Journal of Engineering and Technology Research Vol.1 (2), pp. 039-045, May, 2009
Available online at http:// www.academicjournals.org/JETR

Full Length Research Paper

Identification of insilico 3D structure of amylase (*Drosophila melanogaster*) and comparative computational studies

Ashish Patel[1]*, Ruby Dewangan[1], Sheetal Khatri[2], Jyotsana Choubey[3], Shailendra Kumar Gupta[4] and M. K. Verma[5]

[1]Sub-DIC Bioinformatics Centre, NIT Raipur, Raipur (CG) 492010, India.
[2]Biotechnology Engineering Department, NIT Raipur, Raipur (C.G.) 492010, India.
[3]Biomedical Engineering Department, NIT Raipur, Raipur (C.G.) 492010, India.
[4]Indian Institute of Toxicology Research, Lucknow (U.P.), India.
[5]Civil Engineering Department, NIT Raipur, Raipur (C.G.) 492010, India.

Accepted 13 May, 2009

Comparative studies are useful in the prediction of validated 3D structure of a query protein (Amylase). Modeler is used for structure prediction of amylase. All the structural models are verified by results of ramachandran plot, Whatif server, Prosa (II) and Procheck (Checks the stereochemical quality of a protein structure, producing a number of postscript plots analyzing its overall and residue-by- residue geometry. Predicted and validated structure (after comparative study) is useful in structure based drug designing, protein-DNA interactions, study of protein-protein interactions and protein-ligand binding. Amylase is an enzyme that breaks down starch into sugar. Amylase is present in human saliva, where it begins the chemical process of digestion. The pancreas also makes amylase (alpha amylase) to break down dietary starch into disaccharides which are converted by other enzymes to glucose to supply the body with energy. There are 3 subunits of amylase, α, β and γ. A higher than normal concentration may reflect one of several medical conditions, including acute inflammation of the pancreas, macroamylasemia and perforated peptic ulcer, urine and blood amylase levels may also be elevated with a variety of other conditions, such as ovarian cancer, lung cancer, tubal pregnancy, mumps, intestinal obstruction, or perforated ulcers.

Key words: Amylase, ramachandran plot, stereochemical, 3D structure, modeler.

INTRODUCTION

Amylase is an enzyme that breaks starch down into sugar. Amylase is present in human saliva, where it begins the chemical process of digestion. Foods that contain much starch but little sugar, such as rice and potato, taste slightly sweet as they are chewed because amylase turns some of their starch into sugar in the mouth. The pancreas also makes amylase (alpha amylase) to break down dietary starch into disaccharides which are converted by other enzymes to glucose to supply the body with energy. Plants and some bacteria also produce amylase. As diastase, amylase was the first enzyme to be discovered and isolated (Payen, 1833). Specific amylase proteins are designated by different Greek letters. All amylases are glycoside hydrolyses and act on α-1, 4-glycosidic bonds (Hopkins et al., 1993).

Bioinformatics uses computational approach to answer the biological problems. Answering these questions requires that investigators take advantage of large, complex data sets (both public and private) in a rigorous fashion to reach valid, biological conclusions (Bernstein et al., 1977; Berman et al., 2000; Laskowski et al., 1993). With the explosion of sequence and structural information available to researchers, the field of bioinformatics is playing an increasingly large role in the study of fundamental biomedical /biotechnological problems.

Homology modeling is basically used for the prediction of protein structure and it constructs an atomic-resolution model of a protein from amino acid of query sequence.

*Corresponding author. E-mail: ashish111ppp@gmail.com.

Table 1. List of templates and related information's generated by Modeler are shown below.

S. No.	Templates	No. of Residues in Templates	Resolution	E value	Percentage Similarity
1.	1G94	448	1.74	0	49.10
2.	1SMD	495	1.60	0	54.34
3.	1HX0	495	1.38	0	55.21
4.	1JAE	470	1.65	0	63.75

Table 2. Dope score of models are listed below.

S. No.	Models generated by modeler	Dope score
1.	amy.B99990001.pdb	-54110.55078
2.	amy.B99990002.pdb	-54748.55078
3.	amy.B99990003.pdb	-54526.17578
4.	amy.B99990004.pdb	-54112.88672
5.	amy.B99990005.pdb	-54041.39844

Table 3. RasMol results of models.

S. No.	Features	Model-1	Model-2	Model-3	Model-4	Model-5
1.	Hydrogen Bond	307	303	302	303	301
2.	Helices	19	19	19	19	19
3.	Strands	41	41	41	43	43
4.	Turns	56	60	57	50	50

The quality of the homology model is dependent on the quality of the sequence alignment and template structure (Marti-Renom et al., 2000).

The aim of this study to predict validated 3 dimensional structures from the protein sequence by comparative studies of three structures. Modeling tools and related structural data available on the online databases are used for the comparative studies between 3 structures.

METHODOLOGY

MODELER is a computer program used in producing homology models of protein tertiary structures and sometimes in quaternary structures (MODELER was originally written and is currently maintained by Andrej Sali at the university of california, san francisco) (Fiser and Sali, 2003; Marti-Renom et al., 2000; Sali and Blundell, 1993).

Modeling is a method of designing 3D model for a protein of unknown structure based on one or more related protein of known structure. MODELER is a computer program that most frequently used for homology modeling. Homology modeling requires at least one sequence of known 3D structure with significant similarity with the target sequence. MODELER is a command-line tool and has no graphical user interface. It requires a script file containing MODELER commands. This is an ordinary Python 2.5 script. The script file contains commands for MODELER. A script files to produce models of target sequence from the known structure. Target sequence is used as a query and selection of target is according to the lowest resolution value is done by running the "build_profile.py" file in MODELER. From the log file of "build_profile.py" the sequence having lowest e-value and alignment score greater than 45% considered as the templates and selected templates are compared with the query in "compare.py" file by running the modeler (Table 1). After this, template having the lowest resolution value selected as a template and finally "model-single.py" file is running in modeler for models generation, best model is selected according to its lowest Dope score (Table 2). Model evaluation is done by using the GNUPLOT, gnu plot is a command-driven interactive function and data plotting program. It is case sensitive (commands and function names written in lowercase are not the same as those written in CAPS). All command names may be abbreviated as long as the abbreviation is not ambiguous. Any number of commands may appear on a line (with the exception that load or call must be the final command), separated by semicolons (;). Strings are indicated with quotes. They may be either single or double quotation marks, e.g., load "filename" or cd 'dir'.

RESULTS AND DISCUSSION

5 3D Models of our query protein amylase was generated by the modeler 9v3 (Figures 1, 2, 3, 4 and 5) and visualized by the rasmol with hydrogen bonds helices, strands and turns in the model (Table 3). The maximum number of turns in the model 2 indicates that this structure is more compact than others. The best model (Figure 2), is selected according to its lowest Dope score (Table 2) and then verifying model with the help of Whatif server and procheck. After running the Procheck (Laskowski et al., 1993), ramachandran plot shows (Figure 6a) that 87.3% residues are in the favored region, 11.5% in the additional allowed region, and 0.5% in the generous allowed

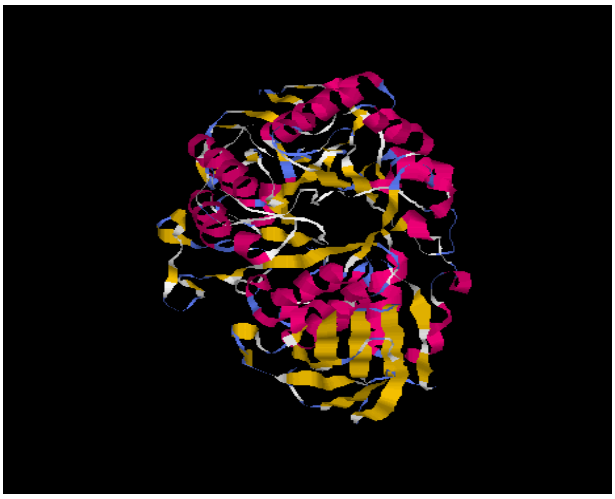

Figure 1. Model (amy.B99990001.pdb).

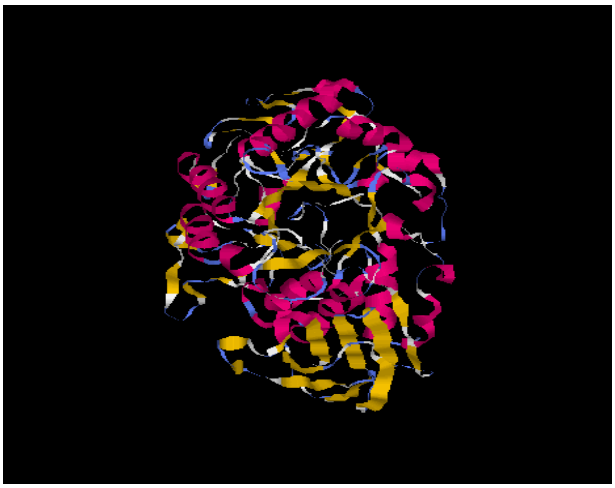

Figure 2. Model (amy.B99990002.pdb).

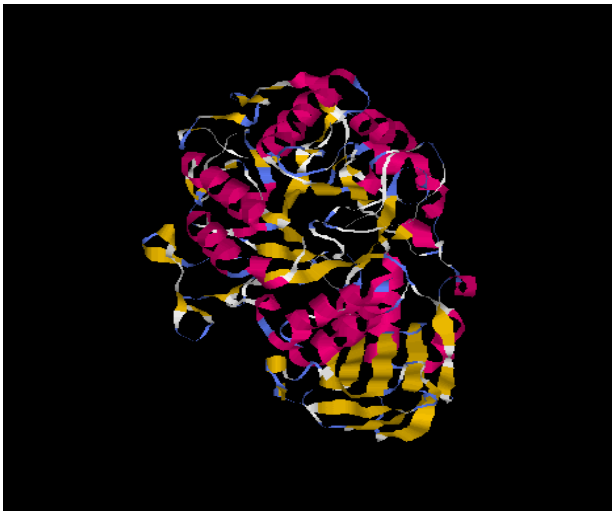

Figure 3. Model (amy.B99990003.pdb).

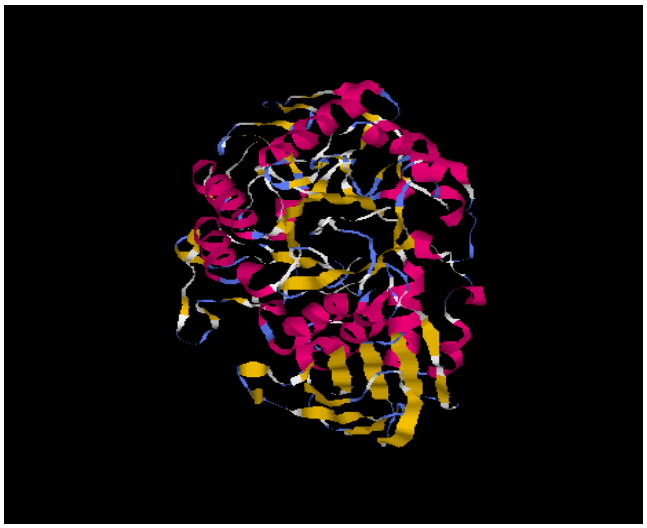

Figure 4. Model (amy.B99990004.pdb).

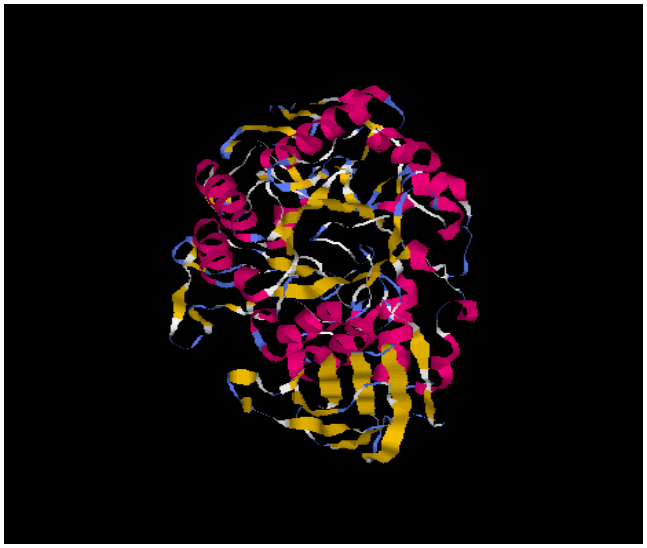

Figure 5. Model (amy.B99990005.pdb).

region and only 0.7% residues in the disallowed region which is acceptable and better than other generated models (Figures 1, 3 ,4 and 5). Ramachandran plot never shows the Gly and Pro, so these residues are shown in other plot (Figure 6b). The plot shows that proline and glysine both are in allowed region. Whatif server used to check the nomenclature of Torsion angles. The residues are sorted by residue type. Evolution of model also shows that generated model is totally based on selected templates (Figures 7a and 7b).

No errors were detected in valine, threonine, isoleucine, leucine and arginine nomenclature. No errors were detected in tyrosine, phenylalanine, aspartic acid, glutamic acid torsion angle conventions. An additional Whatif check involved looking at the fine-packing quality for the homology model for amylase. Quality control values for

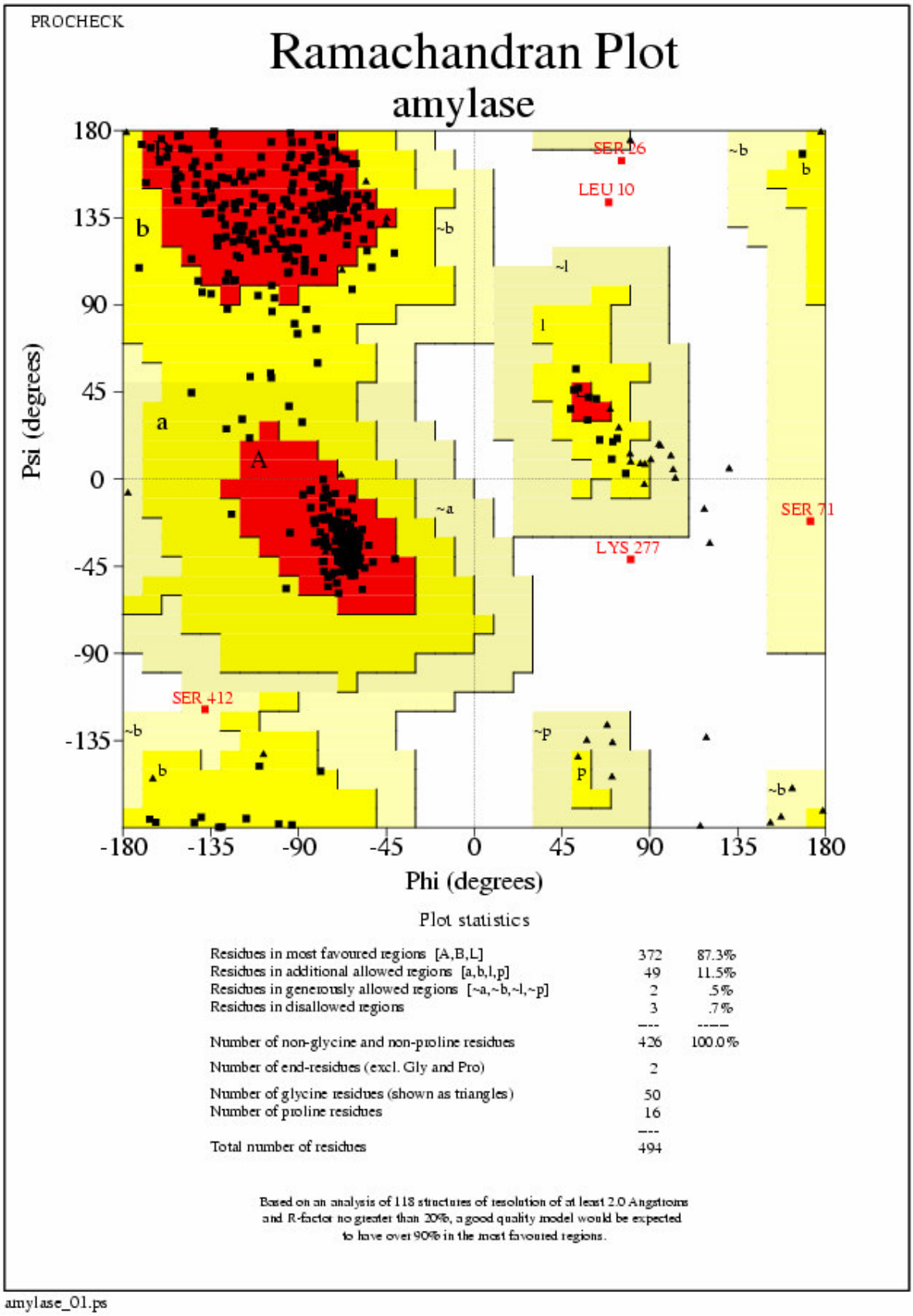

Figure 6a. Ramachandran plot statistics.

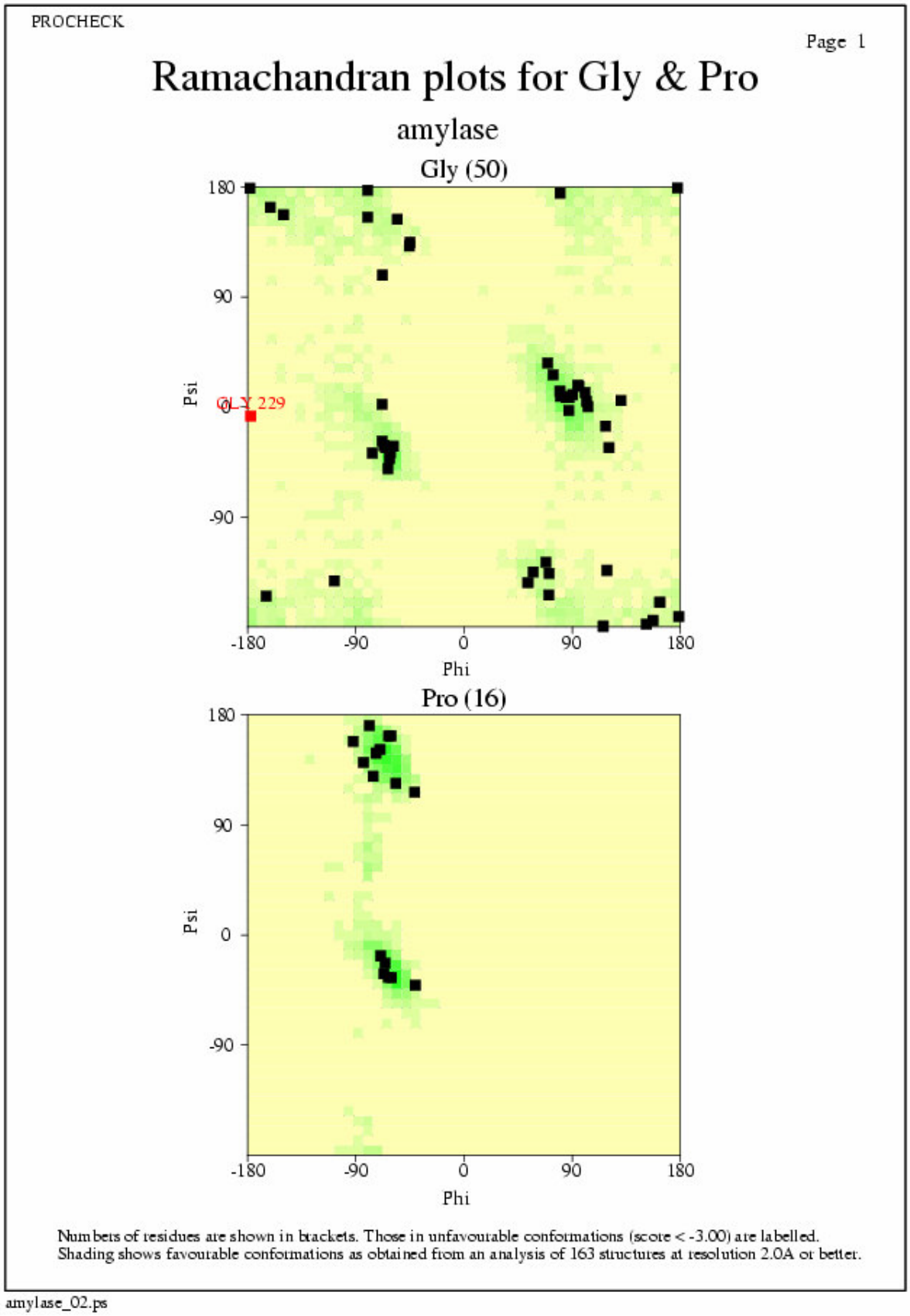

Figure 6b. Ramachandran plot for Gly and Pro.

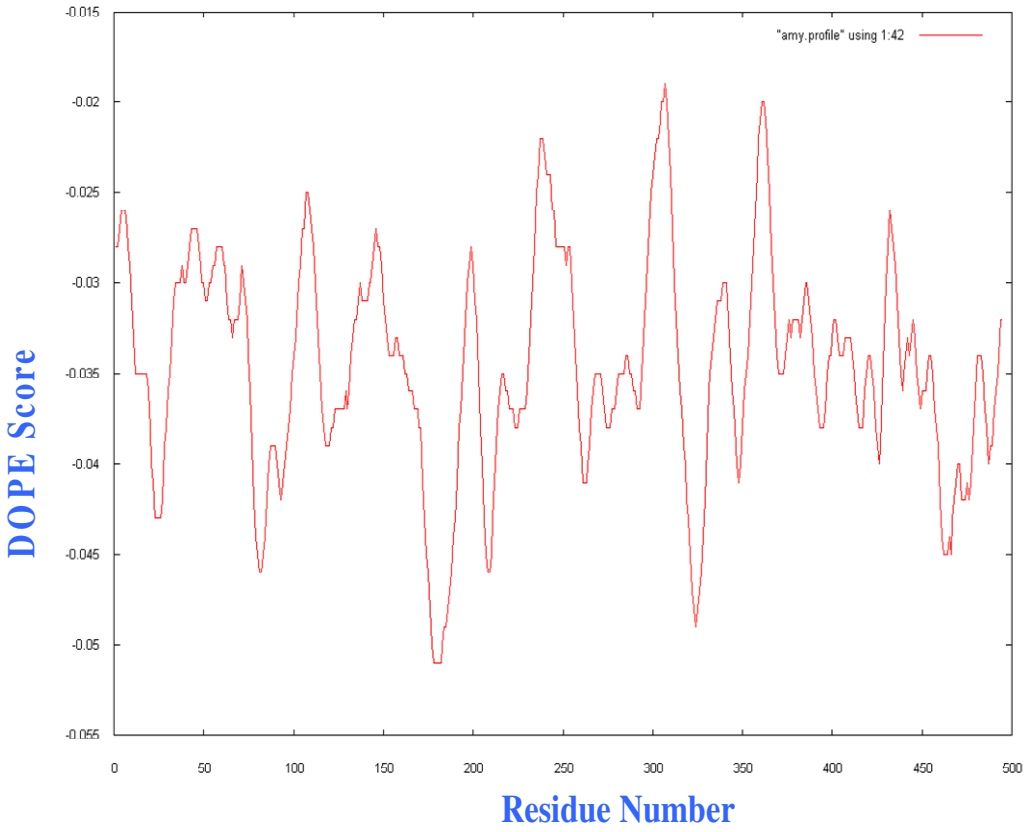

Figure 7a. GNUPLOT for model.

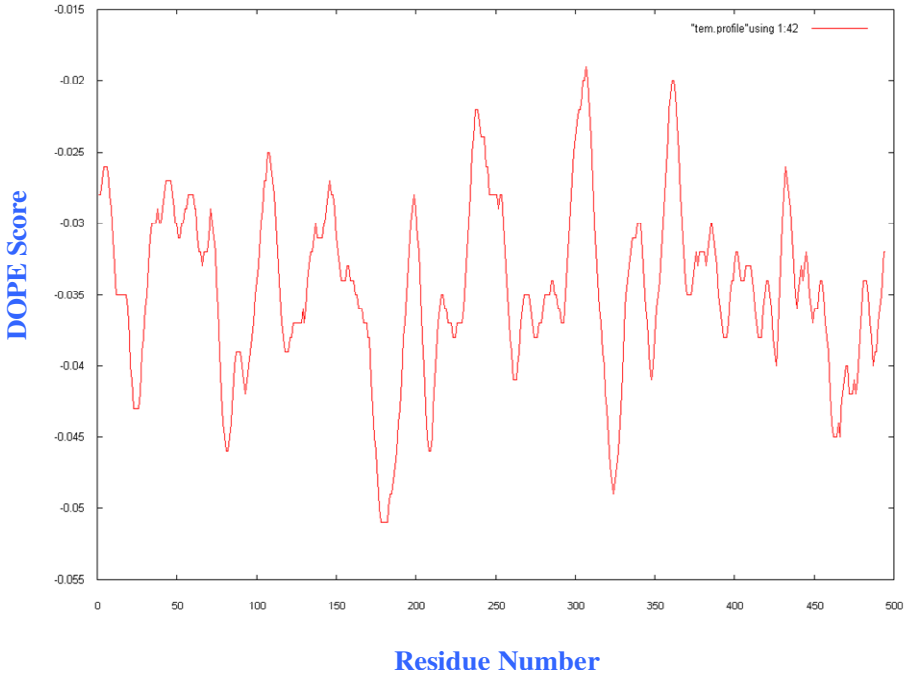

Figure 7b. GNUPLOT for Template.

Table 4. For amylase.

52 PRO	(52) - 103.7 envelop C-beta (108 degrees)
62 PRO	(62) - 105.5 envelop C-beta (108 degrees)
134 PRO	(134) - 103.7 envelop C-beta (108 degrees)
140 PRO	(140) - 101.3 envelop C-beta (108 degrees)
211 PRO	(211) - 108.5 envelop C-beta (108 degrees)
368 PRO	(68) - 104.2 envelop C-beta (108 degrees)
443 PRO	(443) - 119.8 half-chair C-delta/C-gamma (-126 degrees)

individual residues are intended to measure the fit of a residue in the particular part of the structure. If the z-score for the protein is greater than -2.5 (2.5 standard deviations less than the mean), then the structure is acceptable or good. A z-score less than -3 suggests that the structure likely is poor. The z-scores for all types of contacts (BB-BB, BB-SC, SC-BB and SC-SC) were 0.707 and 0.721 for the 1HX0 and model, respectively, indicating an overall improvement in the packing quality for the homology model compared with the template. Puckering amplitudes for all PRO residues are within normal ranges. The proline residues (Table 4) have a puckering phase that is not expected to occur in protein structures. Puckering parameters were calculated by the method of Cremer and Pople.

Normal PRO rings approximately show a so-called envelope conformation with the C-gamma atom above the plane of the ring (phi = + 72 degrees), or a half-chair conformation with C-gamma below and C-beta above the plane of the ring (phi = - 90 degrees) (Morris et al., 1992). If phi deviates strongly from these values, this is indicative of a very strange conformation for a PRO residue and definitely requires a manual check of the data. Be aware that this is a warning with a low confidence level.

Conclusion

The result of comparative structural analysis shows that model (Figure 2) is the best structural model for amylase of *Drosophila melanogaster* because it shows maximum residues(87.3%) in the favored region with the highest Z-score value which is 0.721 (all types of contacts such as BB-BB, BB-SC, SC-BB and SC-SC).

REFERENCES

Bernstein FC, Koetzle TF, Williams GJ, Meyer EFJ, Brice MD, Rodgers JR, Kennard O, Shimanouchi T, Tasumi M (1977). The Protein Data Bank. A computer-based archival file for macromolecular structures. Eur. J. Biochem. 80(2): 319-24.

Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000). The Protein Data Bank. Nucleic. Acids Res. 28 (1): 235-42.

Fiser A, Sali A (2003). Modeler: generation and refinement of homology-based protein structure models. Methods. Enzymol. 374: 461-91.

Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993). PROCHECK: A program to check the stereo chemical quality of protein structures.

Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993). PROCHECK: A program to check the stereo chemical quality of protein structures. J. Appl. Cryst. 26: 283-291.

Marti-Renom MA, Stuart A, Fiser A, Sánchez R, Melo F, Sali A (2000). Comparative protein structure modeling of genes and genomes. Annu. Rev. Biophys. Biomol. Struct. 29: 291-325.

Hopkins J, Maton, Anthea, Charles WM, Susan J, Maryanna QW, David L, Jill DW (1993). Human Biology and Health.

Morris AL, MacArthur MW, Hutchinson EG, Thornton JM (1992). Stereo chemical quality of protein structure coordinates. Proteins 12: 345-364.

Sali A, Blundell TL (1993). Comparative protein modeling by satisfaction of spatial restraints. J. Mol. Biol. 234: 779-815.