In silico Structural Analysis and Binding of Organophosphorus hydrolase of *Kocuria* sp with Chloropyrifos

Nagavardhanam Neti
Department of Botany, Acharya Nagarjuna University, Nagarjuna Nagar -522 510, Andhra Pradesh, India

Vishnuvardhan Zakkula
Department of Botany, Acharya Nagarjuna University, Nagarjuna Nagar -522 510, Andhra Pradesh, India

ABSTRACT
The most popular type of pesticide is the organophosphate (OP) family, which effectively eliminates pests owing to its acute neurotoxicity. Organophosphorus hydrolase is a bacterial enzyme that is capable of degrading a wide range of neurotoxic organophosphate nerve agents. Organophosphorus hydrolase of *Kocuria* sp was isolated but its protein is not having any predicted 3-Dimentional structure available in PDB (Protein databank) as elucidated by X-ray crystallography or NMR. Its structure was determined *in silico* by sequence homology. The gene sequence of the Organophosphorus hydrolase of *Kocuria* sp was known and its protein sequence was subjected to PSI-BLAST at NCBI. There was neither identical sequence available nor the nearest neighbour in the blast analysis. Then an alternative method for finding the homologous protein i.e., fold prediction method was used. The generated model was subjected to several repeated cycles of energy minimization using SPDBV software and the final model was subjected to stereo chemical evaluation. The homology modeled structure of the Organophosphorus hydrolase of *Kocuria* sp was docked by different OP by Molgro virtual docker and the data were presented.

Key words
Chloropyrifos, Organophosphorus hydrolase, SPDBV, Molegro Virtual Docker

1. INTRODUCTION
Organophosphate pesticides (OP) are a group of highly toxic agricultural chemicals widely used in plant protection. Their usage has become an indispensable tool in agriculture for the control of weeds, insects and rodent pests. They are poisonous but play an important role in generating plenty of food to the world population [1, 2]. Compounds of this family are spontaneously hydrolyzed and cause neurotoxicity in mammals [3]. Excessive pesticide usage resulted in accumulation of pesticide residues in crops, soils, and biosphere creating an ecological stress [4]. Chlorpyrifos is a broad spectrum systemic phosphorothioate ester insecticide patented and introduced by Dow Chemical Company in United States of America in 1965 [5]. Chlorpyrifos is available in granules, wettable powder, dustable powder, emulsifiable concentrate [6] and used for the control of a wide range of pests such as cutworms, corn rootworms, cockroaches, grubs, flea beetles, flies, termites, fire ants, aphids, lice, leptinotarsa and other insects. It is applied to different crops including cotton, nuts, corn, fruits, vegetables, ornamental plants and is highly persistent in foliar application. Chlorpyrifos causes hazardous effects to the environment and also toxic to human beings resulting in headache, nausea, muscle twitching, convulsions, birth defects and even death. It is toxic to a variety of beneficial arthropods including bees, beetles and parasitic wasps. It kills fishes and birds in minute concentrations. Plants are affected by delayed seedling emergence, fruit deformities and abnormal cell division [7, 8, 9, 10, 11, 12]. It has antimicrobial property, hence prevents the proliferation of chlorpyrifos degrading microorganisms in soil [13].
In light of its importance in agriculture and a need to degrade it in the environment, the present study has been taken up to analyze the properties of Organophosphorus hydrolase (OPH) encoded by *opd* gene of *Kocuria* species and its binding efficiency by docking with OP *in silico*.

2. MATERIALS AND METHODS
2.1 *In silico* analysis of Organophosphorus hydrolase of *Kocuria* sp
The Organophosphorus hydrolase of *Kocuria sp.* has no 3-Dimentional structure available in PDB (Protein databank), as the 3-Dimentional structure was not elucidated either by using X- ray crystallographic or NMR studies. The protein sequence was subjected to PSI-BLAST at NCBI from the DNA sequence obtained by sequencing. The sequenced DNA sequence was

GATCGTGGATCCTCGATCGGCACAGGCGATCGGATGCAAAC
GAGAAGGGTTGTGCTCAAATCTGCGGCCGCGAGAACTCTGC
TCGGCGGCCTGGCTGGGTGCGCGACGTGGCTGGATCGATCG
GCACAGGCGATGCGATCAATACGTGCGCGTCCTATCACAAT
CTCTGAAGCGGGTTTCACACTGACTCACGAGGACATCTCGG
CAGCTCGGCAGGATTCTTGCGTGCTTGGCCAGAGTTCTTCG
GTAGCGCAAAGCTCTAGCGGAAAAGGCTGTGAGAGGATTGC
GCGCCAGAGCGGCTGGCGTGCGAACGATTGTCGATGTGTCG
ACTTTCGATATCGGTCGCGACGTCAGTTTATTGGCCGAGGT
TTCGCGGGCTGCCGACGTTCATATCTGGCGGCGACCGGCTT
GTGGTTCGACCCGCCACTTTCGATGCGATTGAGGTATGTAG
AGGAACTCACGCTAGTTCTTCCTGCGGTGAGATTCAATATG
GCATCGAAGTACACCGGAATTAGGGCGGGCATTATCAAGGT
CGCGACCACAGGCAAGGCGACCCCCTTTCAGGAGTTAGTGT
TAAAGGCGGCCGCCCGGGCCTCCTTGGCCACCGGTGTTCCG
GTAACCACTCACACGGCAGCAAGTCAGCGCGATGGTGAGCG
AGGCAGGCCGCCATTTTTGAGTCCGAAGCTTGAGCCCTCAC
GGGTTTGTATTGGTCACAGCGATGATACTGATGACTTGAGC
TATCTCACCGCCCTGCTGCGCGGATACCTCATCGGTCTAGA
CCACATCCCGCACAGTGCGATTGGTCTAGAAGATAATGCGA
GTGCATCACCGCTCCTGGGCATCCGTTCGTGGCAAACACGG
GCTCTCTTGATCAAGGCGCTCATCGACCAAGGCTACATGAA
ACAAATCCTCGTTTCGAATGACTGGCTGTTCGGGTTTTCGA
GCTATGTCACCAATATCATGGACGTGATGGATCGCGTGAAC

CCCGACGGGATGGCCTTCATTCACTGAGATCGTAAGCTTTC
ATGACGCCCGCAAGGTCGG

The protein sequence was subjected to PSI-BLAST at NCBI. Protein parameters were analysed by using the tool Prosite [14]. All the protein parameters with respect to amino acid composition, secondary structure prediction, hydrophobicity, isoelectric point etc were analyzed. The generated model was subjected to several repeated cycles of energy minimization using SPDBV software [15] and the final model was subjected to docking.

2.2 Docking of Chloropyrifos onto Organophosphorus hydrolase (OPH) of *Kocuria* sp.

The three dimensional structure of target Organophosphorus hydrolase of *Kocuria* sp. was generated by homology modelling using SPDBV (Swiss Protein Data Bank Viewer) at 2.6 Å RMSD resolution. Phytochemicals Chloropyrifos, debantics, paraxon and phosphamidon were obtained from Dr. Duke database (http://www.arsgrin.gov/duke/), which was searched against pubchem and chemspider database for the 2D structures and then with the help of open babel [http://openbabel.org/wiki/Main_Page] these 2D structures are converted to 3D structures. The 3D structures which are obtained were minimized using Hyperchem's MM+ force field. Molegro Virtual Docker V4.2 was used to detect the active sites and docking was performed by moldock function, which is an implementation of evolutionary algorithms (EAs), focused on molecular docking simulations. Docking was performed with all the potential active sites detected on Organophosphorus hydrolase enzyme. During Docking at first the molecules were prepared and bonds, bond orders, explicit hydrogens, charges, flexible torsions, were assigned if they were missing by the MVD program to both the protein and ligands. From the docking wizard ligands were selected and the scoring function used is Moldock score. The model was calibrated using a data set of more than 200 structurally diverse complexes from the PDB bind database with known binding affinities (expressed in kJ/mol) [16].

3. Results and Discussion

3.1 *In silico* analysis of Organophosphorus hydrolase of *Kocuria* sp

The translated protein sequence was

MQTRRVVLKSAAARTLLGGLAGCATWLDRSAQAMRSIRARP
ITISEAGFTLTHEDISAARQDSCVLGQSSSVAQSSSAKGCE
RIARQSGWRANDCRCVDFRYRSRRQFIGRGFAGCRRSYLAA
TGLWFDPPLSMRLRYVEELTLVLPAVRFNMASKYTGIRAGI
IKVATTGKATPFQELVLKAAARASLATGVPVTTHTAASQRD
GERGRPPFLSPKLEPSRVCIGHSDDTDDLSYLTALLRGYLI
GLDHIPHSAIGLEDNASASPLLGIRSWQTRALLIKALIDQG
YMKQILVSNDWLFGFSSYVTNIMDVMDRVNPDGMAFIH

3.1.1 ProtParam

The parameters computed by ProtParam include the molecular weight, theoretical pI, amino acid composition, atomic composition, extinction coefficient, estimated half-life, instability index, aliphatic index and grand average of hydropathicity (GRAVY). Molecular weight and theoretical pI are calculated as to compute pI/Mw [17].

The number of amino acids present in the Organophosphorus hydrolase of *Kocuria* sp. was 325 with a molecular weight: 35597.9 whose theoretical pI was 9.82.

3.1.2 Amino acid composition

Ala (A)	37	11.4%
Arg (R)	32	9.8%
Asn (N)	6	1.8%
Asp (D)	18	5.5%
Cys (C)	7	2.2%
Gln (Q)	12	3.7%
Glu (E)	9	2.8%
Gly (G)	25	7.7%
His (H)	6	1.8%
Ile (I)	19	5.8%
Leu (L)	33	10.2%
Lys (K)	9	2.8%
Met (M)	8	2.5%
Phe (F)	11	3.4%
Pro (P)	13	4.0%
Ser (S)	30	9.2%
Thr (T)	20	6.2%
Trp (W)	5	1.5%
Tyr (Y)	8	2.5%
Val (V)	17	5.2%

Total number of negatively charged residues (Asp + Glu): 27
Total number of positively charged residues (Arg + Lys): 41

Atomic composition:

Carbon	C	1563
Hydrogen	H	2515
Nitrogen	N	465
Oxygen	O	456
Sulfur	S	15

Formula: $C_{1563}H_{2515}N_{465}O_{456}S_{15}$

Total number of atoms: 5014

Extinction coefficients
Extinction coefficient was 39795 M^{-1} cm^{-1}, at 280 nm measured in water.
Abs 0.1% (=1 g/l) 1.118, assuming all pairs of Cys residues form cystines
Ext. coefficient 39420
Abs 0.1% (=1 g/l) 1.107, assuming all Cys residues are reduced

Estimated half-life
The N-terminal of the sequence considered is M (Met).
The estimated half-life was: 30 hours (mammalian reticulocytes, in vitro).
>20 hours (yeast, in vivo).
>10 hours (Escherichia coli, in vivo).

Instability index
The instability index (II) is computed to be 46.49
This classifies the protein as unstable.

Aliphatic index: 88.95
Grand average of hydropathicity (GRAVY): -0.083

3.1.3 SOPMA

```
0         20         30         40         50         60         70
          |          |          |          |          |          |          |
MQTRRVVLKSAAARTLLGGLAGCATWLDRSAQAMRSIRARPITISEAGFTLTHEDISAARQDSCVLGQSS
Hhhheeehhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhtccceeectttceechhhhhhccccchheeccch

SVAQSSSAKGCERIARQSGWRANDCRCVDFRYRSRRQFIGRGFAGCRRSYLAATGLWFDPPLSMRLRYVE
hhhhhhhhhhhhhhhhhhhhtceeeeeeecccccchhhhhhhhhhhhhheeeehttcccccccchhhhhhh

ELTLVLPAVRFNMASKYTGIRAGIIKVATTGKATPFQELVLKAAARASLATGVPVTTHTAASQRDGERGR
hhhhhhhhhhhhhhcccttcctceeeeeecccccccchhhhhhhhhhhhhhcttcceeeecccccctttcccc

PPFLSPKLEPSRVCIGHSDDTDDLSYLTALLRGYLIGLDHIPHSAIGLEDNASASPLLGIRSWQTRALLI
checcttccccceeeecccccchhhhhhhhhhhtceeeecccccceeeectttthhheeeeccchhhhhhhh

KALIDQGYMKQILVSNDWLFGFSSYVTNIMDVMDRVNPDGMAFIH
hhhhhttchhheeecttheeehhhhhhhhhhhhhtcctttceeee
```

The Sequence length was 325 whose Alpha helix (Hh) accounts 157 amino acids of about 48.31%. The extended strand (Ee) had 58 amino acids accounting 17.85%, Beta turn (Tt) made up of 29 amino acids making up 8.92% and random coil (Cc) made up of 81 amino acids accounting 24.92%. There was no 3_{10} helix (Gg), Pi helix(Ii), Beta bridge (Bb), Bend region (Ss), Ambigous states and other states. The parameters were window width of 17 with a similarity threshold 8 and the number of states is 4 [18] (Fig.1).

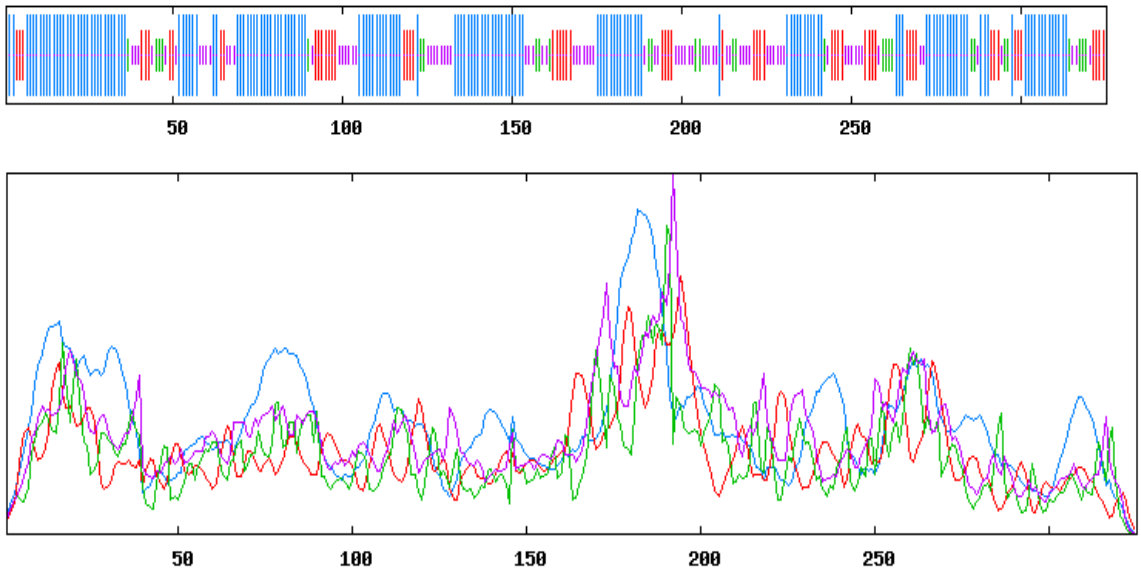

Fig.1: Significant improvement in protein secondary structure prediction by consensus prediction from multiple alignments (SOPMA)

3.1.4 PepWheel

PepWheel draws a helical wheel diagram for a protein sequence. This displays the sequence in a helical representation as if looking down the axis of the helix. It is useful for highlighting amphipathicity and other properties of residues around a helix. By default, aliphatic residues are marked with squares; hydrophilic residues are marked with diamonds, and positively charged residues with octagons, although this can be changed (Fig.2).

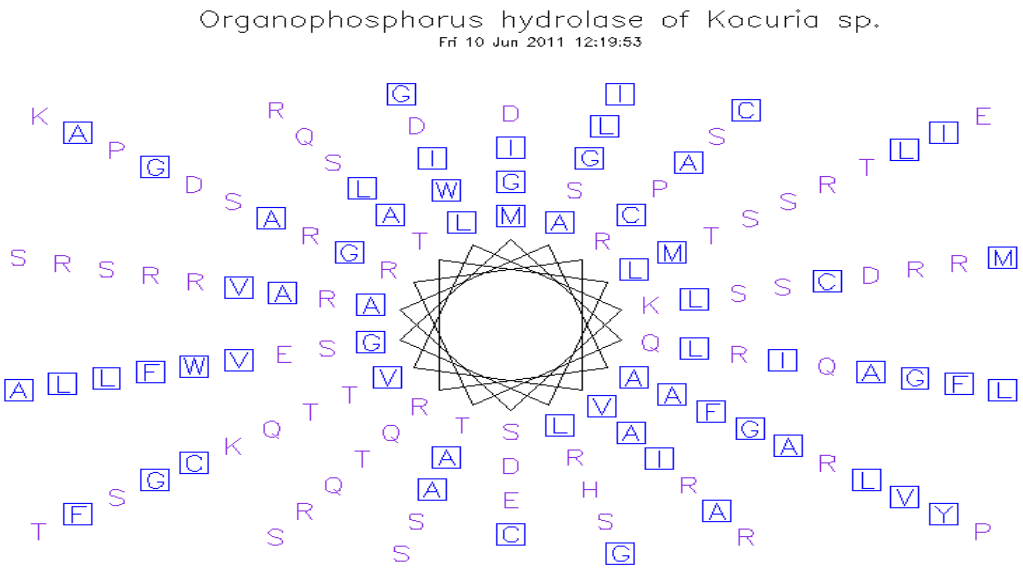

Fig.2: PepWheel for Organophosphorus hydrolase of *Kocuria* sp

3.1.5 3-Dimensional Structure determination of Organophosphorus hydrolase of *Kocuria* sp

PROSITE method (with tools and information) covered by this documentation for the active site residues Leu 146 ; Val 149; Thr 124; Gly 125; Ile 165; Leu 126; Leu142 ; Val 160; Ala 183; Thr 143; Ala 187 and Val 145. This was confirmed by performing Pfam - Protein Family Analysis (http://pfam.sanger.ac.uk/family?acc=PF00704) [19]. Alignment of Protein Sequence (Organophosphorus hydrolase of *Kocuria* sp. and template >2ob3A with chain length of 329 is as follows.

```
Score =  326 bits (836), Expect = 3e-90,   Method: Composition-based stats.
 Identities = 198/298 (66%), Positives = 214/298 (71%), Gaps = 14/298 (4%)

Query: 34  MRSIRARPITISEAGFTLTHEDISAA-----RQDSCVLGQSSSVAQSSSAKGCERIARQS 88
           + ++R  PITISEAGFTLTHE I  +     R      G   ++A   + +G  R
Sbjct: 3   INTVRG-PITISEAGFTLTHEHICGSSAGFLRAWPEFFGSRKALA-EKAVRGLRRARAAG 60


Query: 89  GWRANDCRCVDFRYRSRRQFIGRGFAGCRRSYLAATGLWFDPPLSMRLRYVEELTLV-LP 147
                D    D         +           +AATGLWFDPPLSMRLR VEELT   L
Sbjct: 61  VRTIVDVSTFDIGRDV--SLLAEVSRAADVHIVAATGLWFDPPLSMRLRSVEELTQFFLR 118


Query: 148 AVRFNMASKYTGIRAGIIKVATTGKATPFQELVLKAAARASLATGVPVTTHTAASQRDGE 207
            +++ +    TGIRAGII VATTGKATPFQELVLKAAARASLATGVPVTTHTAASQRDGE
Sbjct: 119 EIQYGIED--TGIRAGII-VATTGKATPFQELVLKAAARASLATGVPVTTHTAASQRDGE 175


Query: 208 RGRPPFLSPKLEPSRVCIGHSDDTDDLSYLTALL-RGYLIGLDHIPHSAIGLEDNASASP 266
           +    F S  L PSRVCIGHSDDTDDLSYLTAL  RGYLIGLDHIP+SAIGLEDNASAS
Sbjct: 176 QQAAIFESEGLSPSRVCIGHSDDTDDLSYLTALAARGYLIGLDHIPYSAIGLEDNASASA 235


Query: 267 LLGIRSWQTRALLIKALIDQGYMKQILVSNDWLFGFSSYVTNIMDVMDRVNPDGMAFI 324
           LLGIRSWQTRALLIKALIDQGYMKQILVSNDW FGFSSYVTNIMDVMDRVNPDGMAFI
Sbjct: 236 LLGIRSWQTRALLIKALIDQGYMKQILVSNDWTFGFSSYVTNIMDVMDRVNPDGMAFI 293
```

Then an alternative method for finding the homologous protein i.e., fold prediction method was used. There is an automated server for protein modeling which searches the homologous protein by fold prediction and sequences are modeled with high degree of accuracy. The

generated model was subjected to several repeated cycles of energy minimization using SPDBV software and the final model was subjected to stereo chemical evaluation. The fold prediction method found out Template Parathion dehydrogenase 2ob3 chain A from protein data bank of *Brevundimonas diminuta* was found to be the best homologue and modeling was carried out. The generated model was subjected to several repeated cycles of energy minimization using modeler software that is performed by satisfaction of spatial restraints and the final model was subjected to stereo chemical evaluation. [20]. After Energy minimization, the energy of the protein model is found to be -16665.717 KJ/mol that fits Ramachandran Plot (Fig.3).

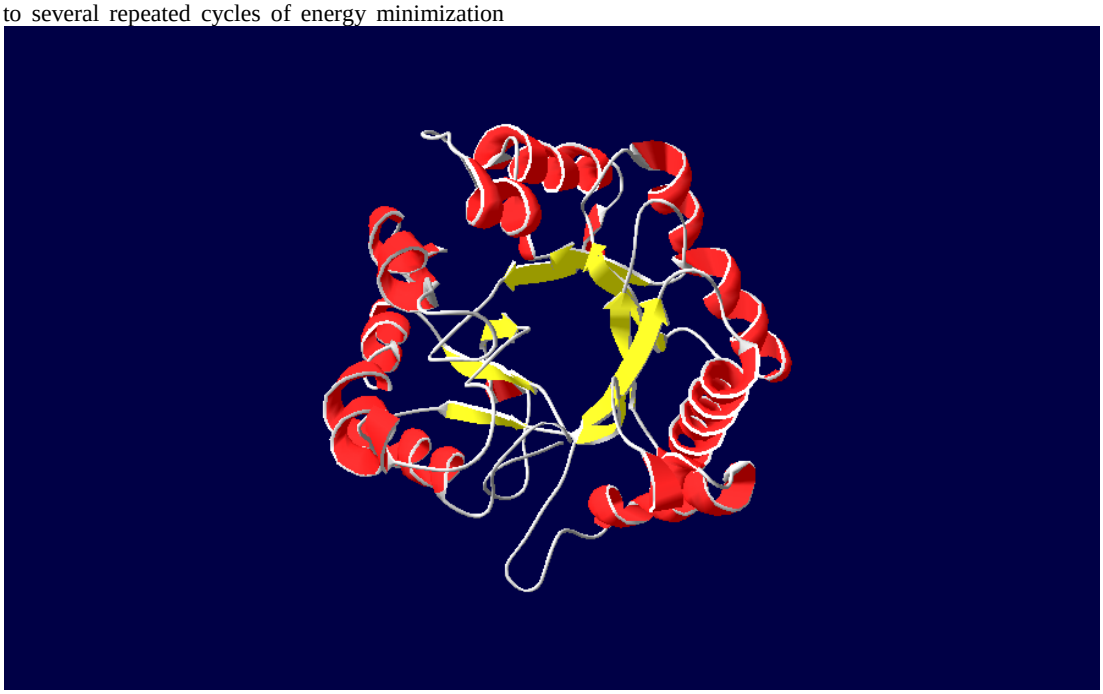

Fig.3: *In silico* 3-Dimensional Structure of Organophosphorus hydrolase of *Kocuria* sp

3.2 Docking of Chloropyrifos onto Organophosphorus hydrolase (OPH) of *Kocuria* sp.

A total of five cavities were able to detected in OPH model of *Kocuria* by using Molegro Virtual Docker and were named cav1, cav2, cav3, cav4 and cav5 with the volume details (Table.1).

The cavity 1 has shown good interaction with chloropyriphos and related organophosphorous pesticides like debantics, paraxon and phosphamidon. The moldock score for Chloropyriphos was 4093.21, Debantic/dietreen was 4099.49, Paraxon was 4993.53 and for Phosphamidon it was 5055.64. The active site residues with which chloropyriphos bind were Leu 146 ; Val 149; Thr 124; Gly 125; Ile 165; Leu 126; Leu142 ; Val 160; Ala 183; Thr 143; Ala 187 and Val 145 (Fig.4).

Table.1 : Top 5 Cavities (Active sites)

Cavity No (cav)	Volume
1	355.84
2	103.93
3	35.84
4	31.74
5	23.04

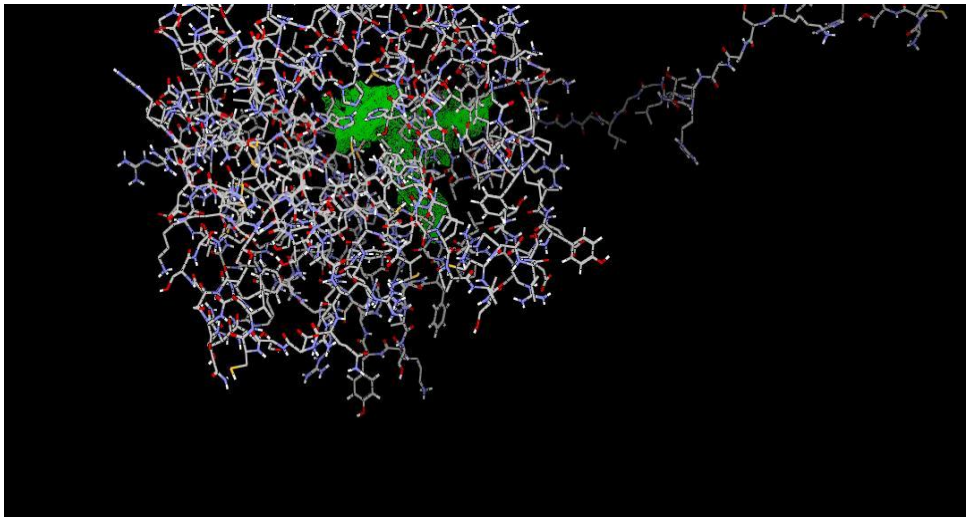

Fig.4: Docking of Chloropyrifos onto Organophosphorus hydrolase (OPH) of *Kocuria* sp.

4. CONCLUSION

The 3- dimensional structure of Organophosphorus hydrolase (OPH) of *Kocuria* sp. was predicted by using SPDBV. Later by using this model it was docked by different OP. The *in silico* model proved that this enzyme is effective in targeting OP. Therefore the same has to be correlated in wet lab and this organism can be used in degrading the OP accumulated in the soil.

5. REFERENCES

[1] Kurzel RB and Certrulo CL. (1981). The effect of environmental pollutants on human reproduction, including birth defects. *Environ. Sci. Technol.* 15: 626-631.

[2] Akhtar S nd Ahmed A. (2002). Pesticides human health and ecosystem. *J. Baqai. Med.Univ.* 5(2): 16-19.

[3] Sogorb MA and Vilanova E. (2002). Enzymes involved in the detoxification of organophosphorus, carbamate and pyrethroid insecticides through hydrolysis. *Toxicology Letters.* 128: 215-228.

[4] Qiao CL, Yan YC, Shang HY, Zhou XT, Zhang Y (2003). Biodegradation of pesticides by immobilized recombinant *Escherichia coli*. *Bull. Environ. Contam. Toxicol.* 71: 370-374.

[5] Murray RT, Vonstein C, Kennedy IR, Sanchez-Bayo F. (2001). Stability of chlorpyrifos for termiticidal control in six Australian soils. *J. Agric. Food Chem.* 49: 2844-2847.

[6] Swathi and Singh DK. (2002). Utilization of chlorpyrifos by *Aspergillus niger* and *A.flavus* as carbon and phosphorus source. 17th World Congress of soil science. Bangkok, Thailand. 14-21.

[7] Thomas K and Nicholson BC. (1989). Pesticide losses in runoff from a horticultural catchment in South Australia and their relevance to stream and reservoir water quality. *Environ. Technol. Lett.* 10: 117-129.

[8] Richards RP and Baker DB (1993). Pesticide concentration patterns in agricultural drainage networks in the lake Erie basin. Environ. *Toxicol. Chem.* 12:13-26.

[9] Giesy JP, Solomon KR, Coats JR, Dixon KR, Giddings JM, Kenaga EE (1999). Chlorpyrifos: Ecological risk assessment i n North American aquatic environments. *Rev. Environ. Contam. Toxicol*. 160: 1-129.

[10] Ragnarsdottir KV. (2000). Environmental fate and toxicology of organophosphate pesticides. *J. Geological Soc*. 157: 859-876.

[11] Wood B and Stark JD. (2002). Acute toxicity of drainage ditch water from a Washington state cranberry-growing region to *Daphnia pulex* in laboratory bioassays. *Ecotoxicol. Environ. Safety*. 53: 273-280.

[12] Galloway T and Handy R. (2003). Immunotoxicity of organophosphorous pesticides. *Ecotoxicology*. 12: 345-363.

[13] Shelton DR and Doherty MA. (1997). A model describing pesticide bio availability and biodegradation in soil. *Soc. Soil Sci. Am. J*. 61: 1078-1084.

[14] www.expasy.org/prosite

[15] Guex N and Peitsch MC. (1997). SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*, 18(15):2714-23.

[16] Thomsen R, Christensen MH. (2006) MolDock: a new technique for highaccuracy molecular docking. *J Med Chem* 49: 3315-3321.

[17] Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A. (2005). *Protein Identification and Analysis Tools on the ExPASy Server;* (In) John M. Walker (ed): The Proteomics Protocols Handbook, Humana Press. 571-607

[18] Geourjon C and Deleage G. (1995). SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. *Comput Appl Biosci*, 11(6):681-4.

[19] Sigrist CJA, Cerutti L, Hulo N, Gattiker A, Falquet L, Pagni M, Bairoch A, Bucher P. (2002). PROSITE: a documented database using patterns and profiles as motif descriptors. *Brief Bioinform.* 3:265-274.

[20] Sali A, Blundell TL. (1993). Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol.* 234(3):779-815.