madras for providing computational facilities.

References:
[1] Arinaminpathy Y et al. Drug Discov Today. 2009 14: 1130 [PMID: 19733256]
[2] Krämer F et al. Eur J Hum Genet. 2000 8: 286 [PMID: 10854112]
[3] Marquardt A et al. Hum Mol Genet. 1998 7: 1517 [PMID: 9700209]
[4] Petrukhin K et al. Nat Genet. 1998 19: 241 [PMID: 9662395]
[5] White K et al. Hum Mutat. 2000 15: 301 [PMID: 10737974]
[6] Tsunenari T et al. J Biol Chem. 2003 278: 41114 [PMID: 12907679]
[7] Sun H et al. Proc Natl Acad Sci. 2002 9: 4008
[8] Kranjc A et al. PLoS One. 2009 4: e4672 [PMID: 19262692]
[9] Baker D & Sali A. Science. 2001 294: 93 [PMID: 11588250]
[10] Skolnick J et al. Nat Biotechnol. 2000 18: 283 [PMID: 10700142]
[11] Chivian D et al. Proteins. 2003 53: 524 [PMID: 14579342]
[12] Matsumoto H et al. J Agric Food Chem. 2001 49: 1546 [PMID: 11312894]
[13] Kong JM et al. Phytochemistry. 2003 64: 923 [PMID: 14561507]
[14] Veigas JM et al. Food Chemistry. 2007 105: 619
[15] Aqil F et al. Nutr Cancer. 2012 64: 428 [PMID: 22420901]
[16] Jorge Arreola & Patricia Pérez-Cornejo, Adv Mol Cell Biol. 2006 38: 181
[17] Bairoch A & Apweiler R, Nucleic Acids Res. 2000 28: 45 [PMID: 10592178]
[18] Bryson K et al. Nucleic Acids Res. 2005 33: W36 [PMID: 15980489]
[19] Jones DT, J Mol Biol. 1999 292: 195 [PMID: 10493868]
[20] Shen H & Chou J, PLoS One 2008 3: e2399 [PMID: 18545655]
[21] Wallner B & Elofsson A, Protein Sci. 2003 12: 1073 [PMID: 12717029]
[22] Laskowski RA et al. J Appl Cryst. 1993 26: 283
[23] Hartzell HC et al. Physiol Rev. 2008 88: 639 [PMID: 18391176]
[24] Qu Z & Hartzell HC, J Biol Chem. 2001 276: 18423 [PMID: 11279188]
[25] Pintea A et al. J Ocul Pharmacol Ther. 2011 27: 315 [PMID: 21663493]
[26] http://pubchem.ncbi.nlm.nih.gov/
[27] Sherman W et al. J Med Chem. 2006 49: 534 [PMID: 16420400]
[28] Friesner RA et al. J Med Chem. 2004 47: 1739 [PMID: 15027865]
[29] Wallace AC et al. Protein Eng. 1995 8: 127 [PMID: 7630882]
[30] Nakajima JJ et al. J Biomed Biotechnol. 2004 5: 241 [PMID: 15577184]
[31] de Brito ES et al. J Agric Food Chem. 2007 55: 9389 [PMID: 17929888]
[32] Li L et al. J Agric Food Chem. 2009 57: 826 [PMID: 19166352]
[33] Carter PH & Ernst E, Surv Ophthalmol. 2004 49: 38 [PMID: 14711439]
[34] Milbury PE et al. Invest Ophthalmol Vis Sci. 2007 48: 2343 [PMID: 17460300]
[35] Kong JM et al. Phytochemistry. 2003 64: 923 [PMID: 14561507]
[36] Dai J & Mumper RJ, Molecules. 2010 15: 7313 [PMID: 20966876]
[37] Bagchi D et al. Mol Cell Biochem. 2006 281: 197 [PMID: 16328973]

License statement: This is an open-access article, which permits unrestricted use, distribution, and reproduction in any medium, for non-commercial purposes, provided the original author and source are credited

Edited by P Kanguane
Citation: Priya et al. Bioinformation 8(16): 742-748 (2012)
Table 1: Induced-fit docking results of Inhibitors showing glide score, glide energy, glide emodel score and Hydrogen bonding interactions with hbest1

| S.No | Compound                  | Glide Score | Glide Energy | Glide emodel | Interactions                               | Distance (Å) |
|------|---------------------------|-------------|--------------|--------------|--------------------------------------------|--------------|
| 1    | Cyanidin 3,5 diglucoside | -11.80      | -64.34       | -71.60       | O-H…O(Leu82)                               | 2.84         |
|      |                           |             |              |              | O-H…O(gln59)                               | 2.87         |
|      |                           |             |              |              | (Arg105)N-H…O                              | 3.18         |
|      |                           |             |              |              | (Arg47)N-H…O                               | 3.44         |
|      |                           |             |              |              | (Arg47)N-H…O                               | 3.08         |
|      |                           |             |              |              | (Arg47)N-H…O                               | 2.85         |
| 2    | Malvidin 3, 5 diglucoside | -5.15       | -71.61       | -92.42       | O-H…O(Arg92)                               | 2.90         |
|      |                           |             |              |              | (Arg92)N-H…O                               | 2.94         |
|      |                           |             |              |              | (Arg461)N-H…O                              | 3.11         |
|      |                           |             |              |              | (Arg534)N-H…O                              | 3.02         |
|      |                           |             |              |              | O-H…O(Thr 87)                              | 3.17         |
| 3    | Petunidin 3,7 diglucoside | -7.98       | -68.39       | -94.47       | O-H…O(Asn136)                               | 2.92         |
|      |                           |             |              |              | O-H…O(Thr87)                               | 2.67         |
|      |                           |             |              |              | O-H…O(Thr 87)                              | 2.94         |
|      |                           |             |              |              | (Arg461)N-H…O                              | 2.93         |
|      |                           |             |              |              | (Arg92)N-H…O                               | 3.27         |
|      |                           |             |              |              | (Arg105)N-H…O                              | 2.77         |
|      |                           |             |              |              | (Arg105)N-H…O                              | 3.04         |
| 4    | Delphinidin 3,5 di glucoside | -9.18      | -77.17       | -107.57      | O-H…O(gly159)                              | 2.88         |
|      |                           |             |              |              | O-H…O(Glu 382)                              | 2.57         |
|      |                           |             |              |              | O-H…O(Ile 49)                               | 2.74         |
|      |                           |             |              |              | O-H…O(Asp 461)                              | 2.79         |
|      |                           |             |              |              | O-H…O(Gln 96)                               | 2.49         |
| 5    | Peonidin 3, 5 diglucoside | -8.47       | -58.82       | -85.17       | O-H…O(Asn 133)                              | 3.14         |
|      |                           |             |              |              | O-H…O(Arg 534)                              | 2.91         |
|      |                           |             |              |              | O-H…O(Gln 460)                              | 2.97         |
|      |                           |             |              |              | O-H…O(Gln 460)                              | 2.69         |
| 6    | Resveratrol               | -8.25       | -30.89       | -48.71       | O-H…O(Tyr 131)                              | 2.93         |
| 7    | Niflumic acid             | -8.80       | -38.19       | -49.34       | O-H…O(Ala132)                               | 2.76         |
|      |                           |             |              |              | (Arg92)N-H…O                               | 2.95         |
|      |                           |             |              |              | N-H…O (Tyr131)                              | 3.10         |