Academic Sciences International Journal of Pharmacy and Pharmaceutical Sciences

ISSN- 0975-1491 Vol 4, Suppl 1, 2012

Research Article

3-D STRUCTURE PREDICTION OF AQUAPORIN-2, VIRTUAL SCREENING AND IN-SILICO DOCKING STUDIES OF GOLD AND SILVER DERIVATIVES, USED AS POTENT INHIBITORS

POMPI SHARMA[1*], ANUPAM BHATTACHARYA[2]

[1]CSIR- North East Institute of Science and Technology, Assam, [2] Birla Institute of Technology, Mesra, Ranchi, India.
Email:pompi.sharma86@gmail.com

Received: 5 Nov 2011, Revised and Accepted: 1 Jan 2012

ABSTRACT

Aquaporin-2 is a channel protein, encoded by AQP2 gene. Mutations in aquaporin-2 can lead to nephrogenic diabetes insipidus, congestive heart failure, liver cirrhosis and preeclampsia. The structure of the protein aquaporin-2 is still unknown. In this study, we modelled the 3-D structure of aquaporin-2 with Modeller 9V9 by using the crystal structure of Aquaporin-1 as the template. The best model was selected by using the lowest DOPE score. Its 3-D structure was evaluated and validated by using PROCHECK comprising 89.2% amino-acids in favouredregion, 0.00% amino-acids in disallowed region of Ramachandran plot. G-factor of the model is -0.06. The overall quality factor of the model (85.656%) was analysed by ERRAT. The energy estimation of the model was done by using ANOLEA and DDFIRE. Gold, silver and mercury derivatives were screened on the basis of the small molecule rules and bioactivity analysis. The Energy minimization of the modelled structure was performed by using Chembio Office Ultra. The gold derivative (CID2012990) showed the maximum binding affinity with active site, lowest MolDock score and highest bioactivity score against AQP2. The ligand or its derivatives might be used as effective drug against nephrogenic diabetes.

Keyword: AQP-2, Homology Modelling, Docking, Gold, Silver derivatives.

INTRODUCTION

The AQP2 or Aquaporin -2 is a water channel protein found in the apical cell membranes of the principal cells in the kidney's collecting duct and in intracellular vesicles of the cytoplasmic cell. The AQP2 is also commonly named as ADH water channel or collecting ducts water channel protein or water channel aquaporin-2 or more. The antidiuretic hormone Vasopressin secreted by the pituitary gland regulate the expression of this aquaporin 2 protein. Upon vasopressin stimulation, AQP2 translocate from sub apical storage vesicles to the apical plasma membrane, rendering the cell water permeable, which in turn causes water reabsorption[1]. The vasopressin binds to the cell surface vasopressin receptor which activates a signalling pathway that causes the aquaporin 2 containing vesicles to fuse with the plasma membrane so that aquaporin 2 can be used by the cell [2]. Water flows through the membranes of all living cells by two distinct mechanisms. Diffusion of water through pure lipid bilayers occurs with high activation energy and through water pores. Multiple observations provided clues to the functioning of water channels in a variety of specialized membranes, and it has long been recognized that the permeation of this pore is remarkably specific, since other small molecules, ions or even protons are not accommodated [3]. The protein aquaporin is encoded by the gene AQP2 present on the Human chromosome 12. AQP-2 is a very hydrophobic membrane-integral protein of a molecular mass of 29 kDa. It is a member of the MIP protein family and is homologous to aquaporin-1. Despite the accumulated knowledge of AQP-2 in functional importance of body water homeostasis, the structural basis of AQP-2 is not well known. The molecular structure of AQP-1, the first identified water channel, has been studied and partially resolved. It was shown that AQP-1 exists in plasma membrane with tetramer formation. Regarding to the structure of the aqueous pore in functionally active AQP-1 monomer, three structural models have been proposed: the hourglass model, the a-helical model, and the b-barrel model. But the validation of these models has not been sufficiently examined. Moreover, there have been no investigations regarding the molecular structures of aquaporin other than AQP-1[4].

METHODS AND MATERIALS

Homology modelling:

The Homology modelling or comparative modelling of protein refers to constructing 3D structural model of a "target" protein from its amino acid sequence and an experimental three-dimensional structure of a related homologous protein. Homology modelling can produce high-quality structural models when the target and template are closely related, which has inspired the formation of structural genomics consortium dedicated to the production of representative experimental structures for all classes of protein folds.

Template Sequence Search

The template selection is the process of finding out a sequence that similar to the target protein sequence and which have a 3D structure available. The method is based on the PDB-BLAST where searching of template done by multiple sequence alignment for those sequences that have structures in protein data bank.

3D Structure Predictions

The homology modelling software MODELER 9v9[5]was used to predict a 3D model of the target protein Aquaporin2. The homology modelling process completed with four sequential steps: template selection, target-template alignment, model construction and model refinement and assessment.

Model refinement

Model refinement was done by minimising the energy of the constructed structure using Chem Bio3D Ultra [6].

Model Assessment

It is a process where the generated model is tested with different methods to find out the errors in the model, disorder regions, and quality of the generated model. The process also deals with the model verification, calculation of energies, calculation of main bond length, main bond angles etc. Different soft-wares as well as tools used for these calculations are as follows.

Procheck server

The PROCHECK server is an analysis tool that provides an idea of the stereo chemical quality of all protein chains of a PDB structure. It highlights the regions of the protein which appear to have unusual geometry and provide an overall assessment of the structure as a whole. The Procheck computes the G-factors which provide a measure of how normal or how "unusual" a stereo chemical property is. It also calculates other properties such as torsional angles and covalent geometry[7].

Errat

Errat is a protein structure verification algorithm that is used for evaluating the progress of crystallographic model building and

refinement. This program works by analysing the statistics of non-bonded interactions between different atom types. A single output plot is produced which gives the value of the error function vs. position of a 9-residue sliding window[8].

QMEAN

The QMEAN server is a well suited tool for model quality estimation which estimation is an essential component of protein structure prediction. Usually, in the course of protein structure prediction a set of alternative models is produced, from which subsequently the most accurate model has to be selected. QMEAN is a composite scoring function which is able to derive both global and local error estimates on the basis of one single model[9].

ANOLEA

ANOLEA (Atomic Non-Local Environment Assessment) is a server that performs energy calculations on a protein chain, evaluating the "Non- Local Environment" (NLE) of each heavy atom in the molecule. It is a server to assess the quality of a three - dimensional protein structure. It uses a statistical potential at the atomic level and gives an energy profile as output. The energy of each pair wise interaction in this non-local environment is taken from a distance-dependent knowledge-based mean force potential that has been derived from a database of 147 non-redundant protein chains with a sequence identity below 25% and solved by X-Ray crystallography with a resolution lower than 3 Å[10, 11].

PROSA analysis

The recognition of errors in experimental and theoretical models of protein structures is a major problem in structural biology. The ProSA program (Protein Structure Analysis) is an established tool which has a large user base and is frequently employed in the refinement and validation of experimental protein structures and in structure prediction and modelling. ProSA calculates an overall quality score for a specific input protein structure. The service specifically addresses the needs encountered in the validation of protein structures obtained from X-ray analysis, NMR spectroscopy and theoretical calculations[12, 13].

VERIFY-3D

The Verify-3D is a structure evaluation server that analyzes the compatibility of an atomic model (3D) with its own amino acid sequence. Each residue is assigned a structural class based on its location and environment (alpha, beta, loop, polar, nonpolar, etc.). A collection of good structures is used as a reference to obtain a score for each of the 20 amino acids in this structural class. The scores of a sliding 21-residue window (from -10 to +10) are added and plotted for individual residues[14].

DDFIRE

The dipolar "DFIRE (dDFIRE) energy function is based on the orientation angles involved in dipole–dipole interactions. This is done by treating each polar atom as a dipole. The orientation of the dipole is defined by the bond vectors that connect the polar atom with other heavy atoms. The dDFIRE energy function is extracted from protein structures based on the distance between two atoms and the three angles involved in dipole–dipole interactions. Each polar atom possesses a reference direction that mimics the orientation of a dipole[15].

Submission of the model

The designed model was submitted to Protein Model Database and an ID was obtained as PM0077394 [16].

RESULTS AND DISCUSSION

Template Selection

The homologues obtained from MODELLER 9v9 were: 1ldfA, 1j4nA, 1rc2B and1tm8A.Among all of them, 1J4N, Chain A, Crystal Structure of the Aqp1 Water Channel, was selected as thetemplate becauseit had the lowest E-value and a good identity score to be a template.

Model Generation

Taking 1j4nA as the template, an alignment was done between both the sequences. A model was generated using MODELLER 9v9. It generated five models each with a DOPE score. The model, AQP2.B99990004, with the lowest DOPE score of -29864.62109 (Table I) was selected as the best model (Fig I).

Table I: Table showing the DOPE scores of the five models constructed by Modeller9v9.

File name	molpdb	DOPE score	GA341
AQP2.B99990001.pdb	1396.29431	-29467.85742	1.00000
AQP2.B99990002.pdb	1240.36133	-29769.29102	1.00000
AQP2.B99990003.pdb	1481.04700	-29742.99219	1.00000
AQP2.B99990004.pdb	1253.44727	-29864.62109	1.00000
AQP2.B99990005.pdb	1315.18054	-29767.68164	1.00000

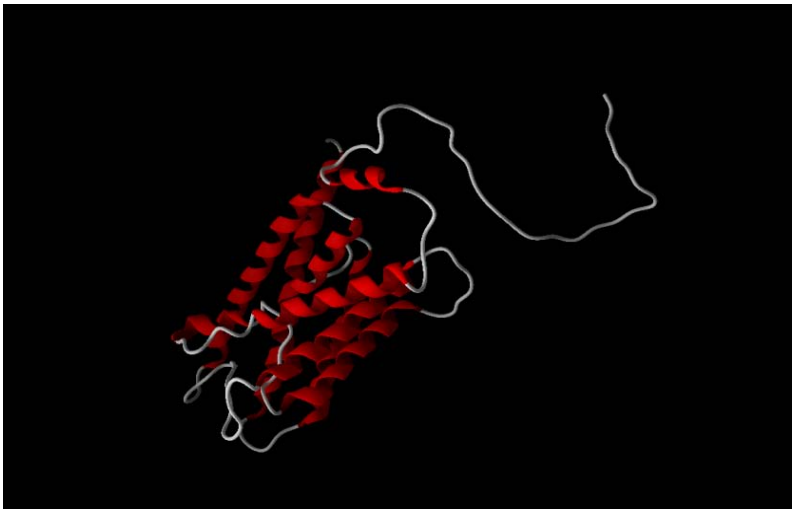

Fig. I: Secondary structure of model AQP2.B9999004 with Molegro Viewer

Model Assessment

The ERRAT results provided the overall quality factor of the generated model to be 85.656 % (Fig. II). ANOLEA results showed that thetotal non-local energy of the protein in (E/kT units) is -2. The results from DDFIRE/DFIRE2 calculated the total energy to be-583.64 which is equal to the summation of DFIRE2 energy and other three angle related energy. VARIFY 3D estimated that the atom no. 57 had the highest compatibility of 0.68 with 3D-1D profile score.

The Ramachandran Plot obtained from the PROCHECK results gave the stereo chemical property of the model (Table II), (Fig.III).

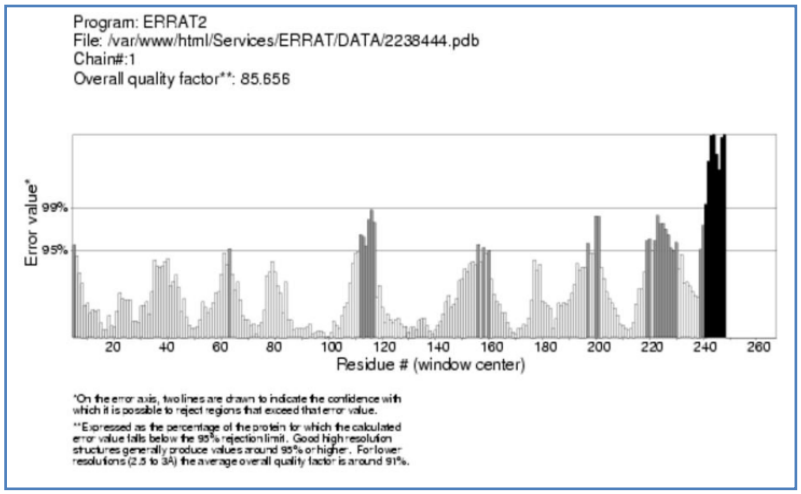

Fig.II: The figure shows the quality of the modelled protein: 85.656%.

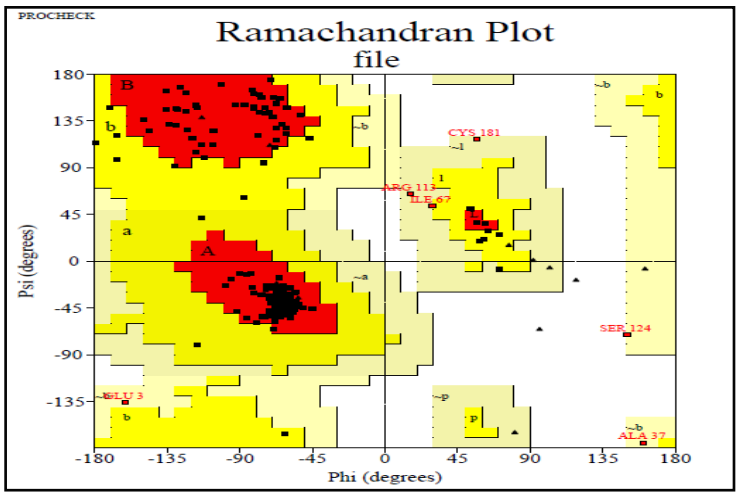

Fig. III: The Ramachandran plot of modelled protein.

Table II: Ramachandran Plot statistics

		No. of Residues	**%-tage**
Most favoured regions	[A, B, L]	206	89.2%*
Additional allowed regions	[a,b,l,p]	19	8.2%
Generously allowed regions	[~a,~b,~l,~p]	6	2.6%
Disallowed regions	[XX]	0	0.0%
Non-glycine and non-proline residues		231	100.0%
End-residues (excl. Gly and Pro)		2	
Glycine residues		24	
Proline residues		14	
Total number of residues		271	
G-Factors			
Parameters	**Score**	**Average Score**	
Dihedral angles:-	0.11		
Phi-Psi distribution			
Chi1-chi2 distribution	-0.22		
Chi1 only	0.18		
Chi3 & chi4	0.48		
Omega	-0.19		
		-0.01	
Main Chain Covalent Forces:			
Main-chain bond lengths	-0.11		
Main-chain bond angles	-0.14		
		-0.13	
Overall average		-0.06	

Selection of compounds

Research reveals that silver and gold compounds are tested as potential inhibitor of aquaporins of plant and human origin. Silver and gold are most potent inhibitors of aquaporins than the widely used mercury containing compounds. The mechanism of silver and gold inhibition is most likely due to their ability to interact with the sulfhydral group of proteins [17, 18].

In this study, we screened silver and gold compounds and their derivatives through virtual screening and selected the compounds for docking and further analysis. The compounds were tested for bioactivity analysis to further identify the inhibitory effect against the modelled AQP-2. Selected mercury compounds did not satisfy the specific criteria, so were not included in the study. 60 gold and 50 silver compounds were selected and bioactivity analysis was performed. Among them, 4 gold derivatives showed highest bioactivity against the channel protein (Table3).

Docking studies were performed with these four gold compounds and the target protein AQP2.

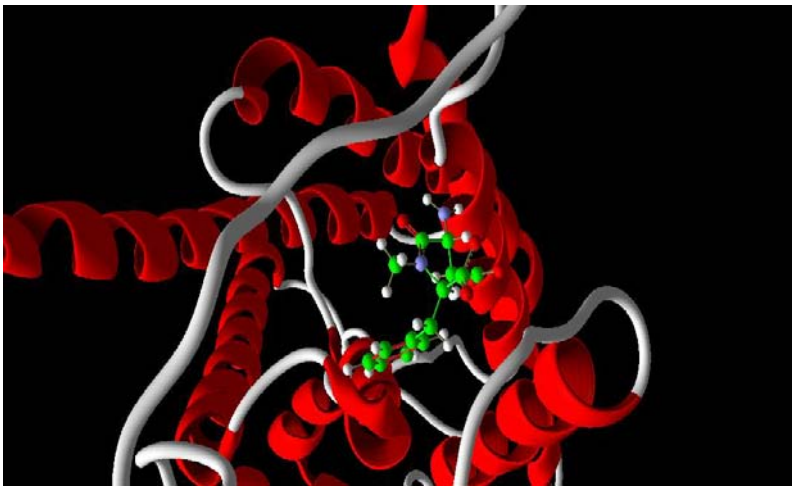

Fig. IV: Docking of the ligand CID 20129902 with the active site of the receptor

Fig. V: Energy minimization of different poses of ligand during docking

Table IV:Best five poses of ligand of which first pose showing least Moldock score

Poses	MolDock score	Rerank score	Interation energy	Internal energy	Hbond energy	Ligand efficiency
1	-102.446	-79.0586	-112.09	9.6437	-6.76682	-4.87839
2	-102.087	-75.0137	-105.689	3.6012	-3.46334	-4.86131
3	-95.9789	-69.0187	-100.36	4.38153	-5.76643	-4.57043
4	-92.0309	-68.9466	-99.9243	7.89332	-3.38312	-4.38243
5	-81.7265	-67.4155	-90.154	8.42746	-1.54975	-3.89174

Docking

Increasing cost of drug development and reduced number of new chemical entities have been a growing concern for new drug development in recent years. In-silico drug design can play a significant role in all stages of drug development from the preclinical discovery stage to late stage clinical development. It helps in selecting a potent lead molecule by virtual screening, thereby reducing the time and cost.[19] Drug design soft-wares have potential role to design novel proteins and drugs in pharma or biotechnology field.[20].

These softwares used to analyse protein-ligand interaction, docking efficiency, ligand efficiency, virtually screened molecules, docking energy with different scoring functions.

In this study, docking was performed for the AQP2 model with the four ligands having the highest bioactivity using Molegro Virtual Docker. The energy of the ligands were initially minimised with Chem Bio Office and then docked with the AQP2. The ligand CID 20129902 showed the lowest Moldock score of -102.445 (Table III) (Fig IV, V) (Table IV). Thus, indicating that it has the highest affinity with the target protein. Also it was found that among the four compounds, this compound had the highest bioactivity with the channel protein. (Table V) (Fig VI).

Table V: Gold derivatives showing the bioactivity and the MolDock score of the gold derivatives

Gold derivatives	Bioactivity score	Moldock score
CID 84284	0.13	875.706
CID 20129902	0.25	-102.445
CID 3083412	0.06	-83.1972
CID 3083412	0.07	906.099

Table VI:Properties of the ligand CID 20129902

Molecular weight (g/ml)	**294.30312 [g/mol]**
Molecular Formula	$C_{14}H_{18}N_2O_5$
Logp	-2.376
H-bond Donor	3
H-bond Acceptor	6
Rotatable bonds	7
Heavy Atoms	21
Canonical SMILE	CN(C(CC1=CC=CC=C1)C(=O)O)C(=O)C(CC(=O)O)N
Polar surface Area	120.93
Volume	264.922

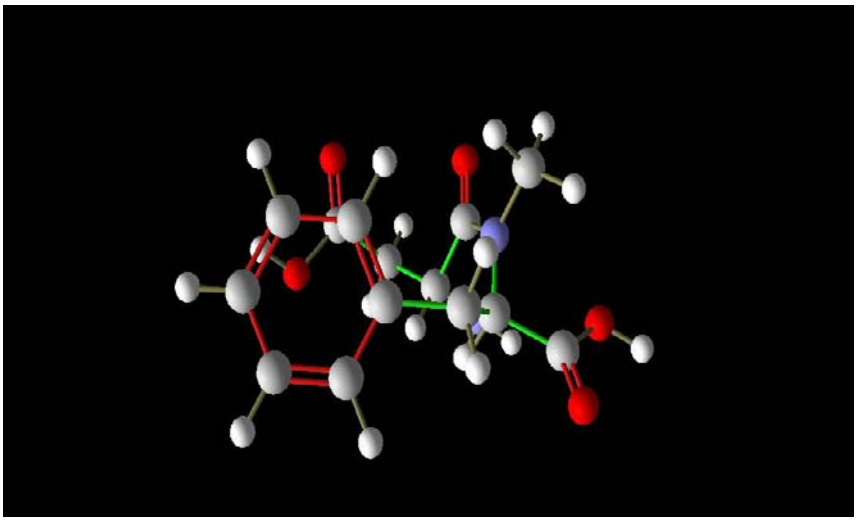

Fig VI: Three dimensional structure of ligand as seen in Molegro Viewer

CONCLUSION

Body water homeostasis is essential for survival of mammals. Water transport occurs through a specialized channel called aquaporin (AQP). AQPs play an important role in reabsorption of water and in concentration of urine in the kidney. Out of 13 aquaporin isoforms, AQP2 is the predominant vasopressin-regulated water channel. The vasopressin tightly regulates body water balance. Upon vasopressin stimulation, AQP2 translocate from subapical storage vesicles to the apical plasma membrane, rendering the cell water permeable, which in turn causes water reabsorption leading to urine concentration. AQP2 mutations cause congenital nephrogenic diabetes insipidus (NDI), a disease characterized by a massive loss of water through the kidney. Structure of AQP2 is not known. In this present work we have modelled a 3D structure of the Aquaporin-2 protein by homology modelling.

Further energy was also calculated and the model validation was done using different online tools which revealed that the overall quality factor of the modelled protein is 85.656%. Further Ramachandran Plot analysis indicated that none of the residues were present in the disallowed region and maximum of the residues, i.e., 89.2% of the residues were in the most favoured region. The modelled structure was submitted to Protein Model Database. The energy of the protein was further minimised with ChemBio Office Ultra. Silver and gold derivatives which act as potent inhibitors of AQP2 were screened for docking with the protein. 50 compounds were screened on the basis of its log P value and bioactivity analysis from PubChem database. And only 4 such derivatives were obtained for gold and none for silver. The four gold derivatives were docked with the modelled structure of AQP2 and it was found that the ligand CID20129902 had the lowest docking score of -102.445, thus showing highest binding affinity with the protein. Further experimental verifications would be needed for the modelled 3-D structure of AQP2 protein. Various other derivatives can be used for docking study.

REFERENCES

1. "Molecular Mechanisms and Drug Development in Aquaporin Water Channel Diseases: Molecular Mechanism of Water Channel Aquaporin-2 Trafficking" Yumi Noda1, and Sei Sasaki; J PharmacolSci 96, 249 – 254 (2004).

2. Lodish, Harvey F. Molecular Cell Biology. New York: W.H. Freeman, 2008. Print. 445.
3. A.Engel1, Y.Fujiyoshi2 and P.Agre3" The importance of aquaporin water channel protein Structures" The EMBO Journal Vol.19 No.5 pp.800–806, 2000.
4. "Structure of Aquaporin-2 Vasopressin Water Channel LiqunBai, Kiyohide Fushimi, Sei Sasaki, and Fumiaki Marumo "THE JOURNAL OF BIOLOGICAL CHEMISTRY Vol. 271, No. 9, Issue of March 1, pp. 5171–5176, 1996.
5. "EasyModeller: A graphical interface to MODELLER"Bhusan K Kuntal**,** PolamarasettyAparoyandPalluReddanna**,** BMC Research Notes 2010, 3:226doi:10.1186/1756-0500-3-226.
6. ChemBioOffice Ultra 2010-A Great Benefit for Academia/CambridgeSoft Solution featuring E-Notebook (www.cambridgesoft.com).
7. "PROCHECK: a program to check the stereochemical quality of protein structures" R. A. Laskowski, M. W. Macarthur, D. S. Moss, J. M. Thornton. J. Appl. Cryst., Vol. 26 (1993), pp. 283-291. Bibtex-import1.
8. http://nihserver.mbi.ucla.edu/ERRAT/.
9. Benkert P, Künzli M, Schwede T. "QMEAN server for protein model quality estimation." Nucleic Acids Res. 2009 Jul 1; 37(Web Server issue):W510-4. Epub 2009 May 8.
10. ANOLEA: a www server to assess protein structures. (PMID:9322034)Melo F, Devos D, Depiereux E, FeytmansEDepartment of Biology, Proceedings. International Conference on Intelligent Systems for Molecular Biology; ISMB. International Conference on Intelligent Systems for Molecular Biology [1997, 5:187-90].
11. http://protein.bio.puc.cl/cardex/servers/.
12. "ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Markus Wiederstein", Manfred J. SipplNucleic Acids Research, Vol. 35, No. suppl 2. (1 July 2007), pp. W407-W410. doi:10.1093/nar/gkm290 Key: citeulike:1443791
13. Markus Wiedersteinand Manfred J. Sippl "ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins" Oxford Journals Life Sciences Nucleic Acids Research Volume35, Issuesuppl2 Pp. W407-W410.
14. Roland Luthy, James U. Bowie, and David Eisenberg, "Assessment of protein models with three-dimensional profiles" Nature, 356, 83-85 (1992).
15. Yuedong Yang and Yaoqi Zhou "Specific interactions for ab initio folding of protein terminal regions with secondary structures"
16. The PMDB Protein Model Database T. Castrignano, P.D. De Meo, D. Cozzetto, I.G. Talamo, A. Tramontano, 2006, Nucleic Acids Res., Vol 34(1), ppD306 - D309.
17. New potent inhibitors of aquaporins: silver and gold compounds inhibitaquaporins of plant and human origin, FEBS Letters, Volume 531, Issue 3, Pages 443-447, Christa M Niemietz, and Stephen D Tyerman.
18. Molecular Mechanisms and Drug Development in Aquaporin Water Channel Diseases: Molecular Mechanism of Water Channel Aquaporin-2 Trafficking. Journal of pharmacological Sciences, 96,249-254(2004). Yumi Noda, and Sei Sasaki.
19. In silico drug design-tool for overcoming the innovation deficit in the drug discovery process.E.n bharath*, s.n manjula, a. Vijay chand,International Journal of Pharmacy and Pharmaceutical Sciences,Vol 3, Issue 2, 2011.
20. Advanced drug designing softwares and their applications in m edical research .B dineshkumar1, p vignesh kumar2, sp bhuvan eshwaran2, analava mitra1. International Journal of Pharmacy and Pharmaceutical Sciences,Vol 2, Issue 3, 2010.