Bioinformation
www.bioinformation.net
Volume 5
Issue 2
open access
Hypothesis

Comparative analysis of organophosphate degrading enzymes from diverse species

Amna Salman[1], Atefeh Taherian Fard[1], Arshan Nasir[1], Habib Bokhari[1]*

[1]Department of Biosciences, COMSATS Institute of Information Technology, Park Road, Chak Shahzad, Islamabad, Pakistan; Habib Bokhari1 Phone: +92-300-512-7684, Fax No: 0092-051-4442805, E-mail: habib@comsats.edu.pk, *Corresponding author

Received January 02, 2010; revised April 13, 2010; accepted June 08, 2010; published July 06, 2010

Abstract:
Different types of organophosphorous compounds constitute most potent pesticides. These chemicals attack the nervous system of living organisms causing death. Different organisms produce enzymes to degrade these chemicals. These enzymes are present in simple microorganisms from archaea, bacteria to complex eukaryotes like humans. A comparison of representative eight shortlisted enzymes involved in the degradation and inactivation of organophosphates from a wide range of organisms was performed to infer the basis of their common functionality. There is little sequence homology in these enzymes which results in divergent tertiary structures. The only feature that these enzymes seem to share is their amino acid composition. However, structural analysis has shown no significant similarities among this functionally similar group of organophosphate degrading enzymes.

Keywords: Pesticide, Organophosphates, Organophosphatases, Common functionality, Sequence and Structure comparison.

Background:
The excessive use of natural resources and large scale synthesis of xenobiotic compounds have raised a number of environmental issues like air and water contamination along with harmful effects on different biota and disruption of the biogeochemical cycle. Presently, the most widely used pesticides belong to the group of organophosphates and consequently the organophosphorus compound poisoning is a worldwide health problem with approximately 3 million poisonings and 200,000 deaths annually **[1,2].**

These compounds have been implicated in several nerve and muscular diseases in humans. Some of the early organophosphates were developed as nerve poisons for human warfare. The organophosphates recommended for non-residential uses are relatively toxic to vertebrates. Their primary mode of action on insects and other animals is by the phosphorylation of acetylcholinesterase enzyme which results in the accumulation of acetylcholine. Higher levels of acetylcholine result in sensory and behavioral disturbances in coordination and depressed motor function. Symptoms such as muscle twitching, weakening, vomiting, abdominal cramps and diarrhea all indicate a worsening condition. Recovery from organophosphate exposure depends upon production of a new enzyme which may further be complicated due to lack of paraoxonase in some individuals **[3, 4].**

Organophosphates are widely used as pesticides today. They are quite effective in controlling pests but exposure to organophosphates not only damages the predators of plants but also adversely affects all sorts of other living organisms in the surroundings. Living organisms are equipped with enzymes that degrade these highly toxic compounds. These organophosphate degrading enzymes are found in many different living organisms ranging from bacteria to human beings **[5, 6]** like Parathion hydrolases (OPH) in Flavobacterium species, Organophosphorous acid anhydrolase in Alteromonas species, Phosphotriesterase homology protein in Escherichia coli, Diisopropyl fluorophosphatase (DFPases) in Squid, Serum Paraoxanases / arylesterases **[1, 2]** in Mammals and Prolidases in Mammals and other organisms **[7].** Though these enzymes are present in different organisms but they still perform the same function. Here, we carried out a comparative study of these enzymes from different organisms in order to analyze these enzymes for their common features that enable them to maintain their functional conservation.

Methodology:
Sequence retrieval:
Organophosphate degrading enzymes from different organism were identified after careful literature review and sequences retrieved from Swiss Prot & TrEMBL databases [http://www.expasy.org/] using keywords such as parathion hydrolases, organophosphatases and organophosphate degrading enzymes (OPD). A total of sixty different enzymes that are involved in organophosphate degradation were retrieved. These sequences were then used independently as a query to search for their homologs using the Basic Local Alignment Search Tool (BLAST) at the National Centre of Biotechnology Information database [http://blast.ncbi.nlm.nih.gov/Blast.cgi]. Eight subgroups were created based on the type of organism in which these enzymes are present and a representative sequence from each of the subgroup was selected for further analysis. As each subgroup showed a high degree of intra group homology therefore one representative sequence was considered adequate.

Multiple sequence alignment:
Representative sequences were aligned using ClustalW that generates multiple sequence alignment file for a set of given sequences. [http://www.ebi.ac.uk/Tools/clustalw/].

Search for motifs:
MEME was used to search for the presence of conserved motifs among these representative sequences. [http://meme.sdsc.edu/meme/meme-download.html].

Secondary structure analysis:
Secondary structures for those enzymes whose structures were not already available in the Protein Data Bank (PDB) were predicted using PSIPRED [http://bioinf.cs.ucl.ac.uk/psipred].

ISSN 0973-2063 (online) 0973-8894 (print)
Bioinformation 5(2): 67-72 (2010)
© 2010 Biomedical Informatics

Bioinformation **Volume 5**
www.bioinformation.net Issue 2 **Hypothesis**
open access

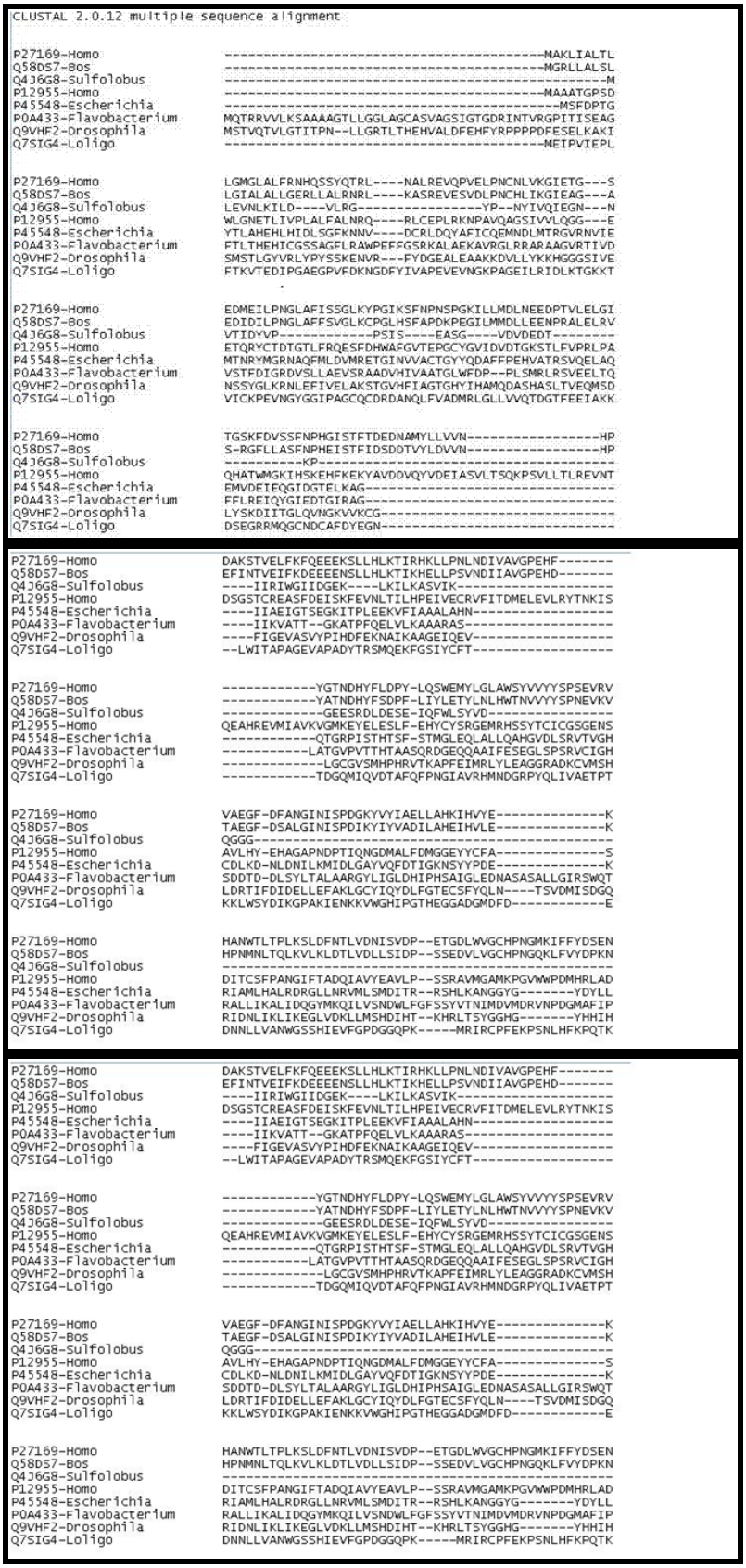

Figure 1: Alignment file generated by ClustalW for eight organophosphate degrading enzymes. No significant homology is observed overall and poorly aligned regions at the ends of proteins can be easily seen.

ISSN 0973-2063 (online) 0973-8894 (print)
Bioinformation 5(2): 67-72 (2010)
© 2010 Biomedical Informatics

Bioinformation **Volume 5**
www.bioinformation.net Issue 2 **Hypothesis**
open access

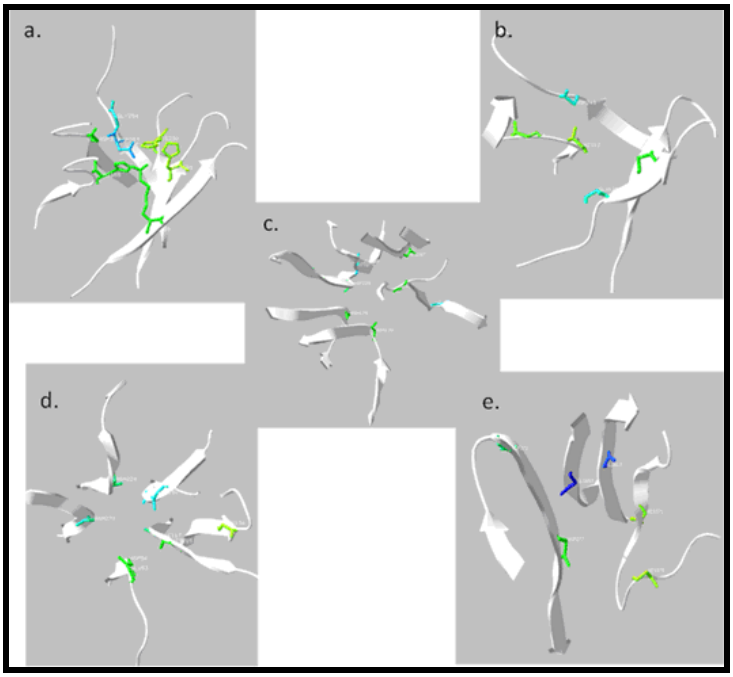

Figure. 2. Active sites of the five experimentally verified structures out of the eight representatives. These structures were retrieved from Protein databank. The active site residues of a. Bacterial PH{PDB ID1P6C} are H55, H7, K169, H201, H230, D253, D301, H254, b. Bacterial PHP{PDB ID 1P6B} are H12, H14, E125, H158, D243, c. Squid DFPase {PDB ID 2IAR} are E21, E37, N120, N175, D229, D232, L273, H274, H287,d.Mamalian Pon1 {PDB ID 1V04} E53, D54, H115, I117, H134, N168, I170, N224, N270,and e.Mammalian Prolidase {PDB ID 2IW2} are D276, D287, H370, E412, E452, H37).

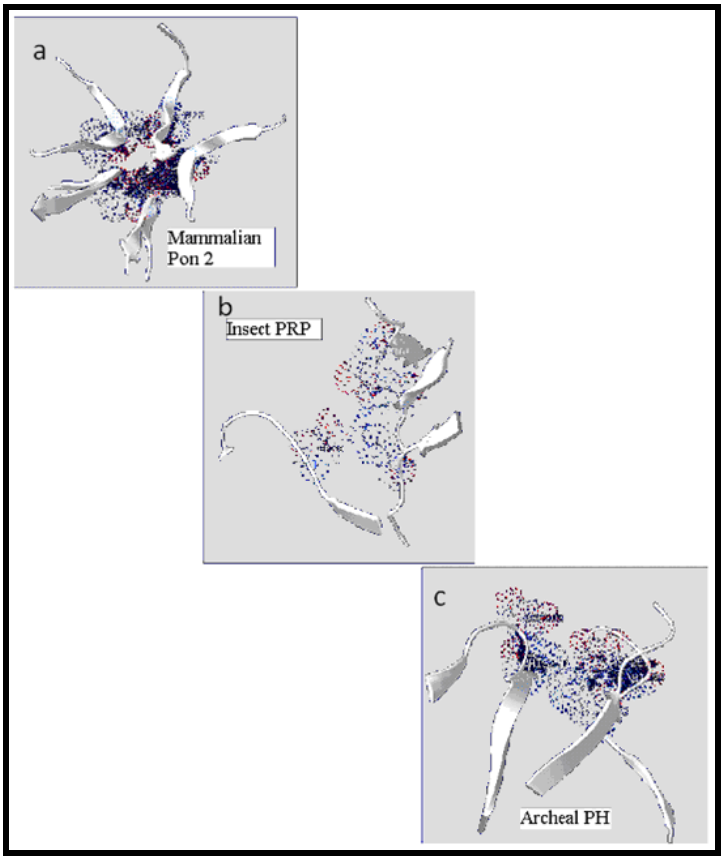

Figure 3: Active sites of the three structures out of the eight representatives that have been predicted using Swiss PDB viewer. The table with the figure shows the active site residues. Active site residues of a. Mamalian pon 2 (E53, D54, H114, I116, H133, N167, I169, N223, N269), b. Insect PRP (H22, H24, E169, H201, H230, D298) and c. Archeal PH (H23, H25, H200, D203, D257).

ISSN 0973-2063 (online) 0973-8894 (print)
Bioinformation 5(2): 67-72 (2010)
© 2010 Biomedical Informatics

Bioinformation
www.bioinformation.net
Volume 5
Issue 2
open access
Hypothesis

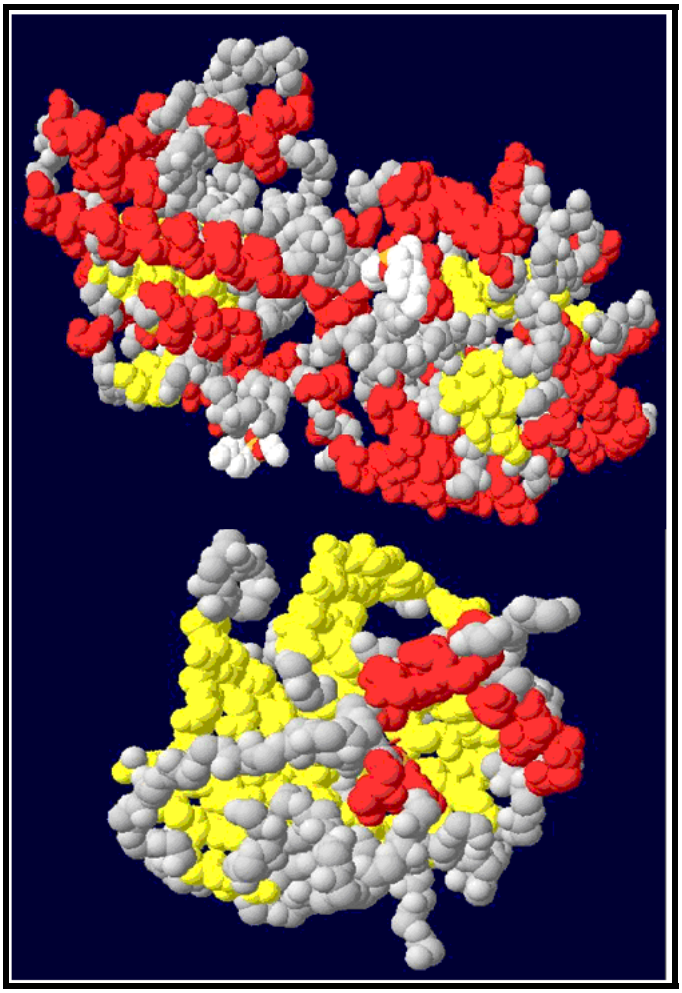

Figure 4: Tertiary structures of two enzymes, Parathion hydrolase precursor (Accession No: P0A433) and Serum paraoxonase/arylesterase 1(Accession No: P27169). Content of alpha helices (colored red) is greater in the Flavobacterium enzyme as compared to the human enzyme. Beta sheets are colored yellow.

Tertiary structure analysis:
Tertiary structures of two of the eight enzymes, Parathion hydrolase precursor (Accession No: P0A433) and Serum paraoxonase/arylesterase 1(Accession No: P27169) were retrieved from the Protein databank (PDB) [http://www.pdb.org/pdb/home/home.do] and visualized using the Swiss PDB viewer [http://ca.expasy.org/spdbv].

Composition analysis:
The detailed analysis of the amino acid composition and properties of the chosen proteins were calculated using AminoComp software [http://144.16.71.10/thgs/aminocomp.html].

Glycosylation & Phosphorylation Site Prediction:
O-linked Glycosylation sites in selected proteins (**Table 1 see supplementary material**) were predicted by NetOGlyc whereas N-linked Glycosylation sites were predicted using NetNGlyc 1.0 Server. For the prediction of phosphorylation sites, NetPhosBac was used. (http://www.cbs.dtu.dk/services/NetPhosBac-1.0).

Results and Discussion:
Sequence retrieval:
Sequences of sixty (60) enzymes that are involved in the degradation of toxic organophosphates were retrieved from the SwissProt and TrEMBL databases. The initial data was subjected to BLAST search in order to retrieve homologous sequences. Once these searches were completed, data was organized into eight homologous groups and a representative sequence was chosen to correspond to each group. **Table 1 (see supplementary material)** enlists the selected sequences.

Multiple sequence alignment:
All the representative sequences were aligned using ClustalW but no significant similarity was found in these sequences specifically the ends of proteins are very poorly aligned (**Figure 1**). In addition, no conserved motif was reported by MEME. Furthermore the sequence analysis showed that the active sites in all these enzymes do not possess conserved amino acid sequences and consequently they give rise to a different active site structure. This can be seen in Figure 2 and Figure 3. Therefore, the possibility of diverse members of this family of enzymes possessing identical motifs can be ruled out. Although in most cases, it is the sequence homology that results in the functional conservation but exceptions to this rule do exist. Examples of this exception are the enzymes GOAG_ECOLI (aminotransferase, E.C. 2.6.1.19) and GSA_ECOLI (glutamate- 1-semialdehyde aminomutase, E.C. 5.4.3.8). Both these enzymes have very little sequence identity but both function as aminotransferases **[14].** Both these enzymes function with the help of a similar cofactor pyridoxal phosphate (PLP). The organophosphate degrading enzymes under study also show very little sequence similarity and are known to have similar divalent metal cofactors.

Secondary and tertiary structure analysis:
Secondary structures were predicted using PSIPRED and it was observed that some of these enzymes seem similar to others but no global

ISSN 0973-2063 (online) 0973-8894 (print)
Bioinformation 5(2): 67-72 (2010)
© 2010 Biomedical Informatics

conservation was found. This can be seen in **Table 1 (see supplementary material)**. Like the secondary structure, tertiary structure analysis showed no characteristic similarity between the two enzymes studied which are serum paraoxonase from humans and Parathion hydrolase precursor from Flavobacterium. Beta sheets are the dominant secondary structure in the human enzyme whereas Flavobacterium enzyme is rich in alpha helices (**Figure 4**).

Studies of the bacterial, mammalian and plant purple acid phosphatases have shown that they do not have any significant sequence homology but the structures of these enzymes are similar. Tertiary structure comparison was carried out using the root mean square values to assess the structural homology between these enzymes. This showed some degree of similarity but could not account for the overall homology of the functioning of these enzymes. This may be caused by the similar amino acid composition of these enzymes **[15].**

Composition analysis:
Analysis of the overall composition of amino acids yielded significant results. The fraction of hydrophobic and hydrophilic residues in these enzymes seemed to be quite conserved. Moreover they also possess the same isoelectric points due to this reason. This is shown in **Table 2 (see supplementary material)**.

Since amino acid composition can play a significant role in substrate selection **[15]** therefore, one of the reasons for the hydrolysis of the same substrate by these enzymes may be the similar amino acid composition. Moreover this may be the basis of the little homology that was observed among some of the enzymes at the secondary and tertiary structure level.

Glycosylation & Phosphorylation Site Prediction Results:
It has been shown that many bacterial enzymes are glycosylated, however glycosylation does not seem to play any critical role in enzymatic activity **[16, 17],** whereas on the other hand it has been shown to enhance the thermostability in a number of enzymes. **[17].** Potential O-glycosylation sites occur less frequently as they were shown to be present only in two proteins i.e. Xaa-Pro dipeptidase (P12955) and Parathion hydrolase precursor (P0A433). The results of NetNGlyc show that there is a relatively higher level of such sites which are N-glycosylated. Potential N-glycosylation sites were seen in all cases studied except Parathion hydrolase precursor (P0A433) and Phosphotriesterase (Q7SIG4). Results of NetOGlyc and NetNGlyc are given in **Table 3 (see supplementary material)**. The occurrence of all three sites in a single protein was also noted in Mammalian prolidase (P12955). Moreover, potential phosphorylation sites were predicted in all proteins and serine amino acid was found to be the most frequent in N-X-S/T motif.

Post-translational modification is a chemical modification that alters the physical and chemical properties of proteins for example: folding, conformation distribution, stability, activity, therefore, these modifications play an important role in the regulation of various intracellular processes. Glycosylation has been found to alter the function of both eukaryotic and prokaryotic proteins **[18, 19].** In agricultural areas, where organophospahtes are the potential risk factor for developing various diseases, it is important to understand various enzymes involved in the detoxification/biodegradation of these toxic materials. For example, recently, op-pesticide degrading bacteria have received considerable attention as the potential agents of bioremediation in clean-up operations of op-pesticide polluted environments and industrial waste **[20, 21].** All these soil isolates synthesize an enzyme called parathion hydrolase (PH) and it would be interesting to gain information using available data with the help of bioinformatics tools.

Conclusion:
We have concluded that the organophosphate degrading enzymes do not have significant sequence similarity however they show some similarity at the structural level. The content of alpha helices and beta sheets in a few of the enzymes is conserved. In tertiary structures, the active site is surrounded by beta sheets in all of them but overall all the structures do not have significant resemblance. However the amino acid content of all of these enzymes is quite similar and conserved. It has been verified that the amino acid content of proteins plays a major role in their similar functioning. Therefore, we believe that the basis for the analogous functioning of these sequentially divergent enzymes is due to the compositional homology of these enzymes. Moreover, it would be interesting to carry out wet lab work to assess the role of each potential glycosylation site by mutagenesis study on the averall activity of the protein.

Acknowledgements:
We thank the Higher Education Commission (HEC) of Pakistan and COMSATS Institute of Information Technology (CIIT), Islamabad for all the support and facilities provided to carry out this work.

References:
[1] M Eddleston *et al. Lancet* **12**:1163 (2002) [PMID: 12387969]
[2] MA Sogorb *et al. Toxicol Lett* **151**: 219 (2004) [PMID: 15177657]
[3] HA Ghanbari *et al. Aging Cell* **3**: 41 (2004) [PMID: 14965355]
[4] NM Reddy *et al. Oncogene* **27**: 5821 (2008) [PMID: 18542053]
[5] WW Mulbry *et al. Appl Environ Microbiol* **5**: 926 (1986) [PMID: 3015022]
[6] DI. Draganov *et al. J Lipid Res* **46**: 1239 (2005) [PMID: 15772423]
[7] G Amitai *et al. FEBS* **273**:1906 (2006) [PMID: 16640555]
[8] http://ca.expasy.org/
[9] http://www.ebi.ac.uk/Tools/clustalw/
[10] http://meme.sdsc.edu/meme/meme-download.html.
[11] http://bioinf.cs.ucl.ac.uk/psipred/
[12] http://ca.expasy.org/spdbv/
[13] http://144.16.71.10/thgs/aminocomp.html
[14] K Pawlowski, *et al. Pacific Symposium on Biocomputing* **5**: 42 (2000).
[15] G Schenk, *et al. Gene* **255**: 419 (2000).
[16] TH Jr Plummer *et al. J Biol Chem* **270**: 13192 (1995).
[17] O Olsen & KK Thomesen, *J Gen Microbiol*, 579 (1991).
[18] L Tzong-Yi *et al. Nucleic. Acids Res.* **34** : 622 (2006)
[19] VX Bisaria & S Mishra, *Crit Rev Biotech* **9**, 61 (1989).
[20] GLJAislabie Aust J Soil Res **33**: 925 (1995).
[21] KN Timmis *et al. Ann Rev Microbiol.* **48**: 525 (1994).

Edited by P. Kangueane
Citation: Salman *et al.* Bioinformation 5(2): 67-72 (2010)
License statement: This is an open-access article, which permits unrestricted use, distribution, and reproduction in any medium, for non-commercial purposes, provided the original author and source are credited.

ISSN 0973-2063 (online) 0973-8894 (print)
Bioinformation 5(2): 67-72 (2010)
© 2010 Biomedical Informatics

Bioinformation Volume 5
www.bioinformation.net Issue 2
open access
Hypothesis

Supplementry material:

Table 1: Chosen representative sequences of organophosphate degrading enzymes used in this study from the different groups that were created through the BLAST searches using query sequence retrieved from the literature. A representative from each group was selected whose accession number, length, cofactor, and helix and strand composition is given.

Group	Accession #	Length	Cofactor	Helix	Strands
Archeal PH	Q4J6G8	96	Zinc	13	10
Bacterial PHP	P45548	292	Zinc	12	10
Bacterial PH	P0A433	365	Zinc	16	10
Insects PRP	Q9VHF2	350	Zinc	16	10
Squid DFPases	Q7SIG4	314	Calcium	-	29
Mammals pon1	P27169	355	Calcium	4	25
Mammals pon2	Q58DS7	354	Calcium	2	28
Mammalian Prolidase	P12955	493	Manganese	14	20

Table 2: Amino acid composition analysis of each group.

Group	Accession #	pI	% NEUTRAL			% CHARGED		
			Polar	Non Polar	Total	Acidic	Basic	Total
Archeal PH	Q4J6G8	6.11	27.95	43.48	71.43	13.65	14.92	28.57
Bacteria PHP	P45548	5.99	34.24	40.43	74.67	13.01	12.32	25.33
Bacterial PH	P0A433	6.18	30.95	47.4	78.35	9.59	12.06	21.65
Insects PRP	Q9VHF2	6.08	32.19	40.29	72.48	12.28	15.24	27.52
Squid DFPase	Q7SIG4	6.00	32.49	40.13	72.62	14.33	13.05	27.38
Mammals *pon*1	P27169	5.92	33.24	43.95	77.19	12.11	10.7	22.81
Mammals *pon*2	Q58DS7	5.97	29.93	45.2	75.13	12.43	12.42	24.85
Mammalian prolidase	P12955	6.04	31.23	42.3	73.44	12.98	13.58	26.56

Table 3: Glycosylation & Phosphorylation prediction.

Group	**Species**	**Accession #**	**Protein Name**	**Superfamily**	***1Pot.Gly. sites #**	***2Pot.Pho. sites #**
Archeal PH	Sulfolobus acidocaldarius	Q4J6G8	Hypothetical protein	Phage tail proteins	1	6
Bacterial PHP	*Esherichia. coli*	P45548	Phosphotriesterase homology protein	Metallo-dependent hydrolases	2	13
Bacterial PH	Falvobacterium	P0A433	Parathion hydrolase precursor	Metallo-dependent hydrolases	2	18
Insects PRP	Drosophila melanogaster	Q9VHF2	Phosphotriesterase-related protein	Metallo-dependent hydrolases	2	17
Squid DFPases	loligo vulgaris	Q7SIG4	Phosphotriesterase	Calcium dependent phosphotriesterase	0	7
Mammals pon1	Homo sapiens	P27169	Serum paraoxonase/arylesterase 1	Calcium dependent Phosphotriesterase	4	14
Mammals pon2	Bos Taurus	Q58DS7	Serum paraoxonase/arylesterase 2	Calcuim dependent Phosphotriesterase	4	16
Mammalian Prolidase	Homo sapiens	P12955	Xaa-Pro dipeptidase	Creatinase/prolidase N-terminal domain	3	20

*1Pot. Gly = Potential Glycosylation
*2Pot. Pho = Potential Phosphorylation

ISSN 0973-2063 (online) 0973-8894 (print)
Bioinformation 5(2): 67-72 (2010)
© 2010 Biomedical Informatics