PREDICTION OF THREE DIMENSIONAL MODEL OF, γ- HEXACHLOROCYCLOHEXANE DEHYDROCHLORINASE (*LINA)* IN *SPHINGOMONAS PAUCIMOBILIS* UT26

Deepti Khewariya* and Richa Singh

Department of Biotechnology, Banasthali University, Jaipur, Rajasthan, India

ABSTRACT

The present study is undertaken to predict the three dimensional structure of γ-Hexachlorocyclohexane dehydrochlorinase (*linA*) in *Sphingomonas paucimobilis* UT26. γ-Hexachlorocyclohexane dehydrochlorinase (*linA*) catalyzes the initial steps in the biotransformation of the important insecticide γ-hexachlorocyclohexane (γ-HCH) by the soil bacterium *Sphingomonas paucimobilis* UT26. Stereo chemical analysis of the reaction products formed during conversion of γ-HCH by *linA* was identified from structural database using homology modeling or comparative modeling approach. Based on the knowledge of the template, a three-dimensional model was predicted and processed in to energy minimization, Ramachandran plot analysis, quality assessment and finally deposited into Protein Model Database.

INTRODUCTION: γ-Hexachlorocyclohexane (γ-HCH; also called γ-BHC and lindane) is a highly halogenated organic insecticide that has been used worldwide. Due to its toxicity and long persistence in soil, most countries have prohibited the use of γ-HCH. However, many contaminated sites remain throughout the world. Moreover, some countries are presently using γ-HCH, mainly for economic reasons, so new sites are continually being contaminated [1].

Sphingomonas (formerly *Pseudomonas*) *paucimobilis* UT26 is a unique microorganism that utilizes γ-HCH as its sole source of carbon and energy under aerobic conditions [2]. The four genes of *Sphingomonas* are *linA*, *linB*, *linC*, and *linD*, the products of which are involved sequentially in the degradation of γ-HCH in UT26 [3]. Three of them encode different types of dehalogenases: dehydrochlorinase, halidohydrolase, and reductive dehalogenase. Among them, γ-Hexachloro cyclohexane dehydrochlorinase (*linA*) encodes an enzyme that catalyzes a unique dehydrochlorination reaction whose mechanism is still unknown [4].

In this study, our aim was to derive a model of the tertiary structure of *linA* by using simple template and a model evaluation through a comparative modeling approach.

MATERIALS AND METHODS:

Retrieval of γ-Hexachlorocyclohexane dehydro chlorinase (*linA*) Protein Sequence: The protein sequence of γ-Hexachlorocyclohexane dehydro chlorinase (*linA*) was retrieved from the NCBI and taken as target sequence. The main reason for choosing this protein was that it is involved in the early steps of γ-HCH degradation [4].It was innate that the three-dimensional structure of this protein was not available in any three-dimensional structural databases. Hence, the current study of developing the three- dimensional structure of γ-

Hexachlorocyclohexane dehydrochlorinase (*linA*) was carried out.

Selection of Structural Template: An effort was made to find an appropriate structural homolog or template for the modeling of the γ-Hexachlorocyclohexane dehydrochlorinase (*linA*). In the beginning, a structural template was obtained from Protein BLAST [5] and Protein Data Bank [6] was used as reference database to identify the closely related sequences.

Target- Template Alignment: The protein sequence of γ-Hexachlorocyclohexane dehydrochlorinase (*linA*) was aligned with the corresponding template by using align-2D module in MODELLER 9V2 [7] which required two files: a file containing target sequence in PIR format and another file containing structural coordinates of template. This step was essential to identify the common conserved residues or active residues present in both the sequences.

Model Building: MODELLER 9V2 was used to envisage the three-dimensional structure of γ-Hexachlorocyclohexane dehydrochlorinase (*linA*) based on satisfaction of spatial restraints. It is a python script based program used to predict the three dimensional model from single template.

Model Evaluation: Once target-template alignment was completed, a three dimensional structure of γ-Hexachlorocyclohexane dehydrochlorinase (*linA*) was predicted. Five sets of models were generated by satisfaction of spatial restraints. The restraints include distance and dihedral angles for the backbone and side chains.

The value of DOPE score for five models was -13817.35059, -13817.35059, -13880.77148, -13545.13086, and -13890.94824. The 5th model was selected due to lowest DOPE score value [7]. Refined model was subjected to a series of tests for testing its internal stability and reliability. Backbone conformation of the refined model was assessed by

the examination of the Psi/Phi Ramachandran plot obtained from RAMPAGE web server [8]. Harmony web server [9] was also used which examines the compatibility between the sequence and the structure of a protein by assigning scores to individual residues and their amino acid exchange patterns after considering their local environment.

RESULTS AND DISCUSSION: An evaluation of the refined model involved two independent tests. The first test was to compare the residue backbone conformations in our refined model with the preferred values obtained from Protein Data Bank of known structures. The results of RAMPAGE web serve explored that 98% residues of refined model of γ-Hexachlorocyclohexane dehydrochlorinase (*linA*) were found to be in the most favoured region of the Ramachandran Plot (see **Figure 1 and 2**) which is more than the cut-off value of 96.1% in most of the reliable models [8]. The stereo chemical

quality of the predicted model was found to be acceptable and low percentage of residues having phi/psi angles in the outlier region.

The second test was carried out by using HARMONY web server (http://caps.ncbs.res.in /harmony/) to check the quality of the models that examines the compatibility between the sequence and the structure of a protein by assigning scores to individual residues and their amino acid exchange patterns after considering their local environment [9]. Here misfolds and folds with elaborate regions of errors can be identified by plotting the scores of the query protein structure against a calibration plot. Protein with misfold regions will have significantly low harmony score and fall well below the straight fit line. The result explored from the HARMONY server is shown in **figure 3** from which it is clear that the structure is acceptable.
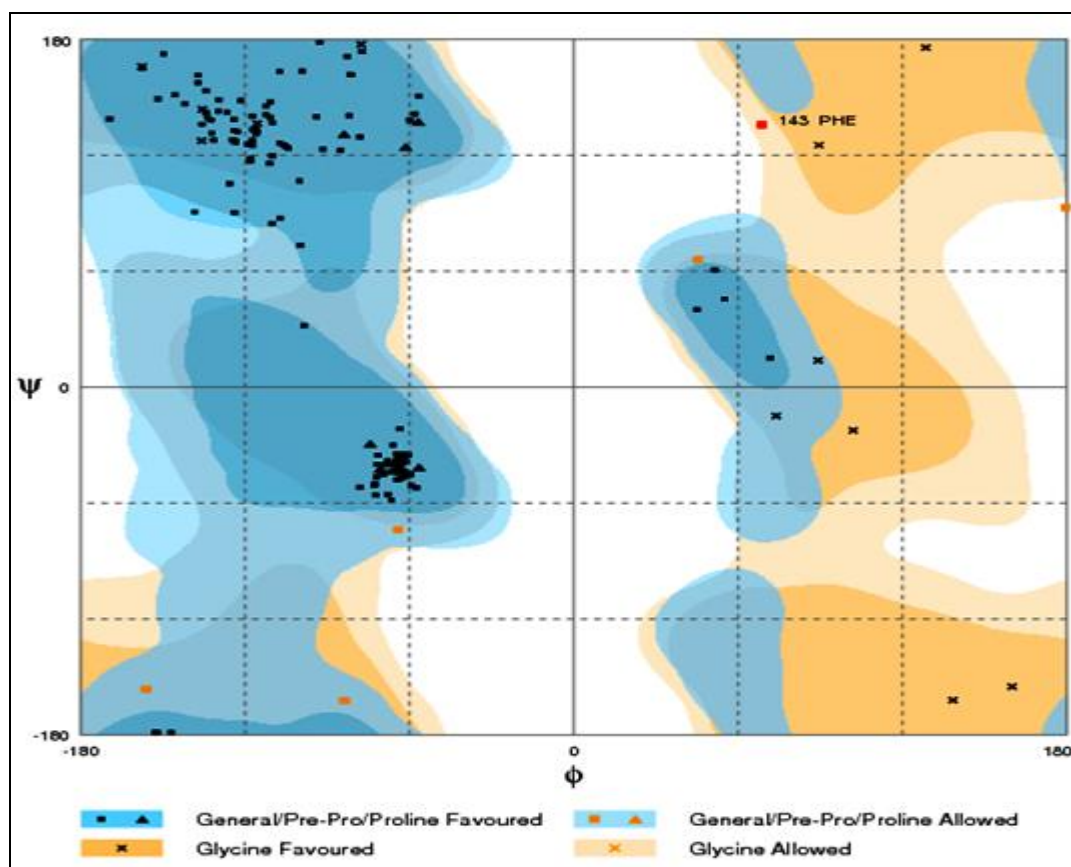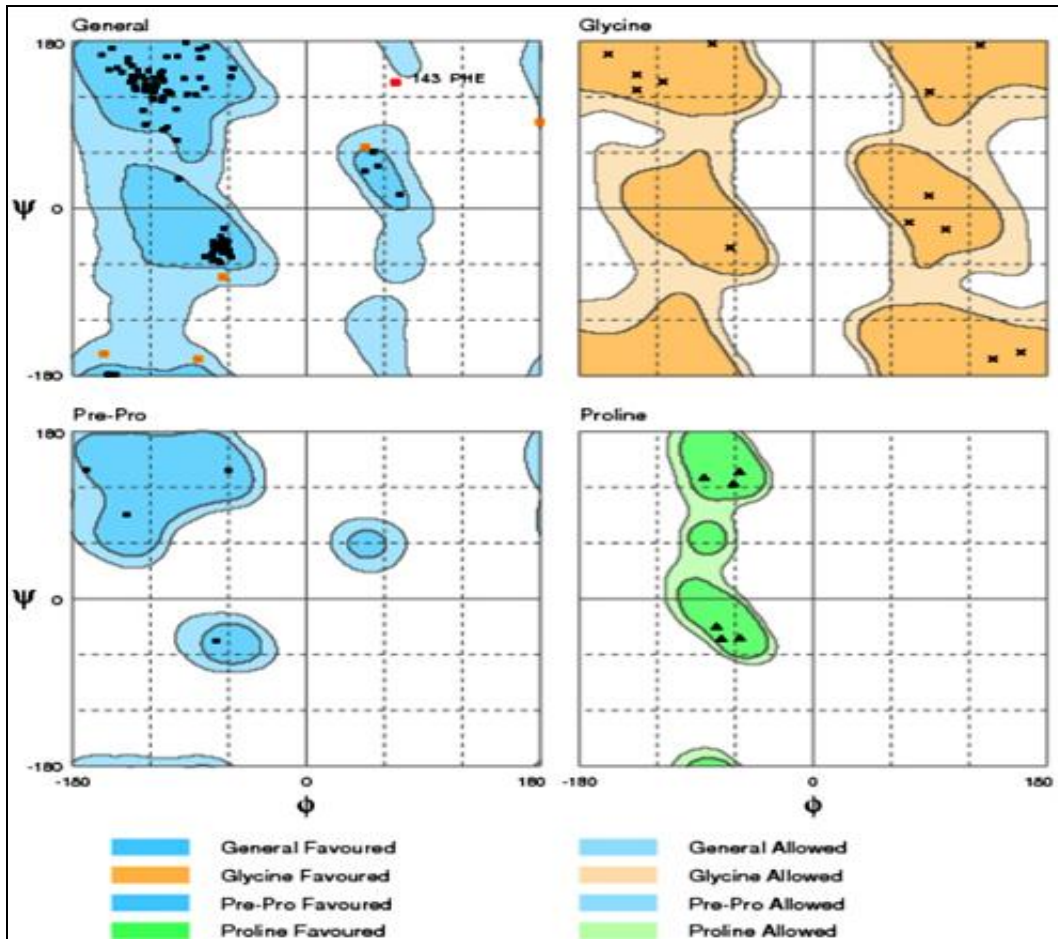

FIG. 1: PROTIEN STRUCTURE QUALITY

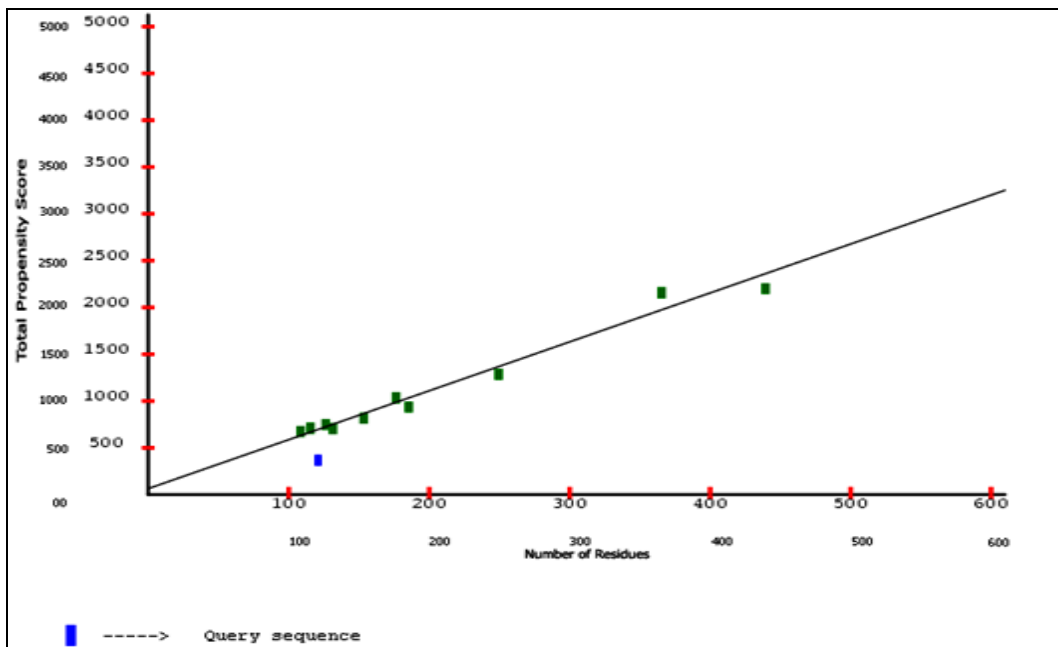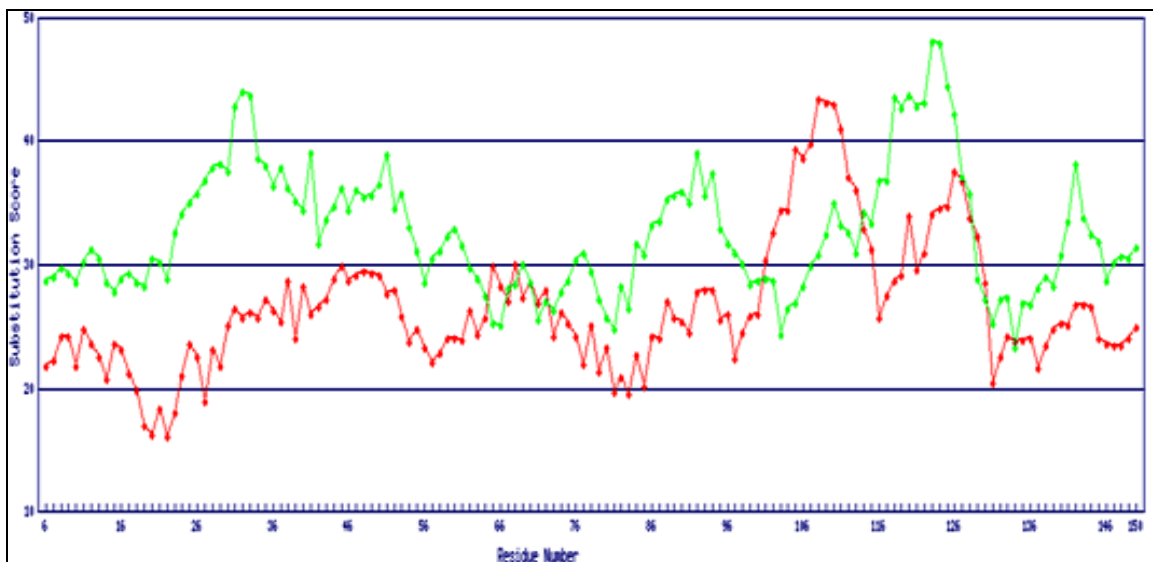**FIG. 2: AMINO ACID DISTRIBUTION IN PREDICTED PROTEIN**

FIG. 3: GRAPH SHOWING COMPATIBILITY BETWEEN SEQUENCE AND PREDICTED STRUCTURE

The result explored from the HARMONY server is shown in **figure 3** from which it is clear that the structure is acceptable. Again the graph represented in **figure 4** provides the smoothened scores between query sequence in comparison to the reverse sequence. The reverse sequence and its scores are used as a control to identify local errors in the proposed protein model. Regions where the reverse sequence acquire better substitution score could indicate possible local errors. The predicted structure designated acceptable local error. At last the predicted model was submitted to PMDB (protein model database) server and its PMDB id is pm0076004.The predicted model is given below. **Figure 5** shows the predicted model of *lin-A* submitted to the PMDB sever and the evaluation methods of this protein structure indicate that the predicted structure is good and acceptable.

Green- Query sequence; **Red**- Reverse sequence

FIG. 4: GRAPH BETWEEN QUERY AND REVERSE SEQUENCE INDICATING MINIMUM LOCAL ERROR IN THE PREDICTED MODEL
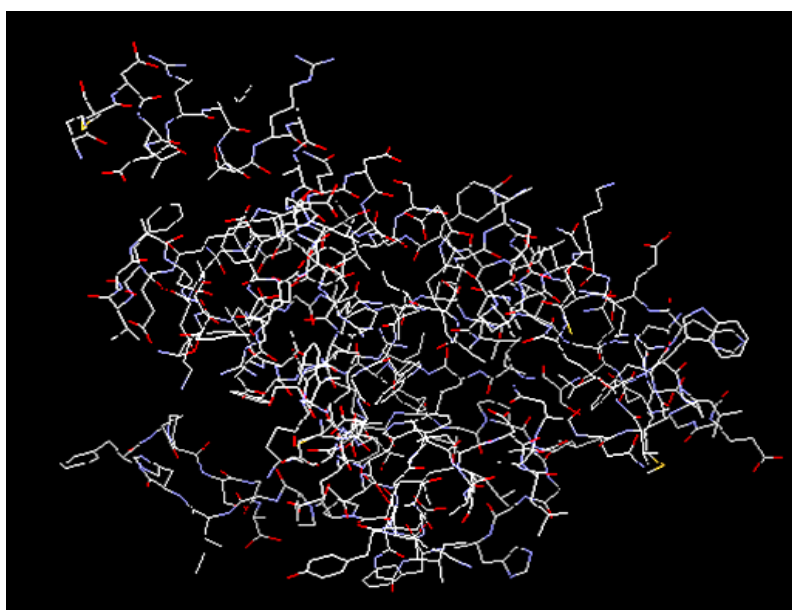

FIGURE 5- PREDICTED MODEL OF γ− HEXACHLOROCYCLOHEXANE DEHYDROCHLORINASE (*LINA*)

REFERENCES:

1. Blais JM, Schindler DW, Muir DCG, Kimpe LE, Donald DB and Rosenberg B: Accumulation of persistent organochlorine compounds in mountains of western Canada. Nature 1998; 395:585–588.

2. Imai R, Nagata Y, Senoo K, Wada H, Fukuda M, Takagi M and Yano K: Dehydrochlorination of γ-hexachlorocyclohexane (γ-BHC) by γ-BHC-assimilating *Pseudomonas paucimobilis*. Agric Biol Chem 1989; 53:2015–2017.

3. Imai R, Nagata Y, Fukuda M, Takagi M and Yano K: Molecular cloning of a *Pseudomonas paucimobilis* gene encoding a 17-kilodalton polypeptide that eliminates HCl molecules from γ-hexachlorocyclohexane. J Bacteriol 1991; 173:6811–6819.

4. Nagata Y, Futamura A, Miyauchi K and Takagi M: Two different types of dehalogenases, *linA* and *linB*, involved in γ-hexachlorocyclohexane degradation in *Sphingomonas paucimobilis* UT26 are localized in the periplasmic space without molecular processing. J Bacteriol 1999; 181: 5409–5413.

5. Altschul SF, Gish W, Miller W, Myers EW and Lipman DJ: Basic local alignment search tool. J Mol Biol 1990; 215: 403-10.

6. Berman H, Henrick K and Nakamura H: Announcing the worldwide Protein Data Bank. Nat Struct Biol 2003; 10: 980.

7. Eswar N and Sali A: Protein Structure Modeling. In: From Molecules to Medicine, Structure of Biological Macromolecules and Its Relevance in Combating New Diseases and Bioterrorism. Ed: J.L. Sussman, P. Spadon, 2009; pp. 139-151, Springer-Verlag, Dordrecht, the Netherlands.

8. Lovell SC, Davis IW, Arendall WB, de Bakker PI, Word JM, et al: Structure validation by C alpha geometry: phi, psi and C beta deviation. Proteins 2003; 50: 437-50.

9. Pugalenthi G, Shameer K, Srinivasan N and Sowdhamini R: HARMONY: a server for the assessment of protein structures. Nucleic Acids Research 2006.

****