GLOBAL JOURNAL OF BIO-SCIENCE AND BIOTECHNOLOGY

www.scienceandnature.org

HOMOLOGY MODELING OF FOOD ENZYMES LIKE PROTEASE, CELLULASE AND PECTINASE

Megha, S.V., Maragathavalli, S., Brindha, S. and Annadurai, B.
Research and Development Centre, Bharathiar University, Coimbatore

ABSTRACT
In this study the homology modeling of food enzyme like protease, cellulase and pectinase from three organisms *Pseudomonas aeruginosa, Ceratocystis paradoxa, and Alternaria cepulae* respectively were conducted. The three dimensional structure of protein help us to understand protein – protein interactions. Multiple sequence analysis was carried out using SWISS PROT and possible models of each enzyme were constructed. The final validation and stability of protein three dimensional structures which was constructed using homology modeling was checked by RAMPAGE

KEYWORDS: SWISS PROT, RAMPAGE.

INTRODUCTION

Homology modeling can be defined as the construction of protein models from its amino acid sequence. Homology modeling relies on the identification of one or more protein structure which is likely resembles the structure of the query sequence. Evolutionarily related proteins have similar sequences and naturally occurring homologous protein have similar protein structures (Kaczanowshi *et al*., 2010). The sequence alignment and template structure are used to produce a structural model of the target. The quality of homology model depends on the quality of the sequence alignment and the template structure. Homology modeling can produce high quality structural models when the target and template are closely related (Williamson, 2000). The protein with undetermined structure is called "model", "unknown" or "sequence" protein. The protein with known three dimensional structures is /is referred as "reference" or "real" proteins.

The three dimensional structure of a protein help us to understand protein –protein interaction or protein – DNA interaction. Primary and structural information of proteins are stored in different places like SWISS-PROT, PDB *etc*. The three dimensional shape of an entire protein is the three dimensional structure. The protein molecule will bend twist to attain maximum stability or lowest enzyme state. The three dimensional structure of protein is attained by the force due to bonding interaction between the side chain group of amino acid. Under physiological condition the hydrophobic side chain of neutral, non-polar amino acids such as phenyl alanine or isoleucin tend to be buried on the interior of protein molecule thereby shielding them from aqueous medium. The alkaline groups of alanine, valine, leucine, isoleucine often form hydrophilic interaction between one another while aromatic groups like phenylalanine and thyrosine often stack together. Acidic and alkaline amino acids will be generally be exposed on the surface of protein as they are hydrophilic. Formation of S-S bond by cysteine is an important aspect of the stabilization of protein three dimensional structure allowing different parts of the protein chain to be held together covalently. Salt bridge, ionic interaction between positively and negatively charged sites on amino acid side chain also help to stabilize three dimensional structure of protein. Proteases execute a large variety of functions and have important biotechnological applications. Proteases represent one of the three largest groups of industrial enzymes and find application in detergents, leather industry, food industry, pharmaceutical industry and bioremediation processes. For an enzyme to be used as an detergent additive it should be stable and active in the presence of typical detergent ingredients, such as surfactants, builders, bleaching agents, bleach activators, fillers, fabric softeners and various other formulation aids. Since protease is one of the widely used enzyme in different industries, its production from *Pseudomonas sp* using milk as the source of protein was carried out and the produced enzyme was purified. Cellulase refers to a group of enzymes which, acting together, hydrolyze cellulose. Cellulose is a linear polysaccharide of glucose residues connected by β-1, 4 linkages. In nature cellulose is usually associated with other polysaccharides such as xylan or lignin. It is the skeletal basis of plant cell walls. According to the research (Spano *et al.,* 1975) cellulose is the most abundant organic source of food, fuel and chemicals. However, its usefulness is dependent upon its hydrolysis to glucose. Acid and high temperature degradation is unsatisfactory in that the resulting sugars are decomposed; also, waste cellulose contains impurities that generate unwanted byproducts under these harsh conditions. Cellulase is a group of enzymes that catalyses cellulolysis. It is mainly produced by fungi, bacteria and some protozoans. It is studies extensively due to their applications in the hydrolysis of cellulose, the most abundant biopolymer and potential source of utilizable sugar, which serves as a raw material in the production of chemicals and fuel (Ali *et al*., 2011, Pradeep *et al*., 2012). Since, Cellulases is used mostly in textiles, food and the bioconversion lignocellulosic waste to alcohol, it becomes industrially important. Because largely is used in the industries, large scale of production from the microbial

strains was done. The enzyme production, isolation and purification were conducted.

Pectinase are enzymes that breakdown pectin, a polysaccharide substrate that is found in the cell wall of plants. Pectinases are general name of pectic enzymes which include pectolyase, pectozymes and poly galacturonase. Pectins are jelly like matrix which cements plant cells together in a cell wall. Pectinase is widely used in fruit juice extraction, wine production, paper and pulp industry, textile processing, waste water treatment; animal feed purification of plant viruses. Pectinase is a growing enzyme of biotechnology sector, showing gradual increase of need in market (Garg *et al*., 2016).The pectinase enzyme was isolated and purified from the crude extract of *Alternaria cepulae*. The purified sample was then subjected to bioinformatic studies of homology modeling to detect the structure of the purified sample.

In addition to that the computational tools and insilico studies are required to preserve and reduce the cost of protease, cellulase and pectinase. Bioinformatics revolutionized the field of molecular biology. The raw sequences information of proteins and nucleic acid can convert to analytical and relative information with the help of soft computing tools. Prediction of protein function is important application of bioinformatics (Prasanth *et al*., 2010). In this chapter Bioinformatics analysis characterization of protease, cellulases and pectinase from *Pseudomonas aeruginosa*, *Cerotocystis paradoxa* and *Alternaria cepulae* respectively were carried out. Hence, the *insilico* studies by using ExPASy, ProtParam tool are used for determining molecular weight. The secondary structure prediction has to be done by SOPMA and GOR IV to show the random coil. Multiple sequence analysis protease, cellulases, and pectinase have to be carried out using SWISS PROT and the possible models of each enzyme were constructed. The final validation and stability of protein three dimensional structure constructed using homology modeling was checked using RAMPAGE. These parameters will help the biochemist and physiologists in extraction, purification, separation and industrial important enzymes.

SWISS MODEL

SWISS MODEL is the field of automated modeling service on the internet. With the help of visualization tool Swiss PDB Viewer, Internet based workspace and SWISS MODEL Respiratory it provides a fully integrated sequence to structure analysis and modeling platform. It provides a computational environment to the scientific community to developing structural information. Over the last decade, the availability of structural information has significantly increased for many organisms thus providing protein modeling and experimental structure determination available.

The SWISS MODEL template library provides annotation of quaternary structure and essential ligand and cofactors to allow for building of complete structural models including their oligometric structures. The improved SWISS MODEL makes extensive use of model quality estimation for selection of suitable template and provides estimates of the expected accuracy of the resulting model. The new website allows user to interactively search for templates, cluster them by sequence similarity, structurally compare alternative template and select the ones to be used for model building, SWISS MODEL is available at www. swissmodel.expasy.org.

In building a homology model comprises four main steps (i) identification of structural templates, (ii) alignment of target sequence and template structure, (iii) model building, (iv) model quality evaluation.

Template search and selection

The SWISS-MODEL Template Library is searched in parallel both with BLAST (Altschul *et al*., 1997) and HHblits (Remmert *et al*., 2012) to identify templates and to obtain target-template alignments. The combined usage of these two methods guarantees good alignments at high and low sequence identity levels (Sadowski and Jones, 2007). In order to select the most suitable templates, the procedure implemented in SWISS-MODEL uses properties of the target-template alignment. Each of the alignment properties is modelled as probability density function (PDF) of the estimate for a resulting model having a certain structural similarity to the target. The use of PDFs has the advantage of at once including the expectation value as well as the accuracy of the estimate for each property. It also takes into account, that some properties are better (more accurate) at predicting the quality at high levels of sequence identity, whereas others are more accurate in the twilight zone of sequence alignments. When combining the estimates of each property, the most likely structural similarity is the value at which the joint distribution is maximized, termed the GMQE (global quality estimation score).

Display of template identification results

The Template results page serves both as an overview of available templates as well as an interactive template selection tool. The top part of the screen contains a summary of the top-ranking templates identified by the template search methods. Three types of views are available: (i) a templates summary table, listing all templates in tabular form and providing an overview of relevant attributes of each template, (ii) an interactive chart showing the templates in relation to each other in Sequence Similarity space and (iii) the sequence Alignment of Selected Templates.

Templates can be selected in any of these views for the subsequent modeling step. Selected templates are automatically shown in the 3D viewer. If multiple templates are selected their structural superposition is shown, allowing instant visualization of structural differences between them. The complete list of all identified templates can be accessed at the bottom of the template results page.

In the Templates summary table, template annotations and target-template alignments can be retrieved by clicking on the arrows at the right end of the table rows to expand the box with the description of the individual template.

GMQE

Global Model Quality Estimation is a quality estimation which combines properties from the target template alignment and the template search method. The result of GMQE score is expressed as a number between 0-1, reflecting the expected accuracy of a model built with that alignment and template. Higher numbers indicate higher reliability. Once a model is built the GMQE gets update

for this specific case by also taking into account the Q Mean score of the obtained model in order to increase reliability of the quality estimation.

Q Mean

Q Mean is a composite scoring function based on different geometrical properties and both global (entire structure) and local (pre residue) absolute quality estimates on the basis of one single model (Benkert, 2009). The Q Mean Z score provides an estimate of the degree of nativeness of the structural features observed in the model and indicates whether the model is of comparable quality to experimental purpose. Higher Q Mean- Z score indicates better agreement between the model structure and experimental structure .Scores of -4.0 and below are an indication of model with very low quality; this is also highlighted by a change of the "thumbs up" symbol to a "thumbs down" symbol next to the score.

Model building and scoring

After templates are selected for model building, either by using the automated or manual selection mode we can built the structure of the query sequence by using the workspace. The SWISS-MODEL Workspace (Biasini *et al*., 2014) is a personal web-based working environment where several modeling projects can be carried out in parallel. Protein sequence and structure databases necessary for modeling are accessible from the workspace and are updated in regular intervals. Tools for template selection, model building, and structure quality evaluation can be invoked from within the workspace.

Models are computed by the SWISS-MODEL server homology modeling pipeline and which relies on ProMod3, an in house comparative modeling engine based on Open Structure. Given a target-template alignment and a template structure, all conserved structural information is transferred to a model corresponding to the target sequence. The final model is selected based on model quality estimation employing statistical potentials of mean force.

Local Quality

The local quality for each residue of the model on (X axis) the expected similarity to the native structure (Y axis).The residues showing a score below 0.6 are expected to be of low quality

Comparison Plot

In comparison plot the model quality scores of individual model are expressed as Zscore in comparison to scores obtained for high resolution crystal structure. The X axis shows the length (amino acid) of the protein. The Y axis is the normalized Q Mean score. Every dot represents a protein structure. The darkest dots are all structure with a global Q Mean Z score between -1 and 1, structures with a Z score between 1 and 2 are grey and if the Z score is more than 2 they are in a light grey colour. The red star represents the model.

Ramachandran Plot

Two torsion angles in the polypeptide chain, also called Ramachandran angles (Ramachandran *et al*., 1963)
describe the rotations of the polypeptide backbone around the bonds between N-Cα (called Phi, φ) and Cα-C (called Psi, ψ). By plotting the φ values on the x-axis and the ψ values on the y-axis ie: plotting the torsional angles graphically shows which combination of angles is possible. The torsional angles of each residue in a peptide define the geometry of its attachment to its two adjacent residues by positioning its planar peptide bond relative to the two adjacent planar peptide bonds; thereby the torsional angles determine the conformation of the residues and the peptide. Many of the angle combinations, and therefore the conformations of residues, are not possible because of steric hindrance. By making a Ramachandran plot, protein structural scientists can determine which torsional angles are permitted and can obtain insight into the structure of peptides.

SYSTEM AND METHODS

SWISS MODEL

Template search and selection

In order to select the most suitable templates, the procedure implemented in SWISS-MODEL uses properties of the target-template alignment. First visit the home page of SWISS MODEL, then in the home page click "start modeling".

Method

To get the Alignment of selected template, first visit the home page of SWISS MODEL, then in the home page click "start modeling". The procedure was explained in general system and methods.

Model building

After templates are selected for model building, either by using the automated or manual selection mode we can built the structure of the query sequence by using the workspace. Protein sequence and structure databases necessary for modeling are accessible from the workspace and are updated in regular intervals. Tools for template selection, model building, and structure quality evaluation can be invoked from within the workspace.

Method

To get the models of the template sequence we have to visit the homepage of SWISS MODEL. The procedure was explained in general system and methods.

RAM PAGE

RAMPAGE helps in the construction of Ramachandran plot. It is important in the conformation of protein and the protein structure. It provide the distribution of conformational angles in proteins, the areas in the plot help us to understand the different torsional angles between planes.the rotation of torsional angle in 180^0 causes the side groups to come closer and results in steric hindrance which results in the structure unstable. Therefore this plot gives us to understand, what are the torsional angles possible for the construction of stable protein molecule.

Method

After construction of model, its stability is checked using Ramachandran plot. We can get the Ramachandran Plot from the site RAMPAGE. The procedure was explained in system and methods.

RESULTS

Protease

FIGURE 1: Sequence alignment of the template sequence of protease

Inference: The sequence alignment was conducted using SWISS MODEL. Fig. 1 shows the list of sequences which have more similarity with the template sequence. From the list the sequence with 42.92% identity with the template was selected for further studies.

Cellulase

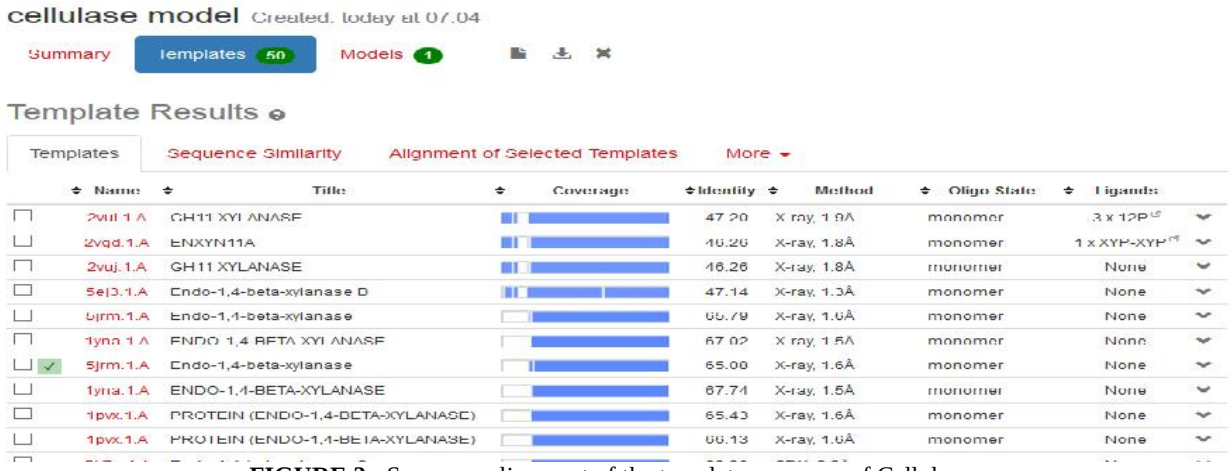


FIGURE 2: Sequence alignment of the template sequence of Cellulase

Inference

The sequence alignment was conducted using SWISS MODEL. Fig. 2 shows the list of sequences which have more similarity with the template sequence. From the list the sequence with 65.08% identity with the template was selected for further studies

Pectinase

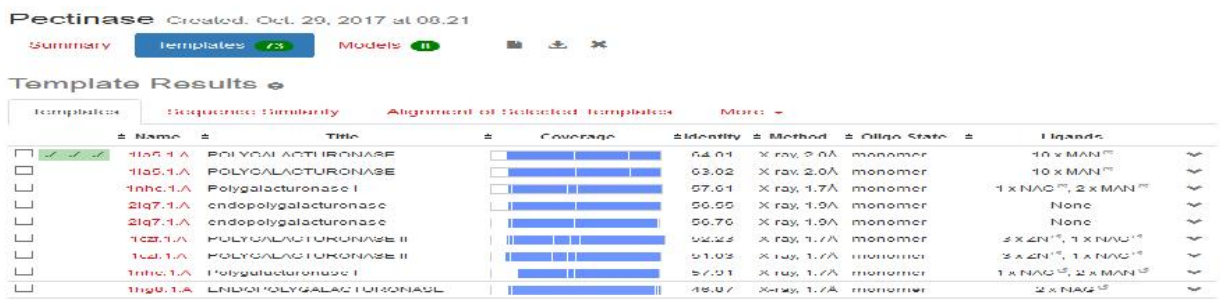


FIGURE 3: Sequence alignment of the template sequence of Pectinase

Inference

The sequence alignment was conducted using SWISS MODEL. Fig. 3 shows the list of sequences which have more similarity with the template sequence. From the list the sequence with 64.01% identity with the template was selected for further studies

Modelling of Protease

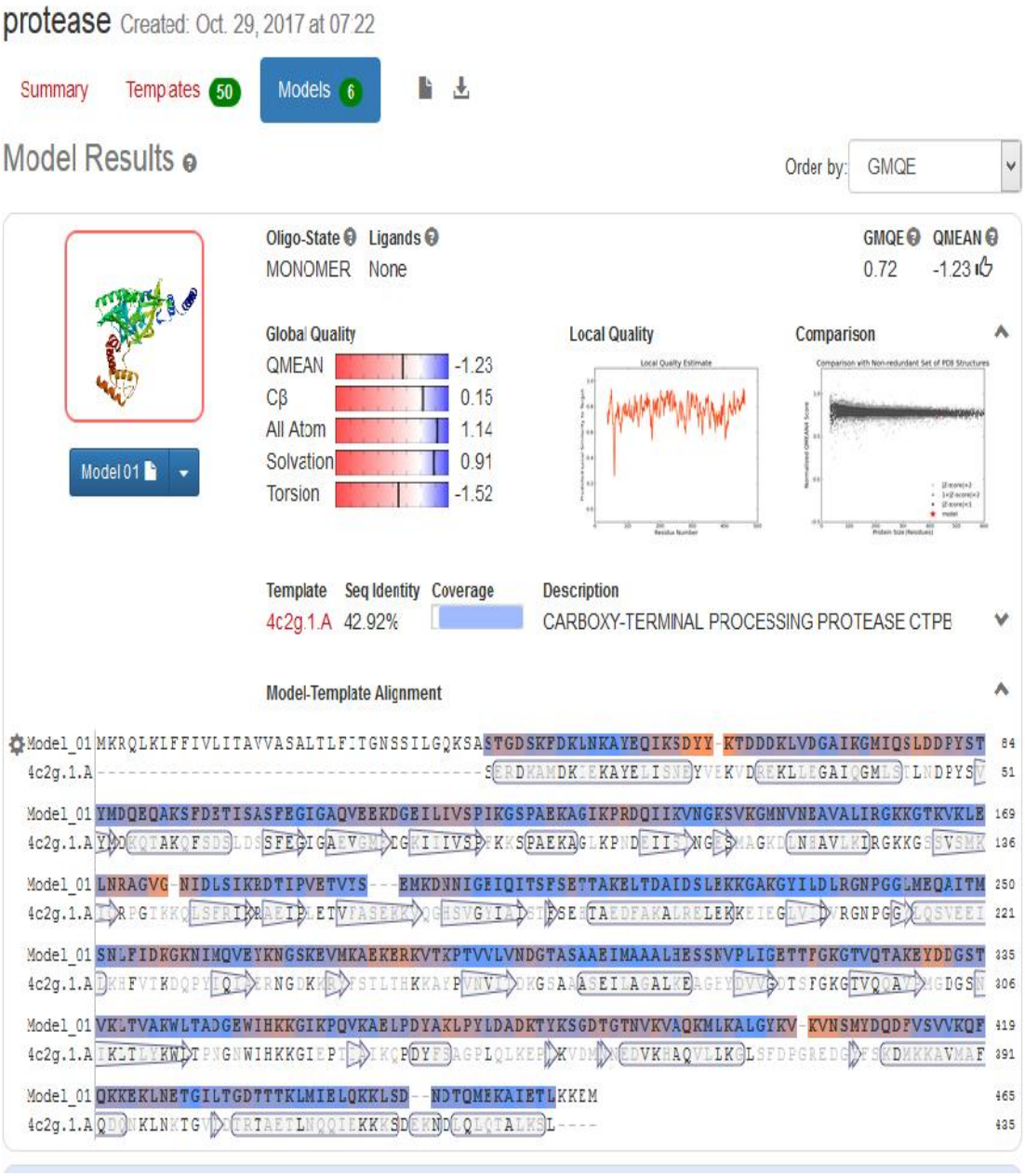


FIGURE 4: Modeling of Protease

Inference: The protein models are constructed using SWISS MODEL. In fig. 4 a model was constructed using the template sequence. The GMQE was found to be 0.72 which is between 0-1 and the value is higher in number which indicates the model has high reliability. The QMEAN value was found to be -1.23 which is greater than -4 and its shown with a "thumbsup" sign indicating that the model constructed is a good quality. The local quality plot for each residue of the model (X axis) the expected similarity to the native structure (Y axis). The residue showing a score below 0.6 are expected to be of low quality. In comparison plot, every dot represents a protein structure. The darkest dots are all structure with a global Q Mean Z score between -1 and 1, structures with a Z score between 1 and 2 are grey and if the Z score is more than 2 they are in a light grey in colour. The red star represents the model.

Modelling of Cellulase

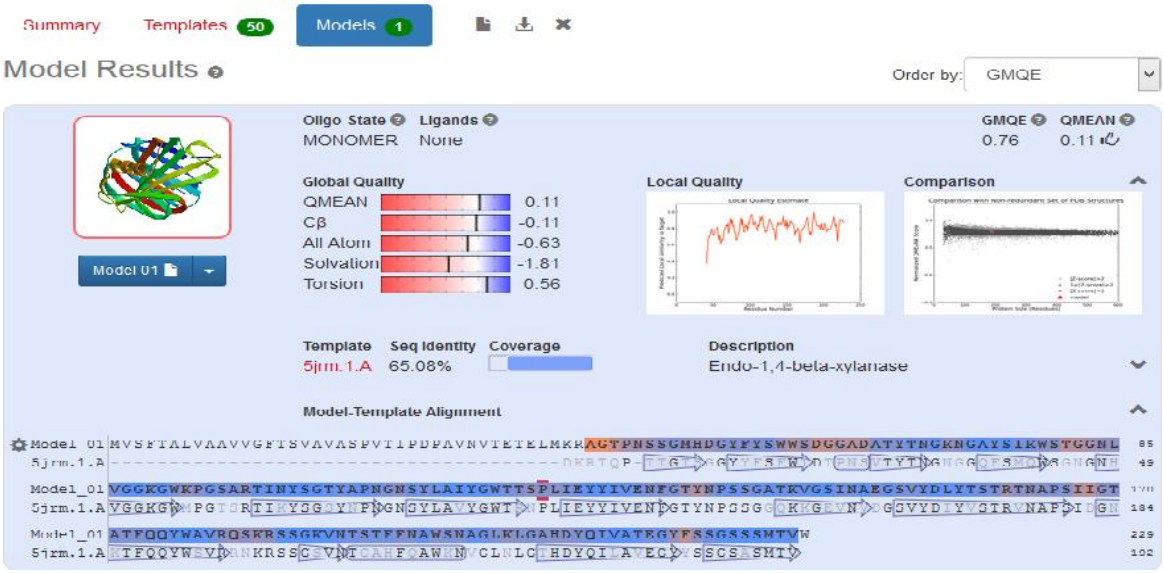


FIGURE 5: Modelling of Cellulase

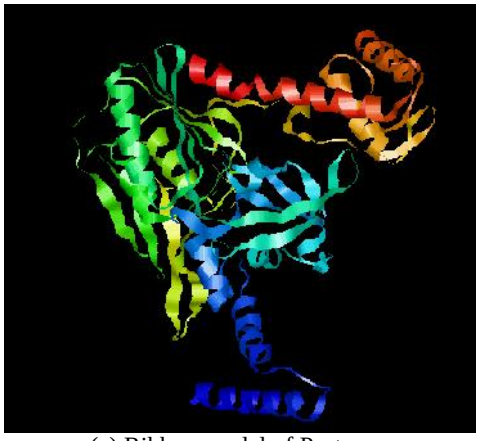
Inference: The protein models are constructed using SWISS MODEL. In fig. 5 a model was constructed using the template sequence. The GMQE was found to be 0.76 which is between 0-1 and the value is higher in number which indicates the model has high reliability. The QMEAN value was found to be 0.11 which is greater than -4 and it's shown with a "thumbs up" sign indicating that the model constructed is a good quality. The local quality plots for each residue of the model (X axis) the expected similarity to the native structure(Y axis). The residue showing a score below 0.6 are expected to be of low quality. In comparison plot, every dot represents a protein structure. The darkest dots are all structure with a global QMean Z score between -1 and 1, structures with a Z score between 1 and 2 are grey and if the Z score is more than 2 they are in a light grey in colour. The red star represents the model.

Modelling of Pectinase

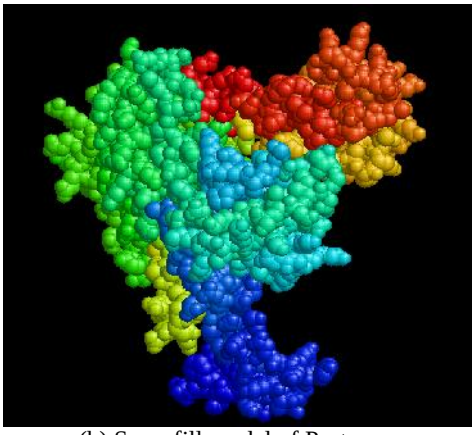
FIGURE 6: Modelling of Pectinase

Inference: The protein models are constructed using SWISS MODEL. In fig. 6 a model was constructed using the template sequence. The GMQE was found to be 0.79 which is between 0-1 and the value is higher in number which indicates the model has high reliability. The QMEAN value was found to be 1.34 which is greater than -4 and its shown with a "thumbs up" sign indicating that the model constructed is a good quality. The local qualityplot for each residue of the model (X axis) the expected similarity to the native structure (Y axis). The residue showing a score below 0.6 are expected to be of low quality. In comparison plot, every dot represents a protein structure. The darkest dots are all structure with a global QMean Z score between -1 and 1, structures with a Z score between 1 and 2 are grey and if the Z score is more than 2 they are in a light grey in colour. The red star represents the model.

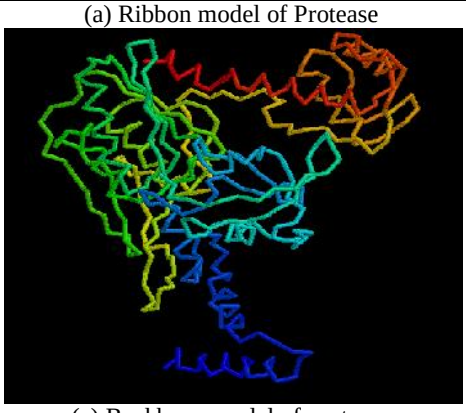
Protease Models



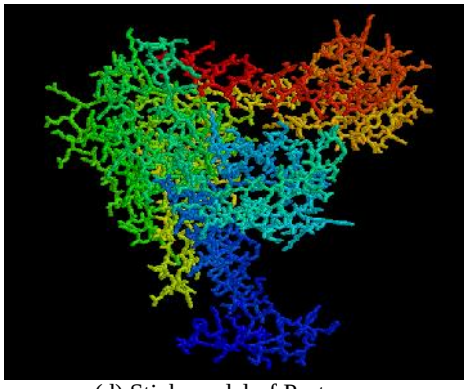
(a) Ribbon model of Protease



(b) Spacefill model of Protease



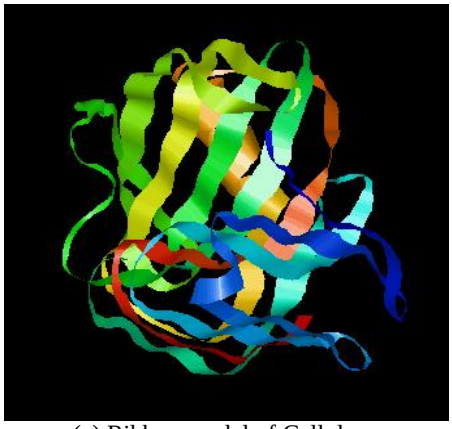
(c) Backbone model of protease



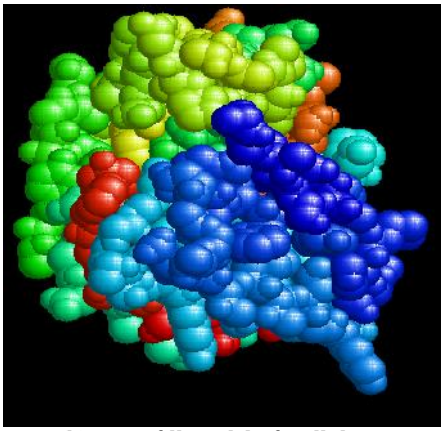
(d) Stick model of Protease

FIGURE 7(a) Ribbon model of Protease (b) Space fill model of Protease (c) Backbone model of protease (d) Stick model of Protease

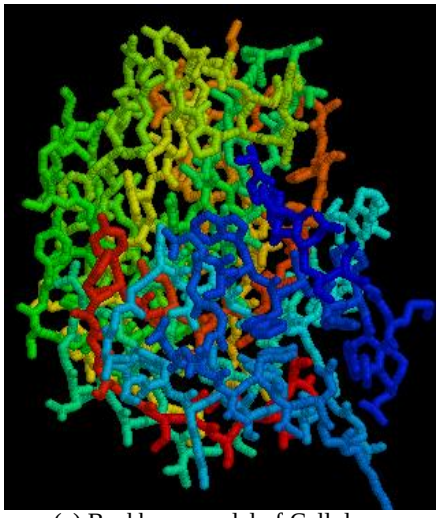
Cellulase Models



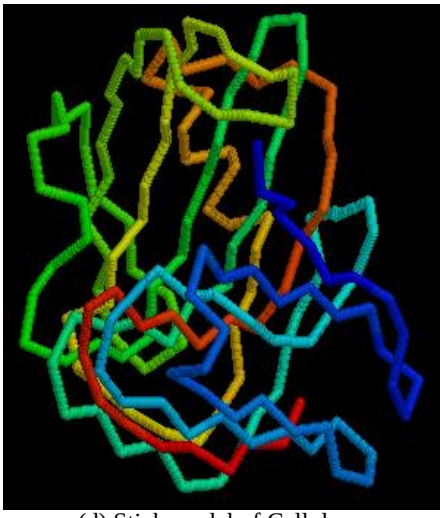
(a) Ribbon model of Cellulase



(b) Spacefill model of Cellulase



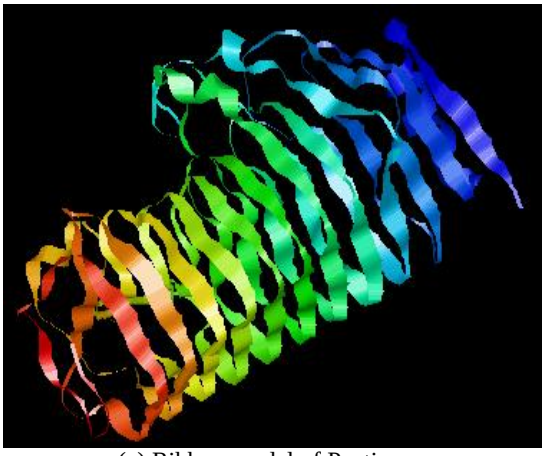
(c) Backbone model of Cellulase



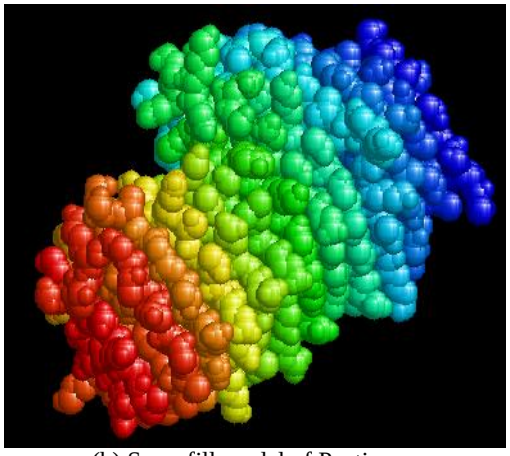
(d) Stick model of Cellulase

FIGURE 8(a) Ribbon model of Cellulase (b) Spacefill model of Cellulase (c) Backbone model of Cellulase (d) Stick model of Cellulase

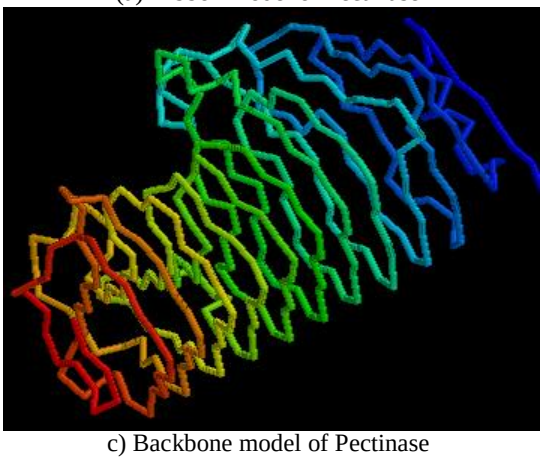
Pectinase Models



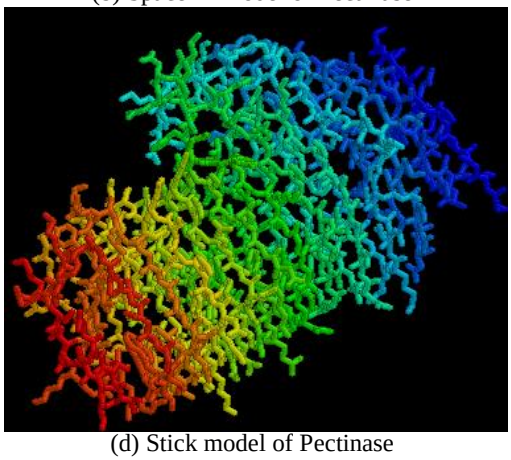
(a) Ribbon model of Pectinase



(b) Spacefill model of Pectinase



c) Backbone model of Pectinase

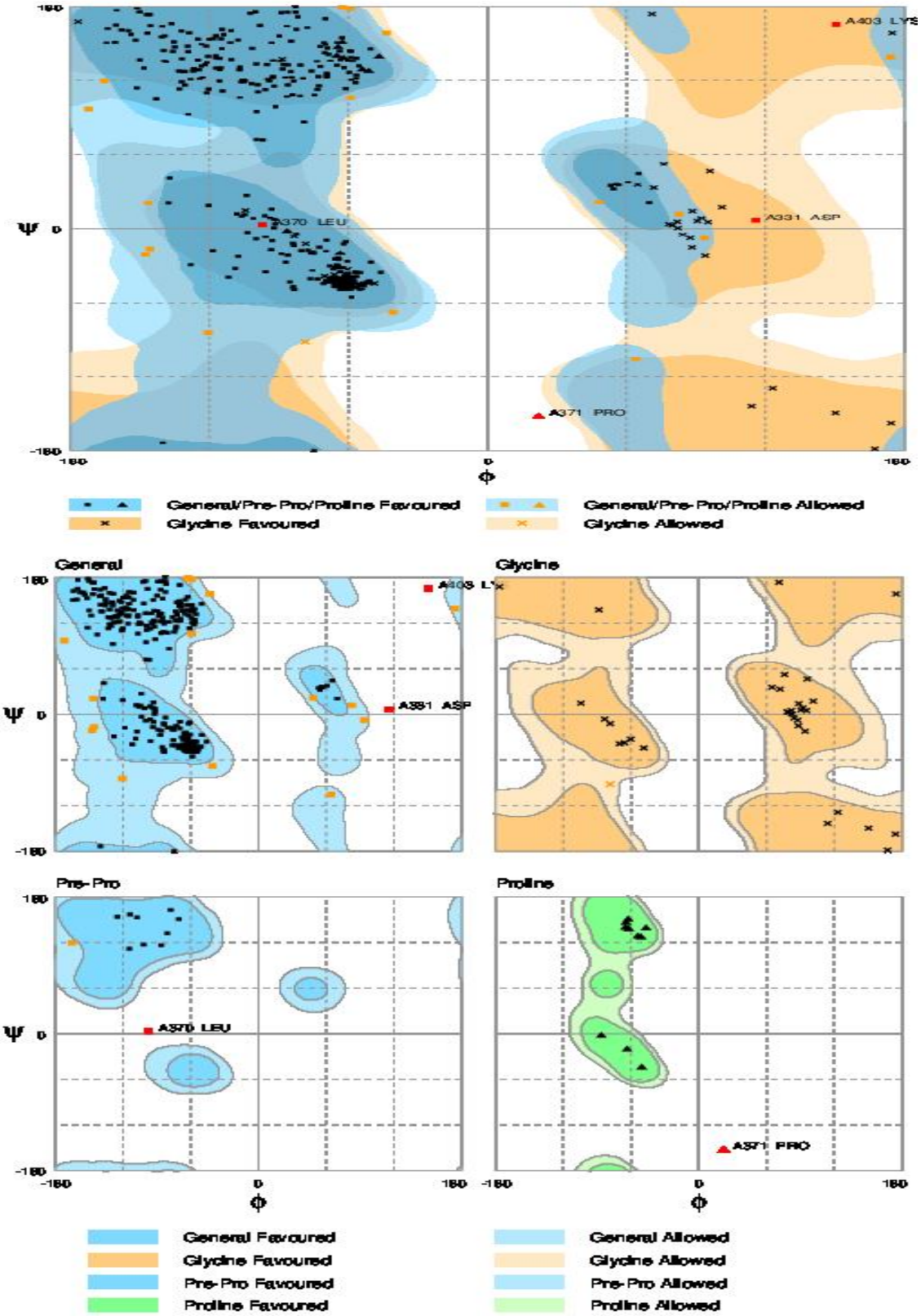


(d) Stick model of Pectinase

FIGURE 9 (a) Ribbon model of Pectinase (b) Spacefill model of Pectinase (c) Backbone model of Pectinase (d) Stick model of Pectinase

RAMPAGE

Protease

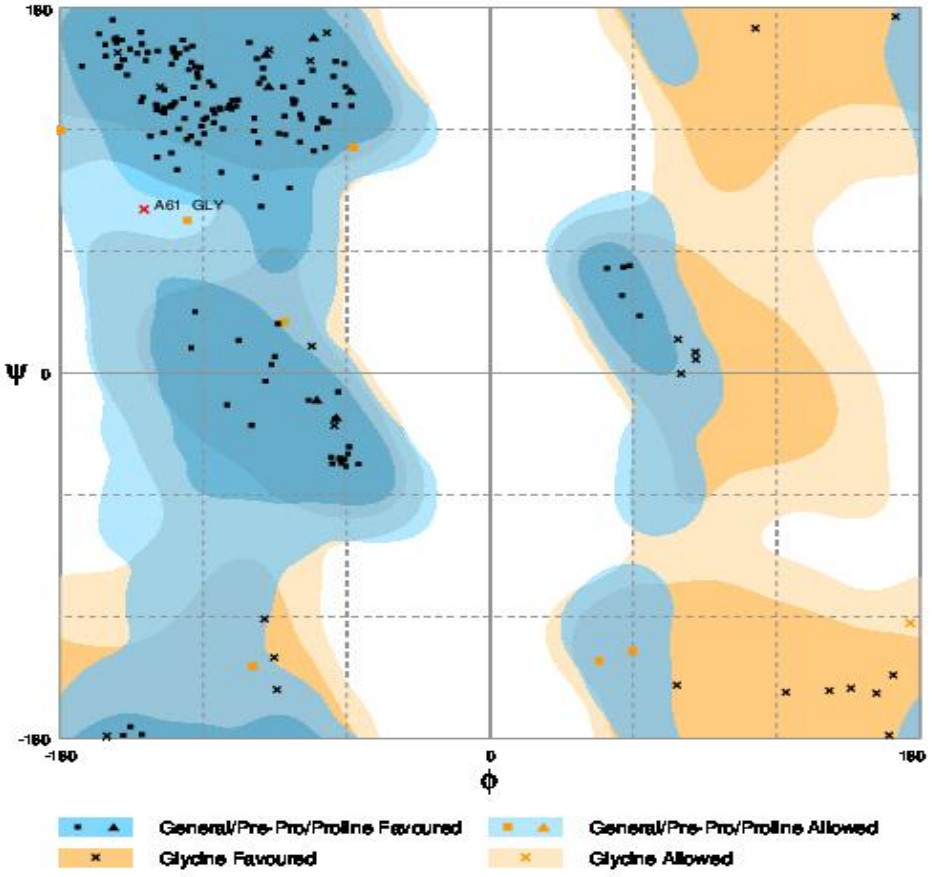


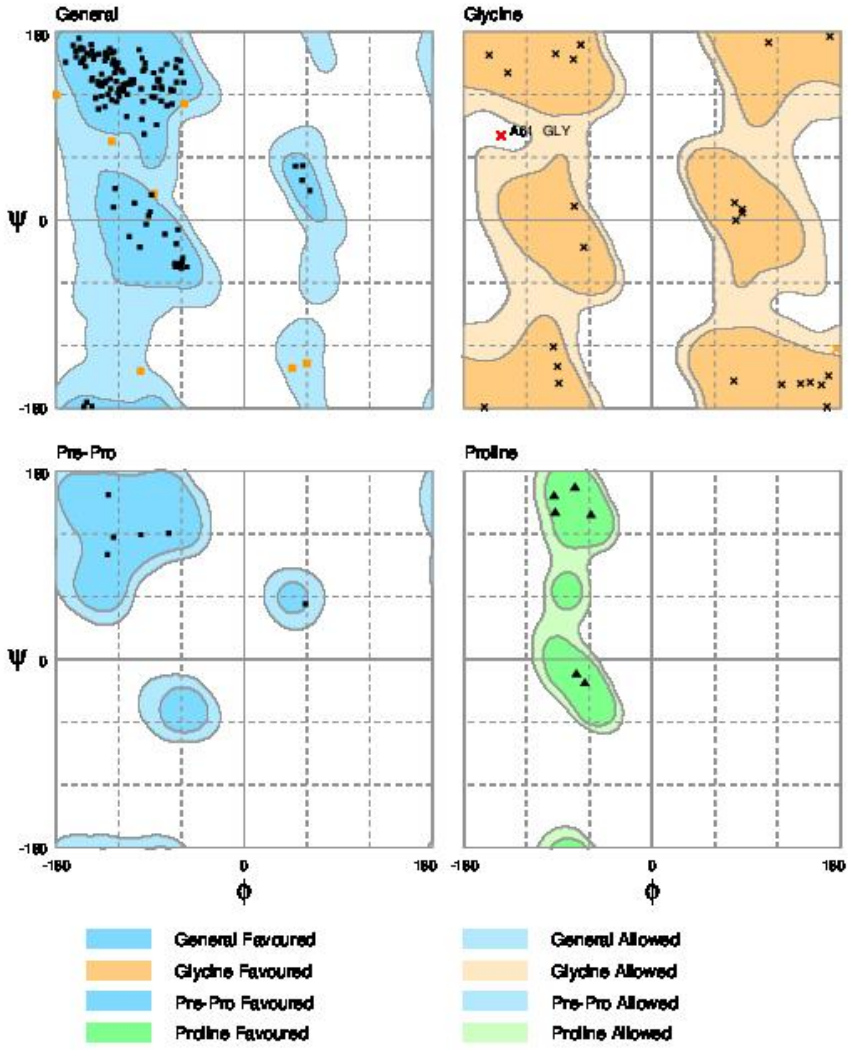
Evaluation of residues

```
Residue [A  59 :TYR] ( -40.72, -67.74) in Allowed region
Residue [A  60 :LYS] ( 174.06, 139.11) in Allowed region
Residue [A  87 :ASP] ( -59.09, 177.99) in Allowed region
Residue [A 120 :SER] (-165.23, 119.54) in Allowed region
Residue [A 135 :ARG] (  82.78,  11.92) in Allowed region
Residue [A 147 :LYS] ( -59.01, 106.10) in Allowed region
Residue [A 176 :GLY] ( -78.48, -91.83) in Allowed region
Residue [A 198 :ASP] (  48.35,  21.57) in Allowed region
Residue [A 200 :ASN] (-145.64, -16.32) in Allowed region
Residue [A 281 :LYS] ( -43.48, 158.44) in Allowed region
Residue [A 296 :SER] (  64.15,-105.39) in Allowed region
Residue [A 345 :THR] ( -62.51, 179.39) in Allowed region
Residue [A 354 :LYS] (-147.60, -21.00) in Allowed region
Residue [A 402 :VAL] (-120.52, -84.14) in Allowed region
Residue [A 404 :VAL] (-146.39,  20.86) in Allowed region
Residue [A 406 :SER] (  93.84,  -7.51) in Allowed region
Residue [A 451 :ASP] (-172.06,  97.11) in Allowed region
Residue [A 331 :ASP] ( 115.91,   6.63) in Outlier region
Residue [A 370 :LEU] ( -97.15,   3.07) in Outlier region
Residue [A 371 :PRO] (  22.06,-151.71) in Outlier region
Residue [A 403 :LYS] ( 150.62, 165.69) in Outlier region
Number of residues in favoured region    (~98.0% expected)   :  401 ( 95.0%)
Number of residues in allowed region     ( ~2.0% expected)   :   17 (  4.0%)
Number of residues in outlier region                         :    4 (  0.9%)
```

FIGURE 10: RAMPAGE of Protease

Inference: To model constructed for protease has to be checked for the quality of the structure. For this Ramachandran plot was constructed Fig. 10 shows the result of the protease sequence submitted for constructing the ramachandran plot. The blue coloured region on the graph shows the general favoured region and the region with blue and orange dots showed the allowed region for **Cellulase** the amino acids for the formation of structure. Orange shaded region shows the favoured and allowed region of Glycine amino acid. Green coloured region on the plot shows the favoured and allowed region for Proline amino acid. Evaluation of the residues shows the allowed region of each amino acid. About 98% of residues are in the allowed region, it shows that the model is acceptable.





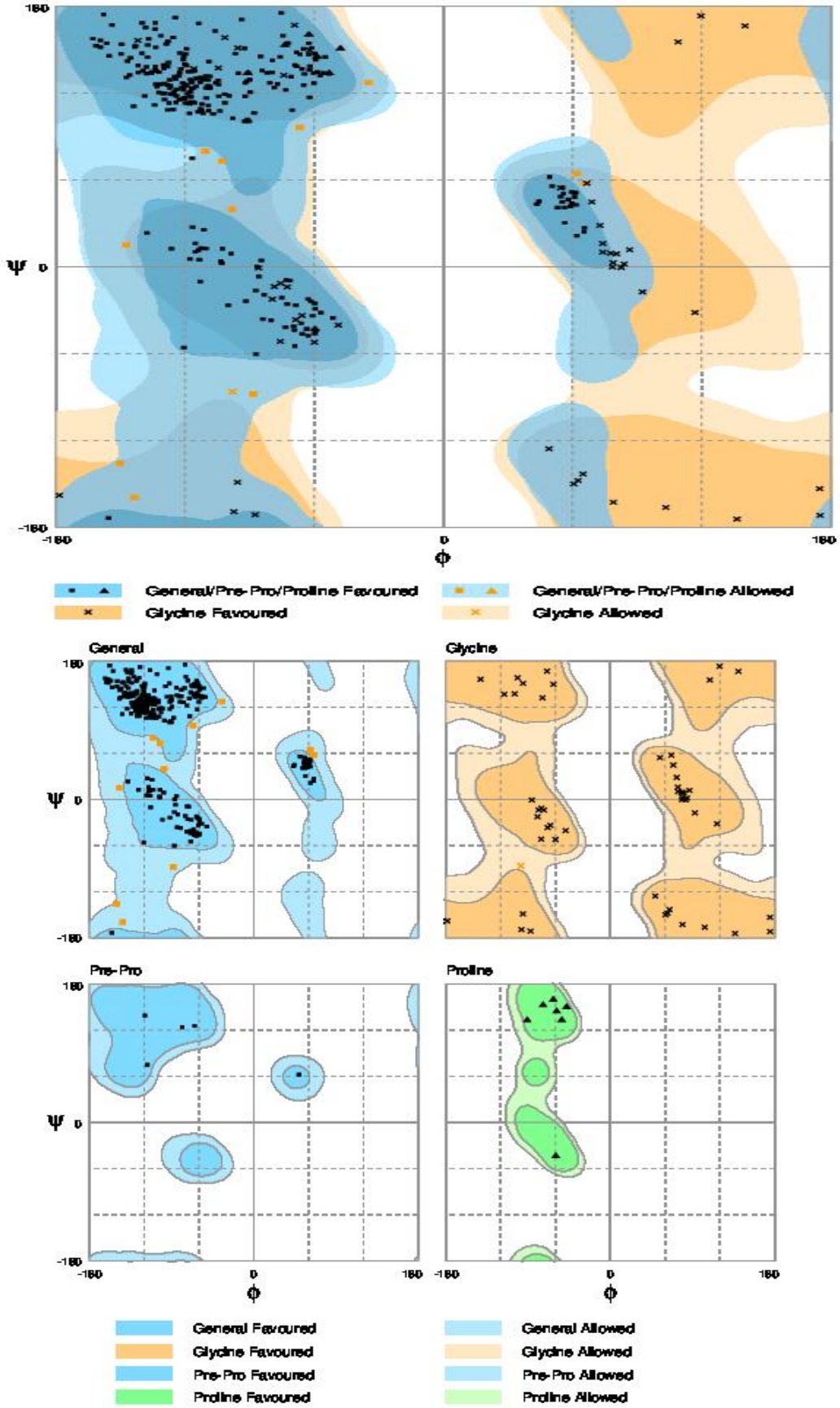
Evaluation of residues

```
Residue [A  72 :ASN] (  45.84,-141.68) in Allowed region
Residue [A  81 :THR] (  59.80,-136.94) in Allowed region
Residue [A  90 :GLY] ( 176.02,-123.12) in Allowed region
Residue [A  96 :ALA] ( -86.33,  25.07) in Allowed region
Residue [A 108 :ASN] (-126.63,  75.19) in Allowed region
Residue [A 110 :ASN] ( -57.21, 111.16) in Allowed region
Residue [A 135 :TYR] (-179.61, 119.60) in Allowed region
Residue [A 209 :ASP] ( -99.27,-144.36) in Allowed region
Residue [A  61 :GLY] (-144.71,  80.60) in Outlier region
Number of residues in favoured region    (~98.0% expected)  :  177 ( 95.2%)
Number of residues in allowed region     ( ~2.0% expected)  :    8 (  4.3%)
Number of residues in outlier region                        :    1 (  0.5%)
```

FIGURE 11: RAMPAGE of Cellulase

Inference: To model constructed for cellulase has to be checked for the quality of the structure. For this Ramachandran plot was constructed Fig.11 shows the result of the cellulase sequence submitted for constructing the ramachandran plot. The blue coloured region on the graph shows the general favoured region and the region with blue and orange dots showed the allowed region for the amino acids for the formation of structure. Orange shaded region shows the favoured and allowed region of Glycine amino acid. Green coloured region on the plot shows the favoured and allowed region for Proline amino acid. Evaluation of the residues shows the allowed region of each amino acid. About 95.2% of residues are in the favoured region,it shows that the model is acceptable.

Pectinase


180
ψ 0
-180
-180
0
180
ϕ
General/Pre-Pro/Proline Favoured
Glycine Favoured
General/Pre-Pro/Proline Allowed
Glycine Allowed
General
Glycine
Pre-Pro
Proline
General Favoured
Glycine Favoured
Pre-Pro Favoured
Proline Favoured
General Allowed
Glycine Allowed
Pre-Pro Allowed
Proline Allowed

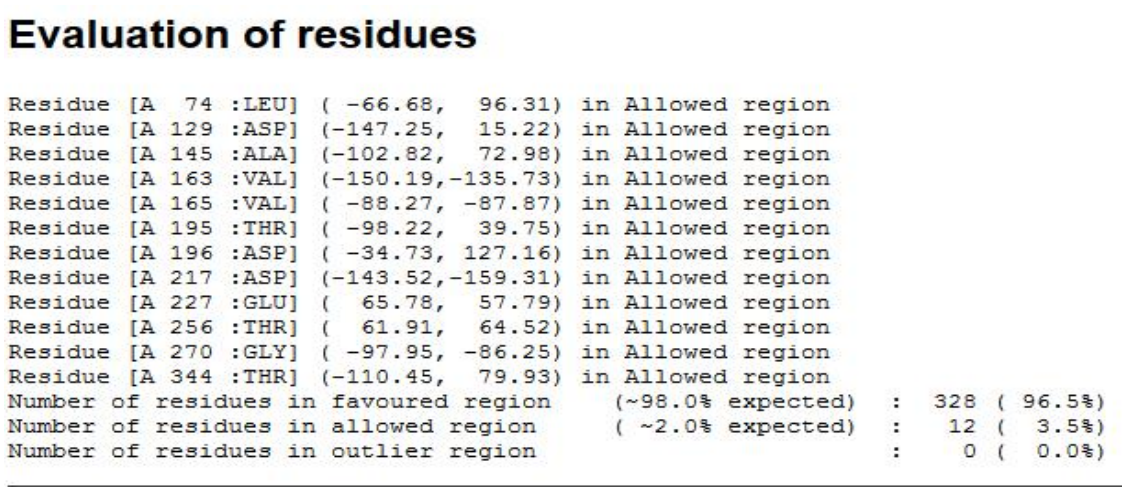


FIGURE 12: RAMPAGE of Pectinase

Inference: To model constructed for pectinase has to be checked for the quality of the structure. For this Ramachandran plot was constructed Fig. 12 shows the result of the pectinase sequence submitted for constructing the ramachandran plot. The blue coloured region on the graph shows the general favoured region and the region with blue and orange dots showed the allowed region for the amino acids for the formation of structure. Orange shaded region shows the favoured and allowed region of Glycine amino acid. Green coloured region on the plot shows the favoured and allowed region for Proline amino acid. Evaluation of the residues shows the allowed region of each amino acid. About 96.5% of residues are in the favoured region, it shows that the model is acceptable.

CONCLUSION

Protease, Cellulases and Pectinase are class of enzymes produced by fungi, bacteria and protozoans. They catalyze proteolysis, cellulolysis and pecteolysis respectively. These enzymes used extensively used in industries, especially in textile, food, paper and pulp, wine production, leather, fruit and juice, animal feed, detergent and in the bioconversion of wastes. The extensive use of these enzymes in industries resulted in research being carried out to isolate better microbial strains and also to develop new fermentation processes with the aim to reduce the product cost. Protease from *Pseudomonas* species, Cellulase from *Ceratocystis sp.* and Pectinase from *Alternaria sp.* were analyzed using computational tools. Homology modeling of the purified samples was carried out resulting in the production of stable structure, which can be later used commercially for the mass production of enzymes. In homology modeling, protease, cellulase and pectinase enzymes were first analysed using sequence alignment and the most identical sequence similarity model was constructed using SWISS MODEL and its stability was confirmed by ploting the ramachandran plot using RAMPAGE. For all the three enzymes a stable model was constructed.

ACKNOWLEDGEMENT

The author S.V. Megha is grateful to R & D Bharatiyar university and Biotechnology department.Megha extend her gratitude to Dr. B. Annadurai, Dr. R. Puvanakrishnan, Scientist at department of Biotechnology CLRI, Chennai, Dr. R. Easwaramoorthy and all the well wishers for useful suggestions.

REFERENCES

Ali, A. J., Ahmed, A.H., Zahraa, A. and Umar, A. (2011) Optimization of Cellulase Production by *Aspergillus niger* and *Tricoderma viride* using Sugar Cane Waste, Journal of Yeast and Fungal Research, 2(2), 19 – 23.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res., 25, 3389-3402

Benkert, P., Kanzil, M. and Schwede (2009) QMEAN server for protein model quality estimation, Nucleic Acid, 37, 510-514.

Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., Kiefer, F., Cassarino, T.G., Bertoni, M., Bordoli, L., Schwede, T. (2014) SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information, Nucliec acid Research 42,252-258

Garg, G., Singh, A., Kaur, A., Singh, R., Kaur, J. and Mahajan, R. (2016) Microbial pectinase : an ecofriendly tool of nature for industries, 3 biotech,Vol 6(1), 47.

Kaczanowski, S., Zielenkiewicz, P. (2010) Why similar protein sequences encode similar three dimensional structure, Theoretical Chemistry Accounts, 125, 643-650.

Pradeep, N.V., Anupama, Vidyashree, K. G., Lakshmi (2012) *In silico* Characterization of Industrial Important Cellulases using Computational Tools, Advances in Life Science and Technology , 4, 8-14.

Prashant, V.T., Uddhav, S.C., Madura, S.M., Vishal, P.D. and Renuka, R.K. (2010) Secondary Structure Prediction and Phylogenetic Analysis of Salt Tolerant Proteins, *Global Journal of Molecular Sciences*, 5 (1),30-36

Ramachandran, G.N., Ramakrishnan, C., Sasisekharan, V. (1963) "Stereochemistry of polypeptide chain configurations".J. Mol. Biol. 7, 95–9.

Remmert, M., Biegert, A., Hauser, A. and Soding, J. (2012) "HHblits: lightning-fast interative protein sequence searching by HMM-HMM alignment.", Nat Methods 9, 173-175.

Sadowski, M.I., Jones, D.T. (2007) Benchmarking template selection and model quality assessment for high-resolution comparative modeling, Proteins, 69, 476-485.

Spano, L., Medeiros, J., Mandels, M. (1975) Enzymatic hydrolysis of cellulosic waste to glucose, pollution abatement DIV., Food SVCS. Labs. Us Army Natick, MA, USA, 7 Jan.

Williamson, A. R. (2000) Creating a structural genomics consortium, Nature Strcture Biology, 7, 953.