

Research Article

HOMOLOGY MODELING AND DOCKING STUDIES OF HUMAN CHITOTRIOSIDASE WITH ITS NATURAL INHIBITORS

Deepsikha Roy and Koel Mukherjee

Department of Bio-Engineering, Birla Institute of Technology, Mesra 835215, Jharkhand

Abstract: Chitinase inhibitors have been found to have anti-inflammatory potential against asthma, allergic diseases and various other disorders. In this study various naturally occurring chitinase inhibitors against human chitinase (chitotriosidases, CHIT1) were studied with the help of protein-ligand docking. The structure of CHIT1 was modeled by homology modeling tool and validated with the help of various computational tools. Following validation, secondary structure, function and solvent accessibility of the protein was analyzed. A molecular dynamics (MD) simulation study was conducted by GROMACS simulation package to study the stability of the structure. This was further followed by docking studies with natural inhibitors like allosamidin, argifin and argadin against CHIT1 by GLIDE docking software. Argadin was observed to have the highest affinity (G-score = -10.955) towards CHIT1 and allosamidin scored the lowest GLIDE score (G-score = -7.741). The structural behavior of the best inhibitor protein complex (CHIT1- argadin) was validated through molecular dynamics simulation studies. A structure based virtual screening on the basis of the binding modes of these inhibitor was performed and best scoring hits were identified. The sequence analysis can be further used for the designing of potent drugs against diseases caused by CHIT1, thereby aiding knowledge in the field of research.

Keywords: Homology modeling; MD simulation; protein-ligand docking, GLIDE

INTRODUCTION

Recent studies have shown that Chitinase (C) / Chitinase-like proteins (CLPs) have a role in the development and progression of allergic inflammation and tissue remodeling (Lee, 2009). Although mammals cannot synthesize or metabolize chitin, a number of chitinolytic enzymes [true chitinases, e.g., chitotriosidases, acidic mammalian chitinase (AMCase)] or chitin-binding proteins [e.g. Ym-1, Ym-2, breast regression protein 39(BRP-39, chondrocyte protein-39)] were discovered in mammals (Elias *et al.*, 2005 ; Myles *et al.*, 2007). Many of the chitinase family proteins are constitutively

expressed in macrophages and epithelial cells of the lung and digestive tracts, consisting of the body's first line of defense against exogenous agents including chitin-containing pathogens (Homer *et al.*, 2006 ; Mizoguchi, 2006). CHIT1 levels in the human body have been proven to be a diagnostic marker for several diseases like Gaucher's disease, malaria, thalassaemia and Glycogen Storage Diseases. Recently, it was discovered that humans have a chitinase activity that was found to be elevated up to 2 orders of magnitude in the plasma of patients suffering from Gaucher's disease (Barranger *et al.*, 1988; Hollak *et al.*, 1994). Purification of this activity from a Gaucher spleen revealed that, although transcribed from a single gene, the enzyme occurs in two major forms of 39 and 50 kDa. The enzyme is able to cleave chitotriose (hence was termed chitotriosidase) but also hydrolyzes colloidal chitin to yield chitobiose and is thus thought to

Corresponding Author: **Koel Mukherjee**
E-mail: koelmukherjee@bitmesra.ac.in

Received: February 6, 2015

Accepted: April 18, 2015

Published: April 21, 2015

be an exochitinase (Renkema *et al.*, 1995; Boot *et al.*, 2001). CHIT1 has a high acid tolerance which has been ascribed to the presence of Asn208, Arg 269, and Gln145 (Olland *et al.*, 2009) near the catalytic residues. It has been suggested that CHIT1 is acting as an endo-enzyme. Endo-activity would benefit from an extended substrate binding cleft with more than two subsites on each site of the catalytic center where cleavage takes place. High oligosaccharide affinity in multiple positive subsites has been shown to be beneficial for trans-glycosylation activity (Zakariassen *et al.*, 2011). The presence of stretches of aromatic residues, including a conserved Trp-Tyr-Trp-Trp motif spanning an area corresponding to what could be subsites -6 to -1 and a Trp-Trp motif in the +1 and +2 subsites (using the nomenclature proposed by Davies *et al.* where subsites are labeled from -n to +n, - n represents the non-reducing end and +n the reducing end, with cleavage taking place between the -1 and +1 subsites. (Davies *et al.*, 1997) suggests extended substrate-binding clefts in the protein.

Consequently, chitinase inhibitors were suggested to have anti-inflammatory potential against asthma and allergic diseases, including atopic dermatitis and allergic rhinitis (Zhu *et al.*, 2004).

Most known chitinase inhibitors are classical reversible inhibitors that compete with the substrate by blocking its binding site, usually mimicking the stacking to one or more of the subsite tryptophans through protein-protein or other hydrophobic interactions and generating hydrogen bonds with potential partners in the active site. Oligo-acetylglucosamine exhibits this binding capability, but long carbohydrate polymers suffer enzymatic degradation while short oligosaccharides have only low binding affinities (Fusetti *et al.*, 2002). Additionally, the synthesis of carbohydrate oligomers is often challenging, thus precluding the development of drug-like compounds of this type. In recent years, numerous crystal structures of chitinase-inhibitor complexes have been published (Tabudravu *et al.*, 2002; Rao *et al.*, 2005; Bortone *et al.*, 2002; Papanikolau *et al.*, 2003; Rao *et al.*, 2003).

Allosamidin, a pseudotrisaccharide first isolated from the mycelium of *Streptomyces* sp.,

(Sakuda *et al.*, 1986; Sakuda *et al.*, 1987) is the most extensively studied chitinase inhibitor. Various biological properties have been reported associated with its activity as a chitinase inhibitor, including inhibition of cell separation in fungi (Kuranda *et al.*, 1991; Sakuda *et al.*, 1990), toxicity towards insect larvae and blocking of malaria parasite penetration into the mosquito midgut (Shahabuddin *et al.*, 1993). Recently, there has been considerable interest in the ability of peptides to mimic carbohydrate-protein interactions and bind to enzymes and lectins. Synthetic compounds that can mimic such peptide ligands, i.e. peptidomimetics, could provide potential drug candidates with improved stability toward proteolytic breakdown but retaining high potency.

Two classes of peptide-based chitinase inhibitors are known: the cyclic proline-containing dipeptides isolated from broth of a marine bacterium (Izumida *et al.*, 1996) and the cyclic pentapeptides argadin and argifin secreted from fungal strains (Omura *et al.*, 2000; Arai *et al.*, 2000). In principle further analogs are readily available by substitution of either amino acid residue. However, although several cyclic dipeptide derivatives have been synthesized, their chitinase inhibition is moderate (high micromolar range) (Houston *et al.*, 2002, 2004). The total syntheses of both argadin and argifin have been achieved (Dixon *et al.*, 2005, 2006) using a combination of solid phase and solution chemistry.

Although total synthesis of allosamidin have been reported previously in Berecibar *et al.*, 1999, the length and complexity of the synthetic routes involved severely limit both the availability of this compound and the scope for structure-based design of novel allosamidin-derived inhibitors. A similar type of problem appears with argifin (Anderson *et al.*, 2008) because its molecular weight and number of hydrogen-bond donors/acceptors well exceed the Lipinski criteria for drug-likeness (Lipinski, 2000). In order to overcome the above drawbacks, this present study is aimed towards searching for a potent inhibitor for CHIT1 by *in silico* methods. The structure of CHIT1 has been modeled followed by validation tests. The structure stability has

been analyzed with the help of MD simulations. This was followed by docking the natural inhibitors with CHIT1 using GLIDE. The best natural inhibitor was used as a reference for screening new inhibitors from ZINC database. The present structure-based drug designing approach can be used to enhance more refined inhibitory property of novel lead molecules against CHIT1 that will aid knowledge in combating various inflammatory and allergic disorders.

MATERIAL AND METHODS

Retrieval of protein sequence

For *in silico* characterization, the sequence of CHIT1 protein of *Homo sapiens* (Accession: AAI05682.1, Gene id:111494114) was retrieved from National Center for Biotechnology Information (NCBI) database (<http://www.ncbi.nlm.nih.gov>) in the FASTA format which were further used for analysis.

Primary and secondary structure analysis

The sequence retrieved from NCBI was characterized using Protparam (Sivakumar *et al.*, 2007). The physicochemical parameters, theoretical isoelectric point (pI), molecular weight, total number of positive and negative residues, extinction coefficient (Gasteiger *et al.*, 2005), instability index (Gill *et al.*, 1989), aliphatic index (Guruprasad *et al.*, 1990) and grand average hydropathy (GRAVY) (Ikai, 1980) were computed using the ExPASy's prediction server Protparam. Secondary structure analysis was carried out using PDBSum server (Laskowski *et al.*, 1997) and Cys_Rec tool (<http://sun1.softberry.com/berry.phtml?topic=Cys-Rec&group=help&subgroup=prop>)

Template selection and alignment of the target

The first task in homology modeling technique requires recognition of the protein structures linked to the target sequence and the subsequently selection of templates. PSI-BLAST (Altschul *et al.*, 1997) was carried out against database of PDB proteins available at the National Centre for Biotechnology Information (NCBI) Web server (<http://www.ncbi.nlm.nih.gov/blast/>).

The modeling tool MODELLER 9.1.2 gave a list of 4 suitable templates which were individually aligned with the target protein sequence by local pair wise-alignment method using EMBOSS-WATER (http://www.ebi.ac.uk/Tools/psa/emboss_water/) tool of EMBL-EBI (<http://www.ebi.ac.uk/>). An appropriate template was selected on the basis on maximum identity, similarity and alignment score.

3D structure prediction and validation

CHIT1 structure is available in Protein Data Bank, but it was inferred that the protein structure has several missing atoms and residues. So, the entire protein structure was remodeled using MODELLER 9.1.2 program (Sali and Blundell, 1993). In order to construct a 3D model for the target protein MODELLER, extracts spatial restraints from two sources. First, homology-derived restraints on the distances and dihedral angles in the target sequence are extracted from its alignment with the template structures. Second, stereochemical restraints such as bond length and bond angle preferences are obtained from the molecular mechanics force field of CHARMM-22 (MacKerell *et al.*, 1998) and statistical preferences of dihedral angles and non-bonded atomic distances are obtained from a representative set of all known protein structures. The modeled structure was analyzed through Ramachandran plot via Rampage server (Hirokawa *et al.*, 1998). The plot gives the percentage of residues that are present in either of the favored, allowed and disallowed regions. SAVES (Structure Analysis and Verification Server) (<http://nihserver.mbi.ucla.edu/SAVES/>) was used to carry out the verifications of the model with PROVE (Pontius *et al.*, 1996), VERIFY3D (Eisenberg *et al.*, 1993) and ERRAT (Colovos *et al.*, 1993).

In order to examine whether the predicted structure of protein remains stable in the presence of explicit solvent from a dynamic point of view, the molecular dynamic simulation was performed with GROMOS 96-43a1 force fields with the periodic boundary conditions(PBC) by using GROMACS 4.0 package (DePristo *et al.*, 2005). The 3Dstructure of the protein was taken in a cubic box with a 3.0 Å edge length. To solvate the

condition the “SPC” water model (spc216.gro file) was used to fill up the box. Before simulation, energy minimization was performed by steep and conjugate gradient (cg) methods to regularize the protein structure geometry. MD simulation was performed for a period of 10ns maintaining the temperature at 300K and pressure at 1atm. After completion of simulation, the trajectory files were generated, which were analyzed with different tools of GROMACS.

Function prediction

On the basis of the sequence information, the functions were further predicted using 3D2GO server (Neu, 1992). 3D2GO predicts the function of your protein using 3 sources of information: Overall topological similarity to structures with known function, Geometric and residue similarity of predicted functional sites to regions of known structures and Sequence homology to functionally annotated sequences. These functions would help to know the type of protein which would in turn help in designing potent inhibitor against it.

Docking

Docking studies with natural inhibitors like allosamidin, argifin and argadin against CHIT1 were done. The protocol of docking of ligands with the receptor was performed using by GLIDE (Grid-based Ligand Docking with Energetics) (Friesner *et al.*, 2004) module of Schrodinger suit-2013. GLIDE has been designed to perform as close to an exhaustive search of the positional, orientational and conformational space available to the ligand as is feasible while retaining sufficient computational speed to screen large libraries. This has been accomplished via the use of a series of hierarchical filters. GLIDE approximates a complete systematic search of the conformational, orientational and positional space of the docked ligand. In this search, an initial rough positioning and scoring phase that dramatically narrows the search space is followed by torsionally flexible energy optimization on an OPLS-AA non-bonded potential grid for a few hundred surviving candidate poses. The very best candidates are further refined via a Monte Carlo sampling of pose conformation; in some cases,

this is crucial to obtaining an accurate docked pose. Selection of the best docked pose uses a model energy function that combines empirical and force-field based terms. Docking accuracy is assessed by redocking ligands from 282 co-crystallized PDB complexes starting from conformationally optimized ligand geometries that bear no memory of the correctly docked pose. Errors in geometry for the top-ranked pose are less than 1 Å in nearly half of the cases and are greater than 2 Å in only about one-third of them (Friesner *et al.*, 2004).

- 1) The protein CHIT1 was downloaded from PDB database and was prepared in the GLIDE module of Schrodinger suit-2013.
- 2) The inhibitors (allosamidin, argadin and argifin) were downloaded from drug bank. Ligand preparation was done in the GLIDE module.
- 3) Grid for the binding site was prepared using the coordinates of the active site residues (Asp-138 and Glu-140).
- 4) GLIDE docking was started using the GLIDE grid generated and the ligands prepared.
- 5) Results were analyzed and the inhibitor with highest binding affinity was selected.
- 6) Best natural inhibitor was used to search its identical compounds and the resultant compounds were docked with CHIT1 protein.
- 7) The best docking compound was obtained using GLIDE.

Glide 2.5 employs two forms of GlideScore: (i) GlideScore 2.5 SP, used by Standard-Precision Glide; (ii) GlideScore 2.5 XP, used by Extra-Precision Glide. These functions use similar terms but are formulated with different objectives in mind. Specifically, GlideScore 2.5 SP is a “softer”, more forgiving function that is adept at identifying ligands that have a reasonable propensity to bind, even in cases in which the Glide pose has significant imperfections. This version seeks to minimize false negatives and is appropriate for many database screening applications. In contrast, GlideScore 2.5 XP is a harder function that exacts severe penalties for poses that violate established physical chemistry principles such as that charged and strongly

- polar groups be adequately exposed to solvent. This version of GlideScore is more adept at minimizing false positives and can be especially useful in lead optimization or other studies in which only a limited number of compounds will be considered experimentally and each computationally identified compound needs to be as high in quality as possible. In this paper, we report SP mode GlideScore (Friesner *et al.*, 2004).
- 8) The stability of the complex structure (protein-ZINC72124668) was again validated using MD simulation for 10ns.

RESULT AND DISCUSSIONS

Structural Analysis

The sequence taken from NCBI was used for primary and secondary structural analysis. In the analysis, various parameters were calculated. Table 1 gives the physicochemical characterization of the protein. ProtParam prediction server of ExPASy calculated Extinction coefficient, instability index, aliphatic index and GRAVY. The extinction coefficient tells that how strongly a protein can absorb light at a given wavelength. Aliphatic index infers a measure of relative volume of a protein occupied by aliphatic side chains while instability index gives a measure of proteins, used to determine whether it is stable or not *in vitro*. The GRAVY score indicates whether the proteins have better

interactions with water molecules or not. pI is the isoelectric point of proteins. $pI < 7$ indicates that the proteins are acidic in nature, while $pI > 7$ indicates that the proteins are basic in nature. pI in this protein has a value of 6.55 indicating it to be slightly acidic. Instability index indicates the stability of protein *in vitro*. Proteins having $\text{Instability Index} < 40$ are considered to be stable while those having value greater than 40 are unstable. The instability index value for this protein is 35.66 (Table 1) indicating that it is stable *in vitro*. Aliphatic index indicates the thermal stability of the globular proteins. The value is 66.82 for CHIT1 which infers that the protein can remain stable in a wide range of temperatures. The GRAVY score of -0.390 indicates that it has better interactions with water molecules and hence it is hydrophilic in nature.

PDBSum server mainly gives the information of a secondary structure in any protein. It infers which secondary structure predominates in a protein. CYS_REC identifies the positions of cysteine, total number of cysteine present and predicts the most probable disulfide bond pattern of pairs in the protein sequence. The disulphide bonds also reflect the stability in 3D structure of the concerned protein. Table 2 gives all these information about CHIT1 protein. The result analysis reveals that beta turns are dominant over all the other secondary structures followed by extended strands, alpha helices and sheets. The results also provide the predicted disulfide

Table 1
Physico-chemical characteristics of Chitotriosidase

Length	Iso-Electric Point	Molecular weight	Extinction Coefficient	Positive Residues and Negative Residues	Instability Index	Aliphatic Index	Grand Average Hydropathy (GRAVY) Score
466	6.55	51681.3	94935	(+)R : 44 (-)R : 45	35.66	66.82	-0.390

Table 2
Secondary structure information of Chitotriosidase

Alpha Helices	Strands	Beta Turns	Sheets	Number of Cysteine Residues	Disulphide Bonds Predicted
12	21	48	7	10	26-440, 51-462, 307-370, 420 460, 450-463

bridges and possible bonding pairs between the cysteine residues. Disulfide bonds are very important for the stability of tertiary and quaternary structures of proteins. Even the function of certain soluble proteins and cell surface receptors is controlled by the cleavage of one or more disulfide bonds. There are 10 cysteine residues in this protein with 5 disulphide bonds indicating the protein to be functionally quite stable.

Structural Modeling

After analyzing the primary and secondary structures of the protein, their 3D structures were taken from PDB for their functional predictions. It was seen that the protein's structure had many missing atoms and residues because of which it was remodeled using Modeler 9.1.2 modeling tool. MODELLER works on the principle of homology modeling and builds models with the most suitable template (among all), which are ranked on the basis of satisfaction of spatial restraints and their DOPE score. The templates with PDB ids 1edqA, 1LL7A, 1goiA, 1vf8A were given by MODELLER after sequence comparison. The templates were aligned with CHIT1 sequence and 1edqA was selected with a maximum alignment score of 779.5. The tool gave a result of 5 models. It has been shown that three-dimensional protein structure is evolutionarily more conserved than would be expected on the basis of sequence conservation alone (Kaczanowski and Zielenkiewicz, 2010). The modeled structures were analyzed via Rampage and SAVES Servers. The analysis inferred that model 1 was the best model based on results from Ramachandran plot, VERIFY3D, ERRAT AND PROVE. Table 3 gives the validation results of all the 5 models.

Figure 1 shows the ribbon representation of the best protein model having each secondary structure in a specific color. Figure 2 is the Z-score plot of the protein. Deviations of the atomic volumes from the standard values, evaluated as the volume Z-scores, are used to assess the quality of protein crystal structures. To score a structure globally, the volume Z-score root mean square deviation (Z-score rms) is computed, which measures the average magnitude of the volume

irregularities in the structure. It has been found that the Z-score rms decreases as the resolution and R-factor improve, consistent with the fact that these improvements generally reflect more accurate models. From the Z-score rms distribution in structures with a given resolution or R-factor, the normal limits in Z-score rms values for structures solved at that resolution or R-factor is determined. Structures whose Z-score rms exceeds these limits are considered as outliers. Such structures also exhibit unusual stereochemistry, as revealed by other analyses. Absolute Z-scores of individual atoms are used to identify problems in specific regions within a protein model(Pontius *et al*,1996).The protein's structure was further validated through molecular dynamics simulation using GROMACS. Figure 3 shows the root mean square fluctuation (RMSF) graph and the radius of gyration (R_g) plot of the protein. RMSF values show the flexibility of different segments of a protein. The figure shows that there is certain flexibility in the structure which can be further stabilized via simulation for a greater period of time stretch. R_g plot shows distance between the atoms and their centre of axis. In the following figure, the structure of protein shows certain fluctuations in the initial period of 700 ps while, there were fewer fluctuations in the R_g values after a period of 1000 ps. Figure 4 shows the total energy (T.E.) plot of CHIT1 protein. It tells that the overall T.E. of the protein during simulation remains stable at a range of -1906000 KJ/mol to -1896000 kJ/mol.

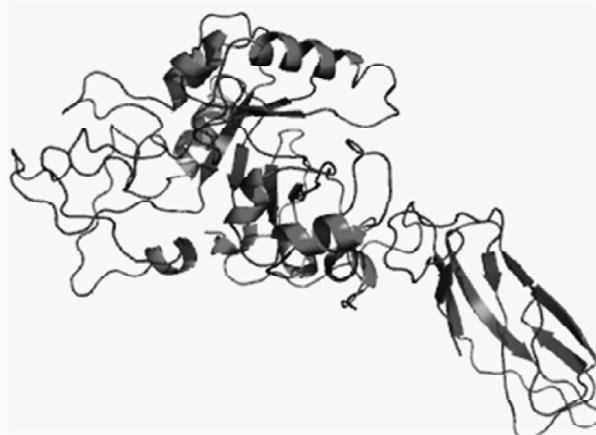


Figure 1: Ribbon representation of chitotriosidase protein. The figure represents the 12 helices and 7 sheets.

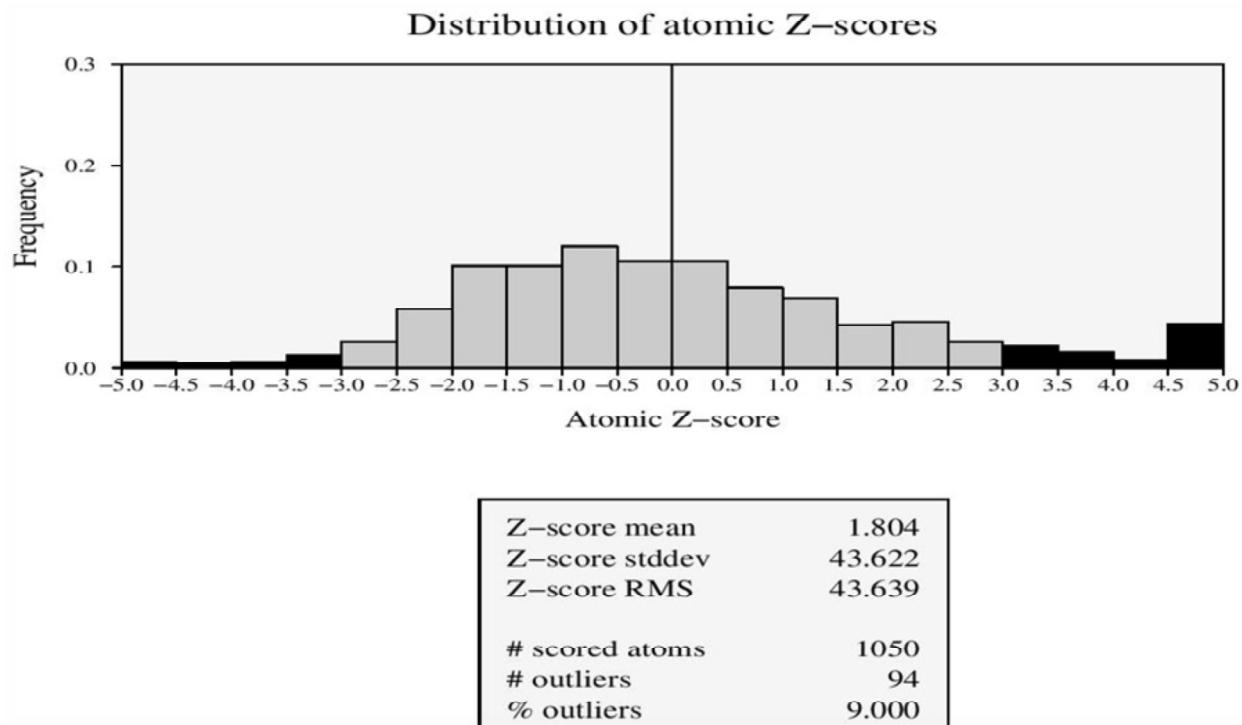


Figure 2: Atomic Z-score plot of the protein by ERRAT validation tool.

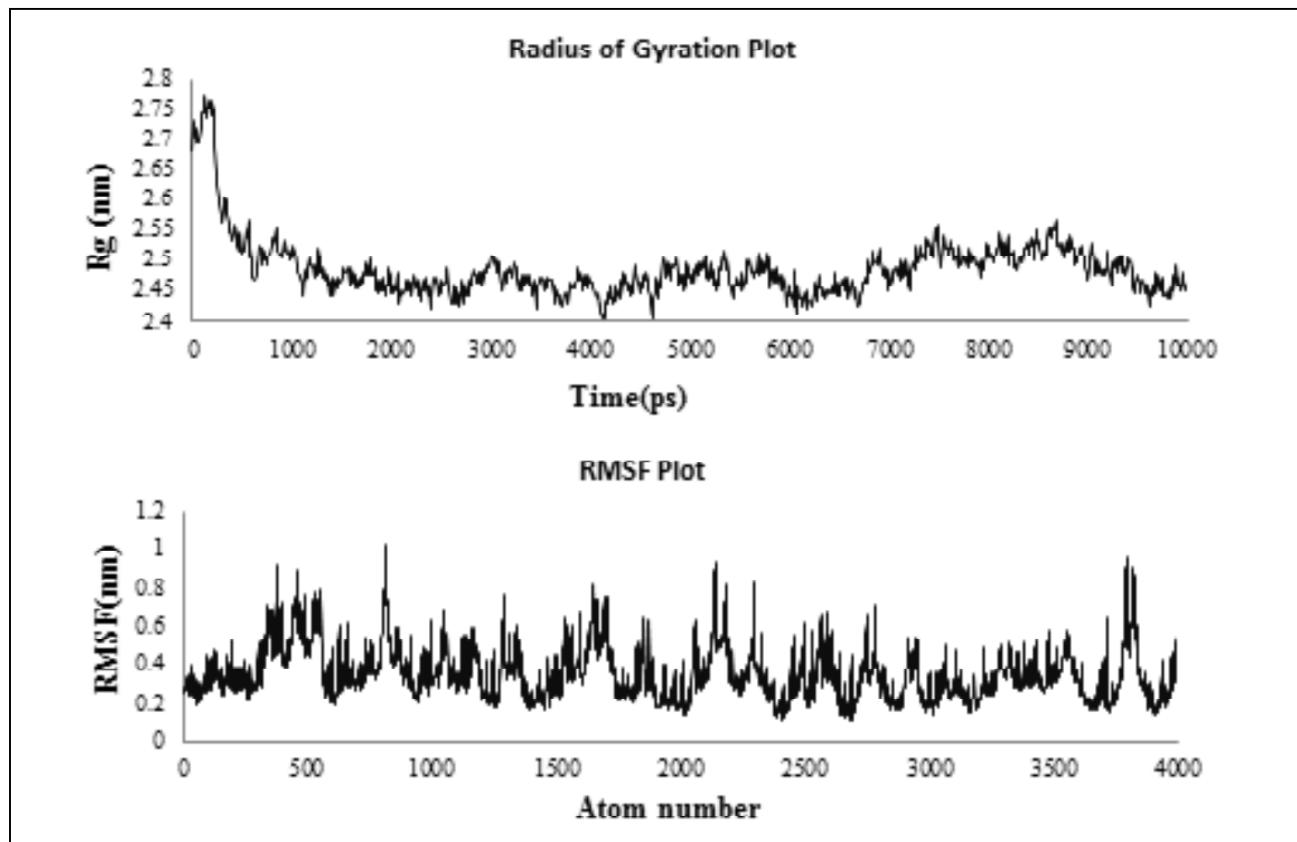


Figure 3: Root Mean square Fluctuation (RMSF) and Radius of Gyration (Rg) Plots. The x-axis represents the atomic number and time scale (ps) respectively for RMSF and Rg curves. The y-axis represents the fluctuation in nanometer scale.

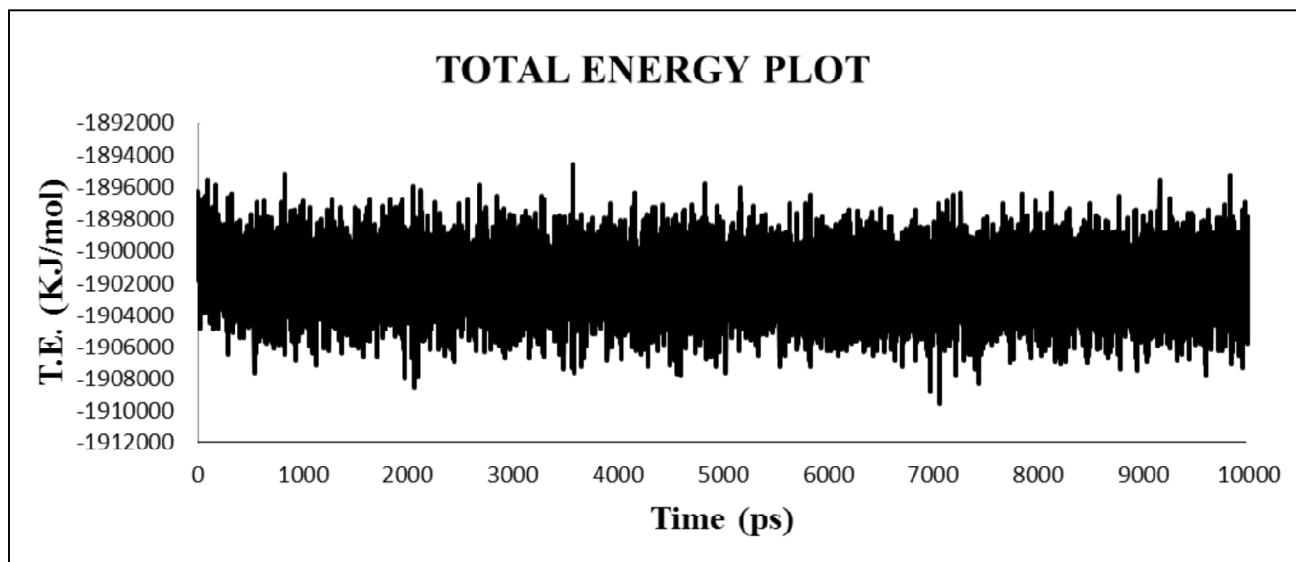


Figure 4: Total Energy plot of the modeled protein. The x-axis represents the time scale (ps) and y-axis represents the total energy in KJ/mol.

Table 3
Verification and validation scores of modeled Chitotriosidase

Model Number	Dope Score	VERIFY3D Score	Number of Residues in favored region in Ramachandran Plot (PROCHECK)	ERRAT Quality Factor	Z-Score (PROVE)
1.	-26643.85	60.96%	82%	36.08	1.804
2.	-26395.83	63.22%	81.4%	20.47	0.818
3.	-27021.18	57.18%	81.4%	23.44	2.709
4.	-27252.55	54.91%	79.3%	30.65	1.818
5.	-26670.44	52.39%	79%	31.09	2.725

Function Prediction

Followed by the stabilization of modeled protein, the functions of the protein were predicted using 3D2GO server. Table 4 displays the predicted function's data, according to which CHIT1 was found to have the highest confidence score for carbohydrate metabolic process. The results suggest that in addition to specific functions during chitin-containing pathogen attack onto the host, it can also have other functions in the human body, which can be further taken into consideration in order to design a potent drug against this protein's over-expression.

Protein-ligand docking

Given the biological and pharmaceutical significance of molecular docking, considerable efforts have been directed towards improving the

methods used to predict docking. After the completion of docking, the GLIDE scores and the hydrogen bond score for all the inhibitors bound complexes were determined. The ligand giving the least glide score was considered to be the most potent one. It was inferred from the docking results that argadin had the highest binding affinity for CHIT1 as compared to allosamidin and argifin. Figure 5a illustrates the interactions between the protein and natural inhibitor argadin. The ligand can be seen to bond with residues Arg -257, Trp -239, Arg-199 and His-149 of the protein via hydrogen bonding (represented by dotted lines). The docked complex of the protein and natural inhibitor is shown in Figure 5b. As per Wilkinson *et al.*, 1983, hydrogen bonding is an exchange reaction where the hydrogen bond donors and acceptors of the free

protein and ligand break their hydrogen bonds with water and form new ones in the protein-ligand complex. Following this inference, 9 compounds (identical to argadin) were taken from Zinc database. **Lipinski's rule of five** also known as the Pfizer's rule of five or simply the Rule of five (RO5) is a rule of thumb to evaluate drug-likeness or determine if a chemical compound with a certain pharmacological or biological activity has properties that would make it a likely orally active drug in humans. The rule was formulated by Christopher A. Lipinski in 1997, based on the observation that most orally administered drugs are relatively small and moderately lipophilic molecules. (Lipinsky *et al.*, 2001; Lipinsky, 2004). The physicochemical properties of the 9 compounds are shown in Table 5. Lipinski's rule states that, in general, an orally active drug has no more than one violation of the following criteria:

- No more than 5 hydrogen bond donors (the total number of nitrogen-hydrogen and oxygen-hydrogen bonds).
- Not more than 10 hydrogen bond acceptors (all nitrogen or oxygen atoms).
- A molecular mass less than 500 daltons.
- An octanol water partition coefficient $\log P$ not greater than 5.

Most of the compounds retrieved from the Zinc database have been found to follow Lipinsky rule. Hence these could be used as oral drugs. Compound ZINC26262591 showed to be most

Table 4
Predicted functions of Chitotriosidase protein using 3D2GO server

GO Term	Description	Confidence
GO:0005975	Carbohydrate metabolic process	0.40
GO:0004553	Hydrolase activity: hydrolyzing O-glycosyl compounds	0.38
GO:0043167	Ion binding	0.29
GO:0043169	Cation binding	0.28
GO:0008152	Metabolic process	0.24
GO:0016798	Hydrolase activity: acting on glycosyl bonds	0.24

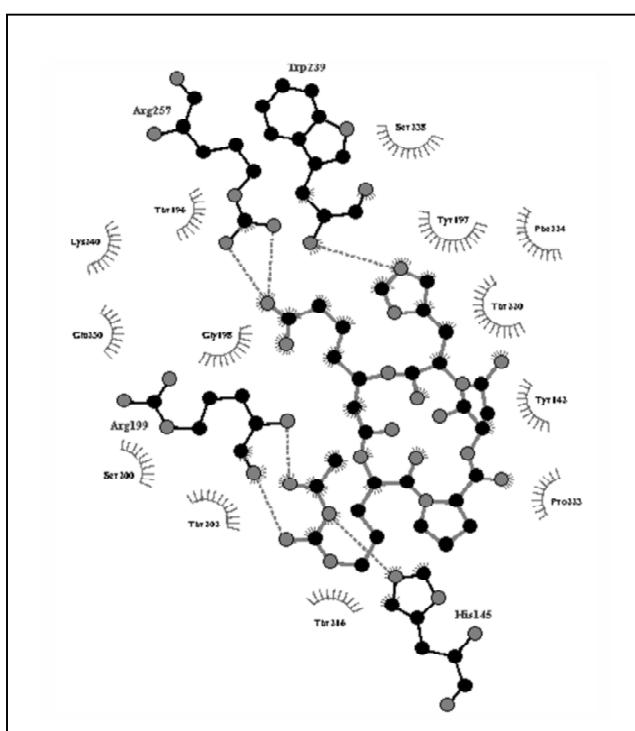
Table 5
List of docked inhibitors and their various scores

Inhibitor	GLIDE Docking Score (kJ/mol)	GLIDE H-Bond Score
ARGADIN	-10.955	-0.945
ARGIFIN	-7.557	-0.433
ALLOSAMIDIN	-7.714	-0.476
ZINC72124668	-9.241	-0.478
ZINC84652312	-8.838	-0.975
ZINC59473877	-8.635	-1.008
ZINC49777277	-8.604	-1.198
ZINC59222538	-7.051	-0.56
ZINC15721431	-6.903	-0.746
ZINC24693311	-6.718	-0.158
ZINC26262591	-6.469	-0.436
ZINC26577404	-5.369	-0.322

orally active in terms of physicochemical properties displayed in Table 6. The docking scores of the 9 compounds and the natural inhibitors with CHIT1 have been shown in Table 5. Ligand ZINC72124668 gave the best result with a GLIDE score of -9.241. Figure 6a illustrates the interaction profile of the protein and ligand ZINC72124668. The ligand can be seen to bind with residues Leu-327 and Glu-328 of the protein via hydrogen bonds. Leucine (Leu) being a non-polar amino acid and glutamic acid (Glu) being a negatively charged one, both of them show strong hydrogen bonding with the ligand. The docked complex of Zinc72124668 and the protein is shown in Figure 6b.

After docking, the stability of the protein-ligand complex was validated by means of MD simulation for 10ns (Rungta *et al.*, 2014). The plots are shown in Figure 7a and 7b. The plot in Figure 7a shows a starting RMSD value at 0nm, but between 1ps -2ps there was a dip in the plot. The RMSD values thereafter are found to be in the range of 1.5nm to 2.5nm till 5ns. In Figure 7b, a RMSD plot of 10ns has been shown. It was found that on extending the simulation time, the plot showed a comparatively narrower range of RMSD values between 0.6nm and 0.8nm from 5ns onwards. This indicated the protein-ligand complex to be stable in nature after 5 ns.

(a)

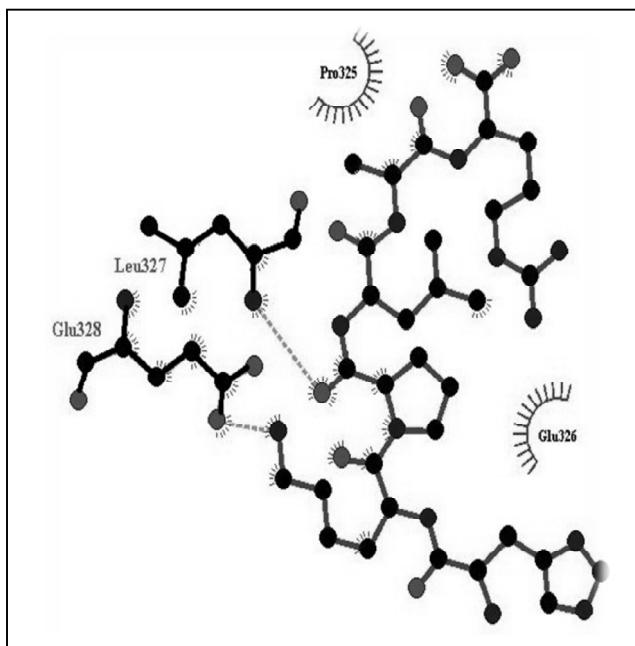


(b)



Figure 5: (a) Interaction profile of chitotriosidase and natural inhibitor argadin. The grey lines show the ligand bonds. The black lines represent non-ligand bonds and the dotted grey lines represent hydrogen bonds. The semi-circled residues represent hydrophobic residues in the protein. (b) Docked complex of chitotriosidase and argadin. The complex shows the hydrogen bonding of the protein with residues Arg -257, Trp -239, Arg-199 and His-149.

(a)



(b)



Figure 6: (a) Interaction profile of chitotriosidase and ligand ZINC72124668. The grey lines show the ligand bonds. The black lines represent non-ligand bonds and the dotted grey lines represent hydrogen bonds. The semi-circled residues represent hydrophobic residues in the protein. (b) Docked complex of chitotriosidase and ligand ZINC72124668. The complex shows the hydrogen bonding of the protein with residues Leu-327 and Glu-328.

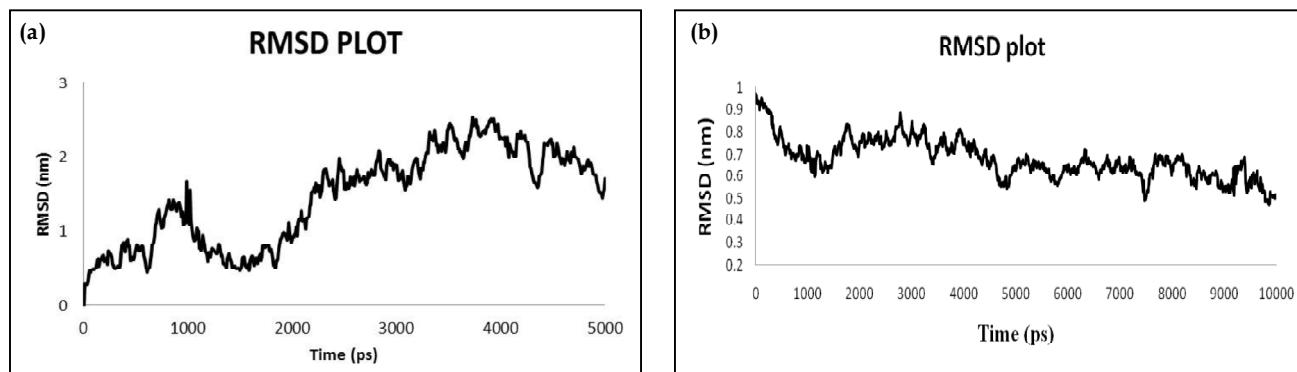


Figure 7: (a) Root Mean square Deviation (RMSD) plot of protein and ZINC72124668 complex. The x-axis represents the time scale (ps) till 5ns and y-axis represents the RMSD (nm). (b) Root Mean square Deviation (RMSD) plot of protein and ZINC72124668 complex. The x-axis represents the time scale (ps) till 10ns and y-axis represents the RMSD (nm).

Table 6
Properties of the 9 compounds from ZINC database

Zinc ID	xlogP	H-Bond Donors	H-Bond Acceptors	Net Charge	Mol. Wt.	Rotatable Bonds
24693311	-4.51	8	19	-1	673.708	11
84652312	-5.01	17	18	2	569.624	20
59222538	-5.34	17	20	2	626.676	19
26262591	-2.14	3	12	-1	405.391	6
26577404	-4.16	8	19	-1	677.692	9
72124668	-4.97	16	19	2	722.893	23
49777277	-4.57	14	21	0	722.805	23
59473877	-5.42	15	19	1	610.653	17
15721431	-4.8	12	14	2	467.531	13

CONCLUSION

For infectious diseases, the development of a single therapeutically useful compound is by no means the end of the story. Such a compound, when used on a large scale, will almost always lead to the occurrence of pathogens which have cleverly developed one or more methods to avoid the harmful effects of the drug. The processes used by resistant organisms and cancer cells to avoid the harmful effects of a drug vary widely. The constant threat of resistance is a major source of concern. The only way to combat resistant pathogens with new drugs with long-lasting utility is to develop cocktails containing multiple compounds acting in diverse ways. The presented work emphasizes the study of protein inhibitor interactions which has a vital role in Computer Aided Drug Designing. Simulation studies conclude that the observed conformational changes in the protein structure might have

occurred due to the binding of the inhibitor molecule.

Increased serum levels of CHIT1 were observed in patients suffering from disorders like Gaucher's disease, bronchial asthma, and atherosclerosis. Therefore, CHIT1 seems to have dual (regulatory and pathogenic) roles depending on the disease. Homology modeling of the protein produced 5 models and the best model was chosen based on scores given by various computational tools (like 82% residues in favoured region in Ramachandran plot and 36.082 quality factor score in ERRAT). Characterization of this modeled protein gave an insight into the structural properties of the protein. MD simulation of the selected model showed the stability of the model by means of RMSD, RMSF and energy plots. Molecules sharing a structural similarity with the best inhibitor could be used for identifying new potent inhibitors against

CHIT1. This structure-based drug designing approach can be used to enhance more refined inhibitory property of novel lead molecules against CHIT1 protein that will aid knowledge in combating various diseases caused due to elevated levels of CHIT1 in the human body.

Acknowledgement

The authors acknowledge BTISnet, Department of Biotechnology, Government of India and Birla Institute of Technology for providing Infrastructure facility for Bioinformatics research.

Abbreviations

AMCase, Acidic Mammalian Chitinase; CHIT1, Chitotriosidase; CLP, Chitinase-like Protein; EMBOSS, European Molecular Biology Open Software Suite; EMBL, European Molecular Biology Laboratory; EBI, FACTS, Fast analytical continuum treatment of solvation; GRAVY, Grand Average Hydropathy; MD, Molecular Dynamics; PDB, Protein DataBank; RMSD, Root Mean Square Deviation; RMSF, Root Mean Square Fluctuation; Rg, Radius of gyration; SAVES, Structure Analysis and Verification Server; T.E., Total energy.

References

- Altschul, S.F., Madden, T.L., Schäffer, 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 Res.* 25, 3389-3402.
- Andersen O.A., Nathubhai A., Dixon M.J., Eggleston I.A., Van Aalten D.M.F. (2008). Structure-Based Dissection of the Natural Product Cyclopentapeptide Chitinase Inhibitor Argifin. *Chem Biol.* 15, 295-301.
- Arai, N., Shiomi, K., Yamaguchi, Y., Masuma, R., Iwai, Y., Turberg, A., Koelb, H. and Omura, S. (2000). Argadin, a new chitinase inhibitor, produced by *clonostachys* sp. fo-7314. *Chem. Pharm. Bull. (Tokyo)*. 48, 1442-1446.
- Barranger, J. A., and Ginnis, E. I. (1988). Glucosylceramide Lipidoses: Gaucher's Disease. (Scriven, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 1677-1698. McGraw-Hill Inc., New York.
- Berecibar A., Grandjean C., Siriwardena A. (1999). Synthesis and biological activity of natural aminocyclopentitol glycosidase inhibitors: mannostatins, trehazolin, allosamidins and their analogues. *Chem. Rev.* 99, 779-844.
- Boot, R. G., Blommaart, E. F. C., Swart, E., Van der Vlugt, K. G., Bijl, N., Moe, C., Place, A., and Aerts, J. M. F. G. (2001). Identification of a Novel Acidic Mammalian Chitinase Distinct from Chitotriosidase. *J. Biol. Chem.* 276, 6770-6778.
- Bortone, K., Monzingo, A. F., Ernst, S. and Robertus, J. D. (2002). The structure of an allosamidin complex with the *Coccidioides immitis* chitinase defines a role for a second acid residue in substrate-assisted mechanism. *J. Mol. Biol.* 320, 293-302.
- Colovos, C., Yeates, T. (1993). Verification of protein structures: patterns of nonbonded atomic interactions. *Protein Sci.* 2, 1511-9.
- Davies, G.J., Wilson, K.S. and Henrissat, B. (1997). Nomenclature for sugar binding subsites in glycosyl hydrolases. *Biochem. J.* 321, 557-559.
- DePristo, M.A., De Bakker, P.I., Johnson, R.J. and Blundell, T.L. (2005). Crystallographic refinement by knowledge-based exploration of complex energy landscapes. *Structure.* 13, 1311-1319.
- Dixon, M.J., Andersen, O.A., Van Aalten, D.M.F., and Eggleston, I.M. (2005). An efficient synthesis of argifin: a natural product chitinase inhibitor with chemotherapeutic potential. *Bioorg. Med. Chem. Lett.* 15, 4717-4721.
- Dixon, M.J., Andersen, O.A., Van Aalten, D.M.F. and Eggleston, I.M. (2006). First synthesis of Argadin: A nanomolar inhibitor of family-18 chitinases. *Eur. J. Org. Chem.* 2006, 5002-5006.
- Elias, J.A., Homer, R.J., Hamid, Q., Lee, C.G. (2005). Chitinases and chitinase-like proteins in T(H)2 inflammation and asthma. *J. Allergy Clin. Immunol.* 116, 497-500.
- Eisenberg, D., Lüthy, R., Bowie, J.U. (1997). VERIFY3D: assessment of protein models with three-dimensional profiles. *Methods Enzymol.* 277, 396-404. Friesner, R.A., Banks, J.L., Murphy, R.B., Halgren, T.A., Klicic, J.J., Daniel, T., Mainz, Repasky M.P., Knoll, E.H., Shelley, M., Perry, J.K., Shaw, D.E., Francis, P. and Shenkin, P.S. (2004). GLIDE: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J. Med. Chem.* 47, 1739-1749.
- Fusetti, F., Von Moeller, H., Houston, D., Rozeboom, H.J., Dijkstra, B.W., Boot, R.G., Aerts, J. and Van Aalten, D.M.F. (2002). Structure of human chitotriosidase. Implications for specific inhibitor design and function of mammalian chitinase-like lectins. *J. Biol. Chem.* 277, 25537-25544.
- Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M.R., Appel, R.D. and Bairoch , A. (2005). Protein Identification and Analysis Tools on the ExPASy Server. The Proteomics Protocols Handbook, Humana Press. 571-607.
- Gill, S.C. and Von Hippel, P.H. (1989). Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.* 182, 319-326.
- Guruprasad, K., Reddy, B.V. and 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. *Prot. Eng.* 4, 155-161.
- Hirokawa, T., Boon-Chieng, S. and Mitaku, S. (1998). SOSUI: classification and secondary structure prediction system for membrane proteins. *Bioinformatics* 14, 378-379.

- Hollak, C.E.M., VanWeely, S., VanOers, M.H.J. and Aerts, J.M.F.G. (1994). Marked elevation of plasma chitotriosidase activity. A novel hallmark of Gaucher disease. *J. Clin. Invest.* 93, 1288-1292.
- Homer, R.J., Zhu, Z., Cohn, L., Lee, C.G., White, W.I., Chen, S. et al. (2006). Differential expression of chitinases identify subsets of murine airway epithelial cells in allergic inflammation. *Am. J. Physiol. Lung. Cell. Mol. Physiol.* 291:L502-L511.
- Houston, D.R., Eggleston, I., Synstad, B., Eijsink, V.G.H. and Van Aalten D.M.F. (2002). The cyclic dipeptide ci-4 inhibits family 18 chitinases by structural mimicry of a reaction intermediate. *Biochem. J.* 368, 23-27.
- Houston, D.R., Synstad, B., Eijsink, V.G.H., Stark, M.J.R., Eggleston, I.M., and VanAalten, D.M.F. (2004). Structure-based exploration of cyclic dipeptide chitinase inhibitors. *J. Med. Chem.* 47, 5713-5720.
- Ikai, A. (1980). Thermostability and Aliphatic Index of Globular Proteins. *J. Biochem.* 88, 1895-1898.
- Izumida, H., Imamura, N., and Sano, H. (1996a). A novel chitinase inhibitor from a marine bacterium *Pseudomonas* sp. *J. Antibiot. (Tokyo)*. 49, 76-80.
- Kaczanowski, S~ Zielenkiewicz, P (2010). Why similar protein sequences encode similar three dimensional structures?. *Theor Chem Acc.* 125, 643-50.
- Kuranda, M. J. and Robbins, P.W. (1991). Chitinase is required for cell separation during growth of *Saccharomyces cerevisiae*. *J. Biol. Chem.* 266, 19758-19767.
- Neu H.C. (1992). The crisis in antibiotic resistance. *Science.* 257, 1064-1072.
- Laskowski, R.A., Hutchinson, E.G., Michie, A.D., Wallace, A.C., Jones, M.L., Thornton, J.M. (1997). PDBsum: a Web-based database of summaries and analyses of all PDB structures. *Trends. Biochem. Sci.* 22, 488-490.
- Lee, C.G. (2009). Chitin, Chitinases and Chitinase-like Proteins in Allergic Inflammation and Tissue Remodeling. *Yonsei Med J.* 50, 22-30.
- Lipinski, C.A., Lombardo, F., Dominy, B.W., Feeney, P.J. (2001). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug. Deliv. Rev.* 46, 3-26.
- Lipinski, C. A. (2004). Lead and drug like compounds: the rule of five revolution. *Drug Discov Today Technol.* 1, 337-341.
- MacKerell, A.D., Bashford, D., Bellott, M., Dunbrack, R.L., Evanseck, J.D., Field, M.J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F.T.K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D.T., Prodhom, B., Reiher, W.E., Roux, B., Schlenkrich, M., Smith, J.C., Stote, R., Straub, J., Watanabe, M., Wiórkiewicz-Kuczera, J., Yin, D., and M. Karplus, M.(1998). All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B.* 102, 3586-3616.
- Mizoguchi, E. (2006). Chitinase 3-like-1 exacerbates intestinal inflammation by enhancing bacterial adhesion and invasion in colonic epithelial cells. *Gastroenterol.* 130, 398-411.
- Myles, O., Wortmann, G.W., Cummings, J.F., Barthel, R.V., Patel, S., Crum-Cianflone, N.F., et al. (2007). Visceral leishmaniasis: clinical observations in 4 US army soldiers deployed to Afghanistan or Iraq, 2002-2004. *Arch. Intern. Med.* 167, 1899-1901.
- Olland, A.M. et al. (2009). Triad of polar residues implicated in pH specificity of acidic mammalian chitinase. *Protein Sci.* 18, 569-578.
- Omura, S., Arai, N., Yamaguchi, Y., Masuma, R., Iwai, Y., Namikoshi, M., Turberg, A., Kolb, H., and Shiomi, K. (2000). Argifin, a new chitinase inhibitor, produced by *gliocladium* sp. FTD-0668. taxonomy, fermentation and biological activities. *J. Antibiot. (Tokyo)*. 53, 603-608.
- Papanikolau, Y., Tavlas, G., Vorgias, C. E. and Petratos, K. (2003). De novo purification scheme and crystallization conditions yield high-resolution structures of chitinase A and its complex with the inhibitor allosamidin. *Acta. Crystallogr., Sect. D: Biol. Crystallogr.* 59, 400-403.
- Pontius, J., Richelle, J., Wodak, S.J. (1996). Deviations from standard atomic volumes as a quality measure for protein crystal structures. *J. Mol. Biol.* 264, 121-136.
- Rao, F. V., Houston, D. R., Boot, R. G., Aerts, J., Sakuda, S. and Van Aalten, D. M. F. (2003). Crystal structures of allosamidin derivatives in complex with human macrophage chitinase. *J. Biol. Chem.* 278, 20110-20116.
- Rao, F. V., Andersen, O. A., Vora, J. A., Demartino, J. A. and VanAalten, D.M. F. (2005). Methylxanthine drugs are chitinase inhibitors: investigation of inhibition and binding modes. *Chem. Biol.* 12, 973-980.
- Renkema, G. H., Boot, R. G., Muijsers, A. O., Donker-Koopman, W. E., and Aerts, J. M. F. G. (1995). Purification and characterization of human chitotriosidase, a novel member of the chitinase family of proteins. *J. Biol. Chem.* 270, 2198-2202.
- Rungra, D., Chauhan, N., and Mukherjee, K. (2014). Identification of Hiv1 Protease Inhibitor through Molecular Modelling and Structure Based Virtual Screening Approach *J. Adv. Bioinform. Appl. Res.* 5, 140-149.
- Sakuda, S., Isogai, A., Matsumoto, S., Suzuki, A. and Koseki, K. (1986), The structure of allosamidin, a novel insect chitinase inhibitor, produced by *Streptomyces* sp. *Tetrahedron Lett.*, 27, 2475-2478.
- Sakuda, S., Isogai, A., Matsumoto, S. and Suzuki, A. (1987). Search for microbial insect growth regulators. II. Allosamidin, a novel insect chitinase inhibitor. *J. Antibiot.* 40, 296-300.
- Sakuda, S., Nishimoto, Y., Ohi, M., Watanabe, M., Takayama, S., Isogai, A. and Yamada, Y. (1990). Effects of demethylallosamidin, a potent yeast chitinase inhibitor, on the cell division of yeast. *Agric. Biol. Chem.* 54, 1333-1335.
- Sali, A. and Blundell, T.L. (1993). Comparative protein modeling by satisfaction of spatial restraints. *J. Mol. Biol.* 234, 283-291.

- Shahabuddin, M., Toyoshima, T., Aikawa, M. and Kaslow, D.C. (1993). Transmission-blocking activity of a chitinase inhibitor and activation of malarial parasite chitinase by mosquito protease. Proc. Natl. Acad. Sci. USA. 90, 4266–4270.
- Sivakumar, K., Balaji, S. and Radhakrishnan, G. (2007). *In silico* characterization of antifreeze proteins using computational tools and servers. J. Chem. Sci. 119, 571–579.
- Tabudravu, J. N., Eijsink , V. G. H., Gooday, G. W., Jaspars, M., Komander, D., Legg, M., Synstad, B. and VanAalten, D. M. F. (2002). Psammaplin A, a Chitinase Inhibitor Isolated from the Fijian Marine Sponge *Aplysinella Rhax*. Bioorg. Med. Chem. 10, 1123–1128.
- Wilkinson A.J., Fersht A.R., Blow D.M., Winter G. (1983). Site-directed mutagenesis as a probe of enzyme structure and catalysis: tyrosyl-tRNA synthetase cysteine-35 to glycine-35 mutation. Biochemistry 22, 3581–3586.
- Zakariassen, H., Hansen, M.C., Jørnli, M., Eijsink, V.G.H. and Sorlie, M. (2011). Mutational effects on transglycosylating activity of family 18 chitinases and construction of a hypertransglycosylating mutant. Biochemistry 50, 5693–5703.
- Zhu, Z., Zheng, T., Homer, R.J., Kim, Y.K., Chen, N.Y., Cohn, L., Hamid, Q. and Elias, J.A. (2004). Acidic mammalian chitinase in asthmatic Th2 inflammation and IL-13 pathway activation. Science. 304, 1678–1682.