Available online at www.scholarsresearchlibrary.com

Scholars Research Library

Scholars Research Library

Annals of Biological Research, 2011, 2 (1) :40-50
(http://scholarsresearchlibrary.com/archive.html)

Annals of Biological Research
Scholars Research Library

ISSN 0976-1233
CODEN (USA): ABRNBW

Structural Model Based Designing of Inhibitors for Glial Fibrillary Acidic Protein

Sagarika Biswas[1]*, Sandeep K. Kushwaha[2], Rahul S. Mandal[1], Saugata Roy[1] and H. R. Das[1]

[1]Institute of Genomics and Integrative Biology, Delhi University Campus, Mall Road, Delhi, India
[2]Department of Biotechnology and Bioinformatics, JUIT, Solan, H.P., India

---

ABSTRACT

Glial fibrillary acidic protein (GFAP) is an intermediate filament Type III protein and is containing three domains. The most conserved domain of GFAP is the rod domain. The present study has been made for in silico prediction to determine the three-dimensional structure of GFAP protein. It has been carried out through molecular modeling using MODELLER 9v5. Its active site residues has been predicted through comparative results of MODELLER 9v5. The ligands were designed using LigandScout 2.0. The designed ligand and receptor interaction studies were carried out through pharmacophore analysis followed by interaction studies using AUTODOCK4. Virtual screening of ligands has been performed by Molegro Virtual Docker. Analogues of ligands were generated through Chemsketch10.0 and ARG (258) was identified as a catalytic residue.

Keywords: GFAP, Protein Modelling, Ligand Designing, Docking, Analogues.

INTRODUCTION

Glial fibrillary acidic protein (GFAP) is an intermediate filament (IF) type III protein and is well known for its biological processes such as cell structure and movement, cell communication and the functioning of the blood brain barrier [1]. It is a major intermediate filament protein of adult brain and is a characteristic of mature astrocytes [2] in central nervous system (CNS). Type III intermediate filaments contain three domains and the most conserved one is the rod domain. The

specific DNA for rod domain differs from the gene sequence of other filament proteins of type III class. The rod domain coils around that of another filament to form a dimmer with the N-terminal and C-terminal of each filament aligned. The DNA sequence in this region may differ with other intermediate filament gene that indicates the high conservation of structural elements of the region [3].

GFAP is closely related to other class III IF proteins like vimentin, desmin, peripherin that are involved in maintaining structure and function of cell cytoskeleton, cell communication and the functioning of the blood brain barrier [4]. The amount of GFAP produced by the cells was found to be regulated by cytokines and hormones. The increased expressions of this protein are commonly referred to as "astrocytic activation". In mature cells, mostly phosphorylation of GFAP has been studied extensively [5]. But the functional importance of alteration in the levels of GFAP is not fully understood. There are multiple disorders associated with improper GFAP regulation. Glial scarring is a consequence of several neurodegenerative conditions as well as injury that severs neural material. The scar was found to be formed by astrocytes interaction with fibrous tissue to re-establish the glia margins around the central tissue core and caused by up-regulation of GFAP [6, 7]. It was also observed that Alexander's disease is directly related to GFAP. The relationship between GFAP and Alexander disease is not completely understood but mutations were observed in the coding region of the GFAP gene [8]. In the present study, GFAP protein has been modeled in order to understand and develop the functional property of GFAP. Computational combinatorial chemical techniques were used for chemical compound library generation and computational approaches like docking and screening was used for identification of inhibitors and ligands designing. We believe that the designed ligand and its analogues can play an important role in GFAP associated diseases and disorders.

MATERIALS AND METHODS

2.1. Template identification and sequence analysis of GFAP proteins

GFAP sequence (NCBI-GI: 251802,) was obtained from NCBI (www.ncbi.nlm.nih.gov). The program BLAST-P [9] has been used to detect similar protein sequences to GFAP. A set of templates (1GK7, 3B9A, 3KLT, 1GK4) was obtained from Protein Data Bank (PDB) that was showing moderate identity with GFAP. Multiple sequence alignment of template and target has been performed using ClustalW for GFAP modeling [10].

2.2. Modeling of GFAP and quality analysis studies

The 3D structure of GFAP has been modeled on the basis of multiple-templates of high-resolution crystallographic structures. Homology modeling was performed using MODELLER9v5 [11]. This program models protein tertiary structure by satisfaction of spatial restraint using standard parameters sets. The generated three dimensional model includes all non-hydrogen main-chain and side-chain atoms. Generated model has been refined using energy minimization techniques to optimize stereochemistry and to remove bumps and steric clashes among non-bonded interactions using the commands of MODELLER9v5 [12]. Parameters like covalent bond distances and angles, stereo chemical validation, atom nomenclature were validated by Ramachandran plot using PROCHECK [13] and WHAT-IF [14]. The overall

quality factor of non-bonded interactions between different atoms type were measured using ERRAT program.

2.3. Structural characterization and ligand designing

GFAP was structurally characterized through online tool and offline software and active site prediction has been carried out on the basis of comparative analysis of results [15]. Ligands of GFAP were designed using Ligand Scout 2.0. Ligand Scout software automatically calculates a potential pharmacophore by considering the distances and the angles between the corresponding chemical functional groups of the ligands and the target-proteins that were used for the ligand generations [16].

2.4. Docking studies and virtual screening

Molecular-docking-based virtual screening was found to be an important tool in drug discovery that has been used significantly to reduce the number of possible chemical compounds to be investigated [17]. Screening of best compatible ligands to target were obtained through docking. Interaction studies of designed ligand was carried out using AutoDock4.0. Ligand analogues were generated using ChemSketch10.0 and Molegro has been used for virtual screening.

RESULTS AND DISCUSSION

GFAP is mainly related to brain disorders. Most of the previous studies showed that serum GFAP is highly and significantly associated with the volume of brain lesion [18,19]. Most neurobiologists believe that the levels of GFAP and the state of assembly into filaments are important in modulating astrocyte motility and shape, especially through extensions of astrocytic processes [20]. However regulation of GFAP in brain disorder patients as well as patients with bone fractures but no brain injury is not completely understood. We believe that modeling of GFAP may lead us to the identification of active site of GFAP that may help in understanding the pathogenesis of GFAP associated disease.

Modeling of GFAP was tedious task due to large uncovered region and unavailability of suitable template (Fig.-1). Since the PDB structure of GFAP was not available, modeling of GFAP was carried out using four reference templates (1GK7, 3B9A, 3KLT, 1GK4). Only ungapped portions of templates have been considered from identified templates for 3D modeling of GFAP.

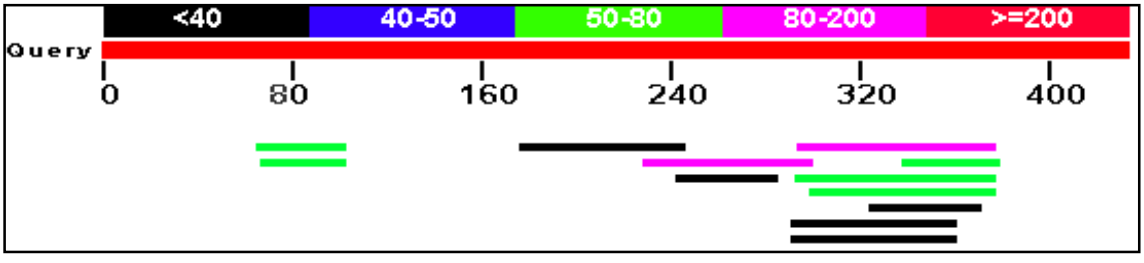


Figure-1 Template search for query sequence of GFAP through BLAST

We used MODELLER9v5 program that uses the spatial constraints, determined from the crystal structure of a template protein, to generate 3D model of the target protein. Initially, the number

of incompetent model has been generated with uncovered part of the query sequence. Non template supportive portion of query sequence has been covered by fold recognition prediction using LOMETS (local meta-threading server for protein structure) and the generated model has been shown in Fig.-2.

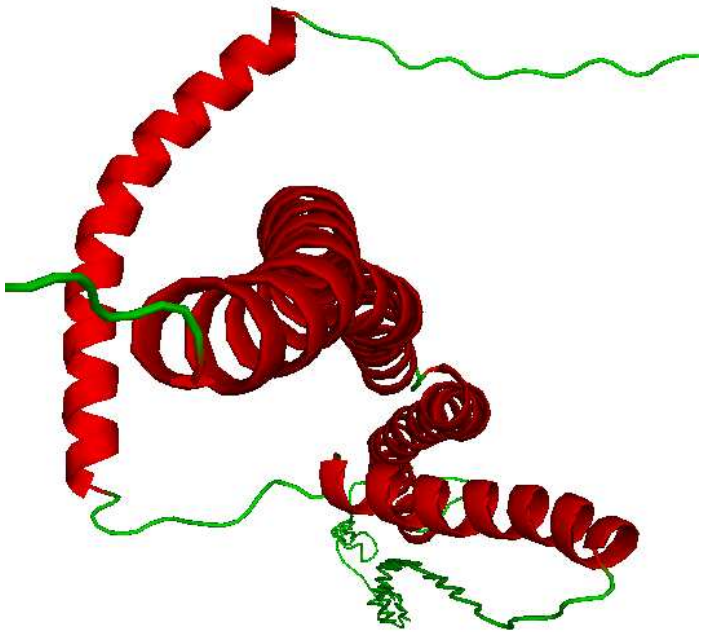


Figure-2 Three dimensional visualization of the modelled structure of GFAP

The homology model of GFAP satisfies sterochemical restraints that were carried out by PROCHECK and WHAT-IF. The generated 3D model of target proteins has been analyzed and validated by Ramachandran plot (Fig.-3) through PROCHECK program of the SAVS metaserver (http://nihserver.mbi.ucla.edu/SAVS/). In modelled structure 96.1% residues were in most favoured region and no residues were lying in the disallowed regions of Ramachandran plot. The overall quality factor for modelled structure was reported to be 76% through ERRAT program of SAVS. From WHAT-IF analysis, the Z-score and RMS Z-score of average packing quality of generated model was found to be -2.921 and 1.194 respectively.

Structural characterization has been found to be very much essential for drug designing. Thus we performed BLAST of our query sequence. BLAST search result for query sequence (Fig.-4) was showing domain filament and HAD_like superfamily on the basis sequence conservation. The query sequence of protein had filament head (33-66), intermediate filament protein (68-376) and Haloacid dehalogenase like Hydrolase (190-285). Filament head represents the N-terminal head region of intermediate filaments which binds to DNA. Phosphorylation of the head region can affect filament stability [21]. The head has been shown to interact with the rod domain of the same protein (Fig.-4).

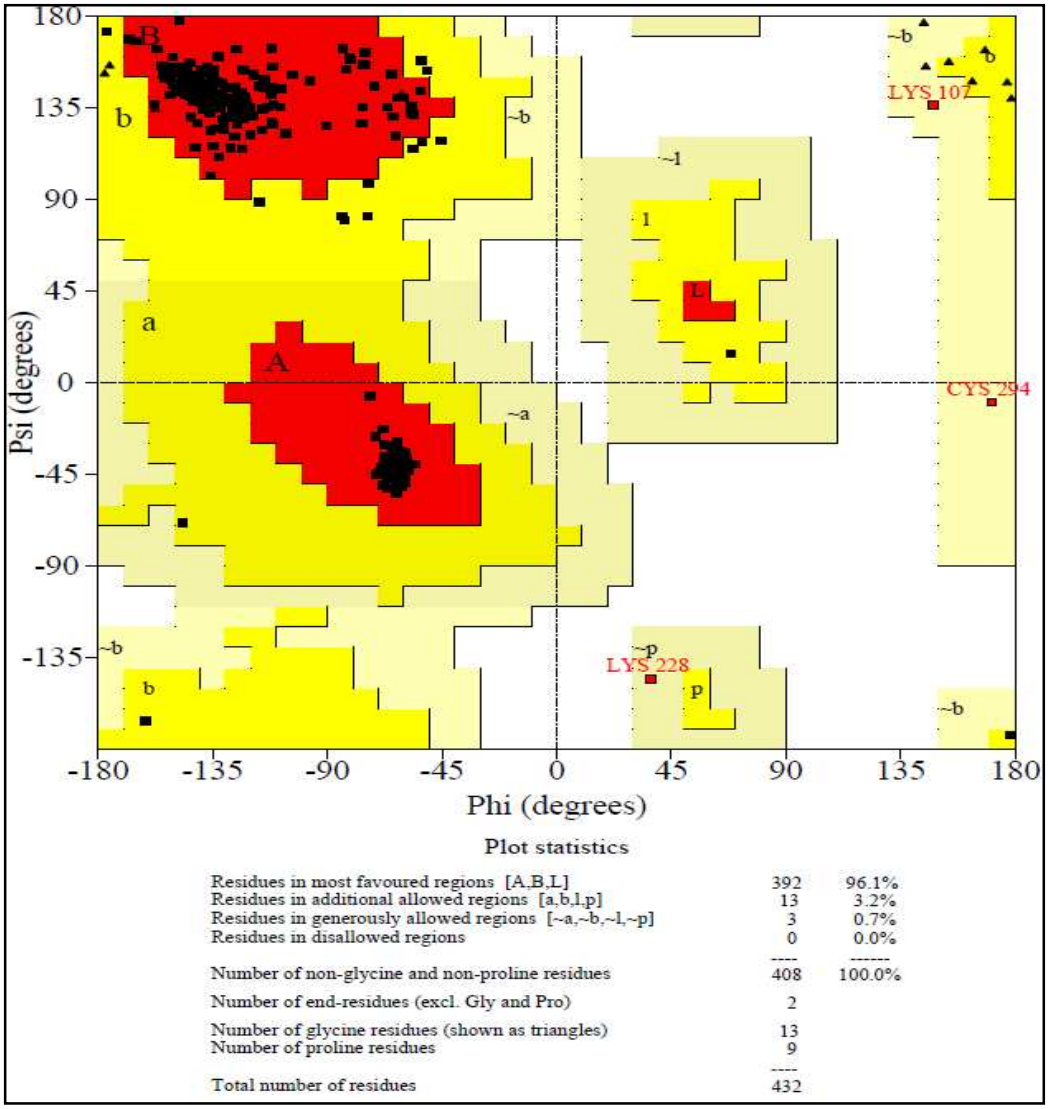


Figure-3 Shows torsion angles of phi (*φ*) and psi (*ψ*) in the generated model through Ramachandran plot.

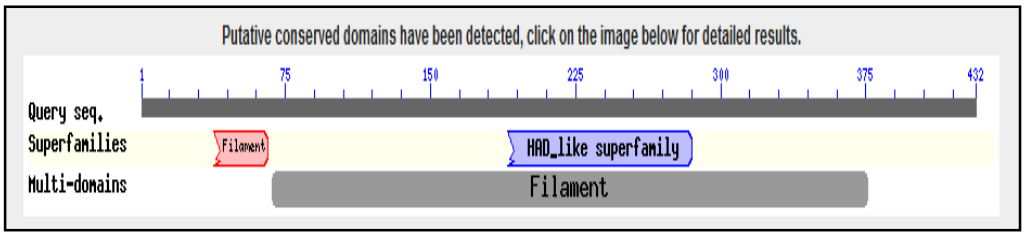
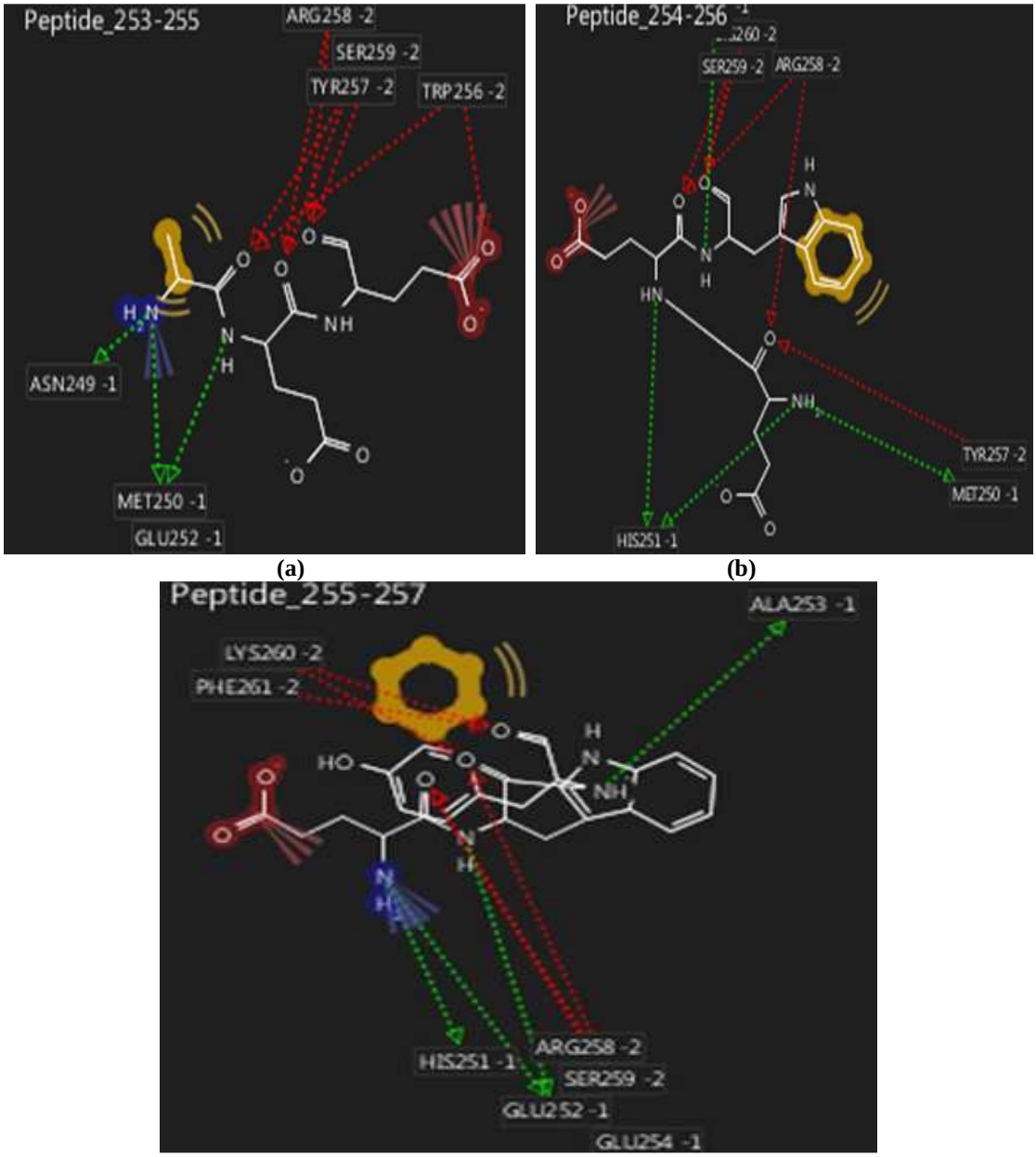


Figure-4 BLAST search result of GFAP

Template 3KLT was found to have high similarity (88%) with query sequence and possess same domain. Alignment of query with templates was showing the conservation of sequence from residue 254 to 269 i.e. EEWYRSKFADLTDAA. The sequence based on prediction has been

made through E1DS server and three catalytic sites are 254-259 (EEWYRS); 306-307 (LE); 394-401(LDTKSVSE). Active sites and residues (EEWYRS) were further verified through structural based prediction method using Pocket finder, SURFACE RACER4.0, SURFNET and LIGSITE. Ligand has been designed using Ligand Scout 2.0 to identify active site residues (EEWYRS) of GFAP. Number of ligands were generated through different pharmacophores formed by the predicted residues.

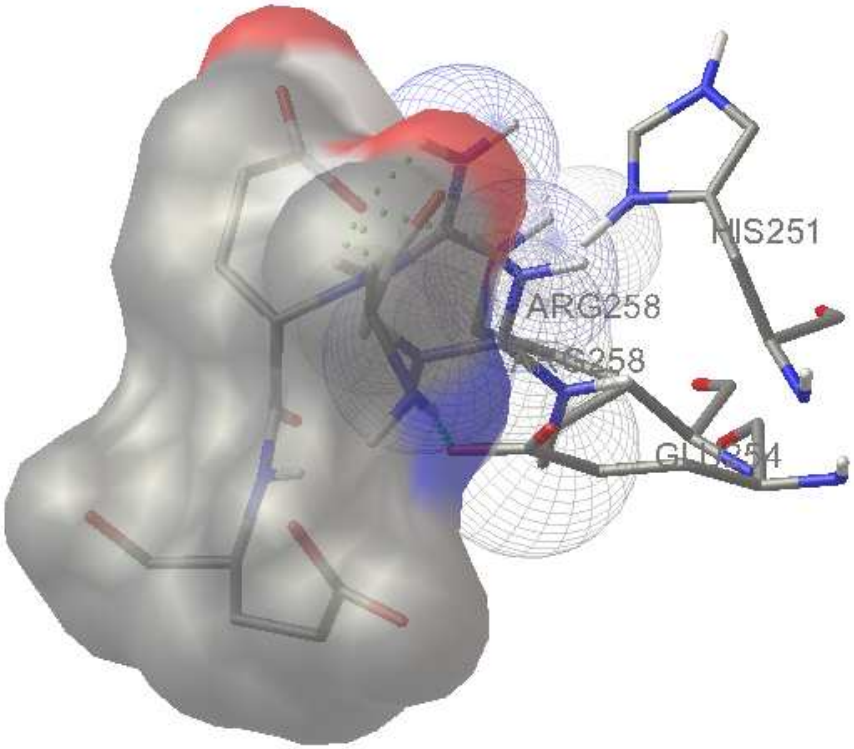


(a) **(b)**

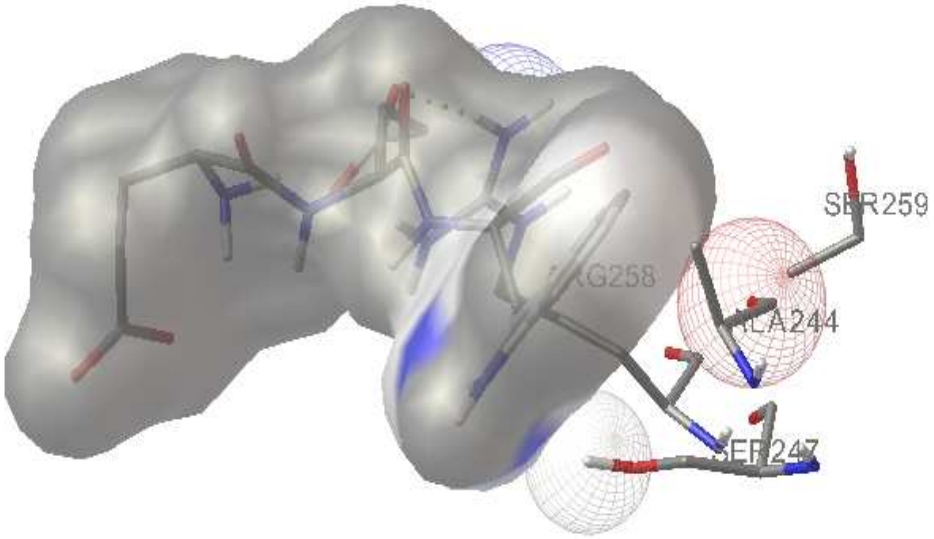
(c)

Figure-5 Interacting residues of designed ligand (Reference ligand-1(a), 2(b) and 3(c)) through created pharmacophore of Ligand Scout 2.0.

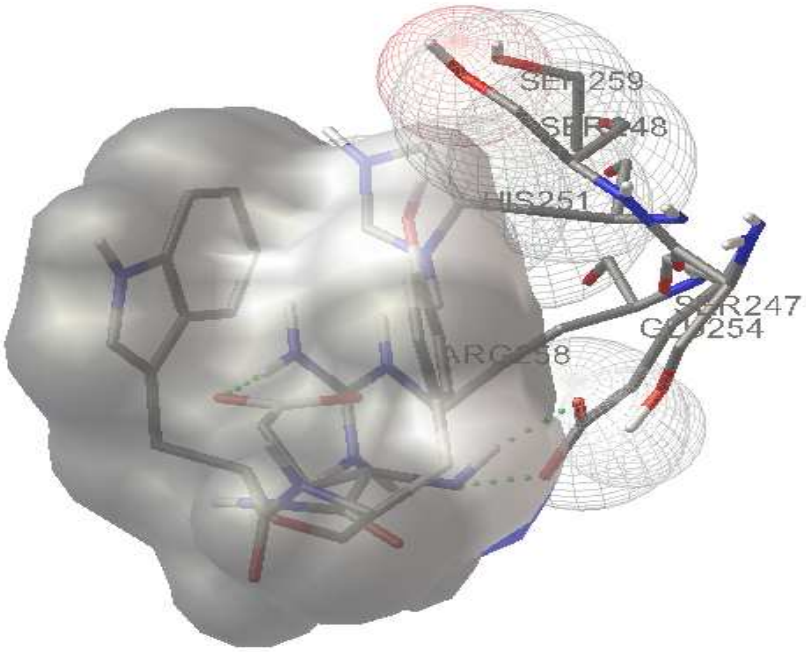
Three designed ligands were selected as reference ligands on the basis of desired interaction with active site residues through pharmacophore studies. Interactions and compatibility studies of ligands were further studied through docking process. Pharmacophore interactions with designed reference ligands are shown in Fig.-5(a-c).



(a)



(b)



(c)

Figure-6. Interacting residues of designed ligand (Reference ligand-1(a), 2(b) and 3(c)) through Autodock 4.0.

Compatibility of designed ligand and GFAP protein was checked by Docking. Docking has been used to predict the strength of association between ligand and receptor. It generally predicts the preferred orientation of one molecule to a second when bound to each other to form a stable complexes. Interactions between ligands and GFAP was further studied through AUTODOCK 4.0 which are shown in Fig.-6(a-c).

Docking studies of all the designed ligands were observed to interact with ARG-258 in docked complex (GFAP and Ligand). A comparative study of binding energy, inhibition constant (KI), intermolecular energy, internal energy and torsional energy, have been performed which are shown in Table-1. Binding energy is the sum of intermolecular energy, internal energy and torsional energy.

Table-1. Comparative docking results of designed reference ligand through Autodock 4.0.

Properties	Reference Ligand-1	Reference Ligand-2	Reference Ligand-3
Binding Energy (Kcal/Mol)	-4.04	-4.57	-3.26
Inhibition Constant (Ki)	1.1mM	444.45µM	4.06mM
Intermolecular Energy (Kcal/Mol)	-1.68	-3.76	-2.57
Internal Energy (Kcal/Mol)	-5.65	-4.66	-4.53
Torsional Energy (Kcal/Mol)	3.29	3.34	3.84

From docking results, it has been found that all three designed ligands were interacting within the cavity and ARG (258) was identified as the most interacting residues among all the other residues such as SER (259), GLU (254).

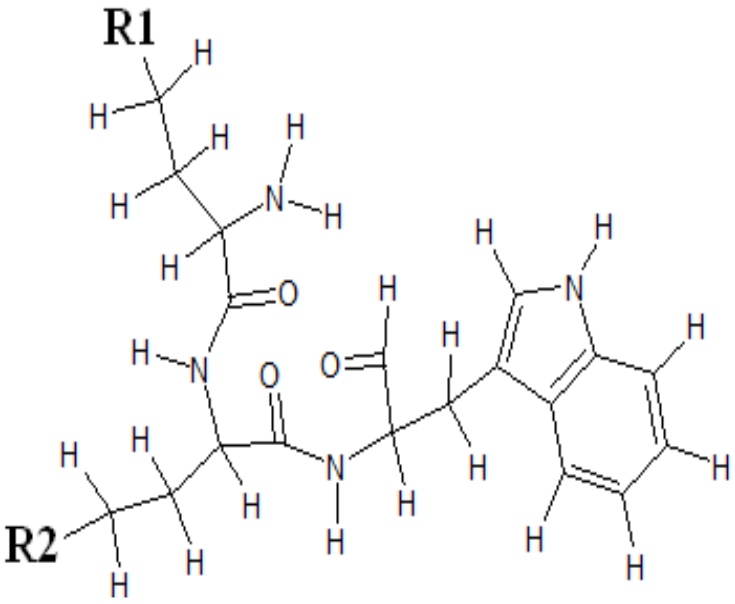


Figure-7. Reference Ligand-2 with non-interacting sites R1 and R2.

Table-2. Comparative study of binding affinity, moldock score and hydrogen bonding of reference ligand with analogues for inhibitory sites

R1	R2	BindingAffinity	MVD Score	HBond
O	O	-13.2576	-69.9629	-1.38221
NCH_2OH	O	-17.4388	-84.3609	-3.2701
O	CH_3COO	-18.2192	-69.067	-2.60252
O	SO_3	-15.8784	-81.9897	-3.3285
O	HPO_3H_3	-17.5063	-92.446	-5.82433
H_2NO_3	O	-18.6898	-78.2512	-2.19226
O	CH_3CON	-17.1631	-82.8651	-5.46681
O	H_2SO_2	-15.2391	-92.8139	-2.29838
O	H_2PO_3	-19.6796	-91.723	-3.54808
O	NO	-16.6949	-67.4472	-4.97514
OCH_2OH	O	-19.3235	-65.239	-4.06498
O	H_2NSO_2	-14.2156	-84.0046	-1.899423
O	$OPHO_3H_3$	-17.9621	-91.342	-4.49557
O	NO_2	-20.3643	-90.2304	-5.01569
NCH_2OH	O	-15.7594	-76.737	-2.62995
O	SO_2CL	-18.2864	-84.5328	-2.59623
O	CH_3O	-16.7664	-75.732	-1.73349
O	NSO_3	-20.4172	-80.1921	-3.70483
O	SO_4	-19.0201	-78.651	-2.75008
O	SO_2F_3	-27.1078	-81.3241	-4.71539
H_2NO_2	O	-17.6217	-87.7839	-3.10955
CH_3COO	O	-19.0857	-73.8254	-2.88456
SO_3	O	-28.4251	-81.3332	-3.20623
HPO_3H_3	O	-14.1324	-78.3004	-2.21843
H_2NO_3	O	-20.5635	-86.4237	-5.27512

Thus ARG (258) could be catalytic site which can be used for further studies. Compatibility of protein and ligands were measured through binding energies (BE). Reference ligand-2 showed better binding energy (-4.57) than reference ligand-1(-4.04) and reference ligand-3 (-3.26). Moreover, the inhibitor constant (Ki) for reference ligand-2 (444.45µM) also gave less inhibitory concentration than reference ligand-1(1.1mM) and reference ligand-3(4.06mM), indicating a high affinity of reference ligand-2 towards receptor. Hence reference ligand 2 (Fig.-7) was used for further studies i.e. analogue designing.

Various analogues were constructed by adding different functional groups at R1 and R2 sites to obtain the ligands having greater binding affinity with the receptor. Additional functional groups at R1 increased the affinity of ligands at the desired inhibitory site [18]. However, attachment of different functional groups at R2 site created an environment by interacting with neighbouring residues. The designed analogues finally gave better calculations for binding affinity, MolDock score and re-ranking score than their corresponding reference ligands, listed in Table 2.

CONCLUSION

In the present study an attempt has been made for in silico prediction for wet lab support in determination of three-dimensional structure of GFAP using molecular modelling and simulation techniques. Model generation and refinement have been done using systematic implementation of various computational techniques such as sequence analysis, homology modelling and energy minimization. ARG (258) has been identified as catalytic residue. Designed ligand and analogues can play an important role in identifying the pathogenesis of GFAP associated diseases and disorders.

Acknowledgement
I am thankful to Ms. Pallavi Chauhan for some technical support and cooperation.

REFERENCES

[1] RA Quinlan; M Brenner; JE Goldman; A Messing; *Exp. Cell Res*, **2007**, 313, 2077 – 2087.
[2] D Dahl; *DC* Rueger; *A* Bignami; K Weber; M Born; *Eur. J Cell Biol*, **1981**, 24, 191 -196.
[3] S A Reeves; L J Helman; A Allison; MA Israel; *Proc. Natl. Acad. Sci*. U.S.A.,**1989**, 86, 5178-82.
[4] T. Kaneko; S Kasaokaa; T Miyauchia; M Fujitaa; Y Odaa; R Tsurutaa; T Maekawaa; *Resuscitation*, **2009**, 80/7, 790-794.
[5] E. Bongcam-Rudloff; M Nistér; C Betsholtz; JL Wang; G Stenman; K Huebner; CM Croce; B Westermark; *Cancer Res*., **1991**, 51, 1553-60.
[6] M. Inagaki; Y Gonda; K Nishizawa; S Kitamura; C Sato; S Ando; K Tanabe; K Kikuchi; S Tsuiki; Y Nishi; *J Biol Chem*, **1990**, 265/8, 4722-9.
[7] W. Liedtke; W Edelmann; PL Bieri; FC Chiu; N J Cowan; R Kucherlapati; C S Raine; *Neuron*, **1996**,17/4, 607–15.
[8] M Brenner; AB Johnson; OB Tanguy; D Rodriguez; JE Goldman; A Messing; *Nat. Genet*, **2001**, 27/1, 117–20.

[9] S F Altschul; T L Madden; A A Schäffer; J Zhang; Z Zhang; W Miller; D J Lipman; *Nucleic Acids Res*, **1997**, 25/17, 3389–3402.
[10] R Chenna; H Sugawara; T Koike; R Lopez; T J Gibson; D G Higgins; J D Thompson; *Nucleic Acids Res*, **2003**, 31, 3497-3500.
[11] A Sali; T L Blundell; *Molecular Medicine Today*, **1995**, 1, 270-277.
[12] A Sali; E Shakhnovich; M Karplus; *Nature*, **1994**, 369, 248-251.
[13] R A Laskowski; JAC Rullmann; M Arthur; R Kaptein; J M Thornton; *Journal of Biomolecular NMR*, **1996**, 8, 477-486.
[14] G Vriend; *J. Mol. Graph*, **1990**, 8/1, 52-56.
[15] S Kushwaha; P Chauhan; *The Internet Journal of Infectious Diseases*, **2010**, 8/1,
[16] G Wolber; T Langer; *J. Chem. Inf. Model*., **2005**, 45/1, 160 -169.
[17] W J A Van Geel; H P M Reus De; H Nijzing; M M Verbeek; P E Vos; K J B Lamers; *Clin Chim Acta*, **2002**, 326, 151– 4.
[18] P E Vos; KJB Lamers; JCM Hendriks; M Haaren; T Beems; C Zimmerman; W Geel van; H de Reus; J Biert; MM Verbeek; *Neurology*, **2004**, 62, 1303– 10.
[19] N J Laping; B Teter; N R Nichols; I Rozovsky; *C E Finch; Pathology*, **1994** ,1, 259-275.
[20] S Wu; Y Zhang; *Nucleic Acids Research*, **2007**, 35, 3375-3382.
[21] P Chauhan; M Shakya; *Bioinformation*, **2009**, 4/6, 223-228.