Structural and Functional Characterization of Novel Phosphotyrosine Phosphatase Protein from *Drosophila Melanogaster* (Pupal Retina)

Rubina Naz[1], Anwar Iqbal[2], Asma Saeed[3], Alamzeb Khan[4], Meshal Shutaywi[5], Adriana Lavinia Cioca[6], Narcisa Vrinceanu[7*]

[1] Institute of Chemical Sciences, Gomal University, Dera Ismail Khan, Pakistan

[2] Department of Chemical Sciences, University of Lakki Marwat, Lakki Marwat, Pakistan.

[3] Center of Biochemistry and Biotechnology, Gomal University, Dera Ismail Khan, Pakistan.

[4] Department of Pediatrics, Yale School of Medicine, Yale University, New Haven, CT, 06511, USA.

[5] Department of Mathematics, College of Science & Arts, King Abdulaziz University, P.O. Box 344, Rabigh 21911, Saudi

[6] CMI Dr. Cioca Adriana-Lavinia, 22 Dorobantilor Street, Sibiu 557260, Romania.

[7*] Faculty of Engineering, Department of Industrial Machines and Equipments, "Lucian Blaga" University of Sibiu, 10 Victoriei Boulevard, 550204, Sibiu, Romania.

Corresponding Author: Narcisa Vrinceanu

Address: Faculty of Engineering, Department of Industrial Machines and Equipment, "Lucian Blaga" University of Sibiu, 10 Victoriei Boulevard, 550204, Sibiu, Romania.

Contact: vrinceanu.narcisai@ulbsibiu.ro, Tel: +40 721428641

Abstract

A novel pair of protein Tyrosine Phosphatases in *Drosophila Melanogaster* (pupal retina) has been identified. Phosphotyrosyl protein phosphatases (PTPs) are structurally diverse enzymes increasingly recognized having fundamental role in cellular processes including effects on metabolism, cell proliferation and differentiation. This study presents comparative homology modeling of low molecular weight phosphotyrosine protein phosphatase (PTPs) from *Drosophila melanogaster* (Dr-PTPs) and their complexation with potent inhibitor HEPES. The 3D structure was predicted based on sequence homology with bovine heart low molecular weight PTPs (Bh-PTPs). The sequence homology is approximately 50% identical to each other and to low molecular weight protein tyrosine phosphatases (PTPs) in other species. Comparison of the 3D structures of Bh-PTPs and Dr-PTPs (primo-2) reveals a remarkable similarity having a four stranded central parallel β sheet with flanking α helices on both sides, showing two right-handed β-α-β motifs. The inhibitor shows similar binding features as seen in other PTPs. The study also highlights the key catalytic residues important for target recognition and PTPs activation. The structure guided studies of both proteins clearly reveal a common mechanism of action, inhibitor binding at the active site and will expected to contribute towards the basic understanding of functional association of this enzyme with other molecules.

Keywords: PTPs; Drosophila melanogaster; homology modeling; sequence homology; enzyme substrate interactions;

1. Introduction

Low molecular weight phosphotyrosine protein phosphatases (PTPs), previously known as low molecular weight acid phosphatases, catalyzes the hydrolysis of tyrosine phosporylated proteins, low molecular weight aryl phosphates and natural and synthetic acyl phosphates [1,2]. Although the activity of PTPs on serine and threonine phosphorylated proteins are very poor with the exception of flavin mononucleotide (FMN) [3,4]. Tyrosine phosphorylation plays a vital role in the regulation of the variety of developmental processes. These processes include several cell functions like growth, cell differentiation, metabolism, cell cycle and cyto-skeletal functions. Furthermore, the phosphorylation state of tyrosine and seγ/thr of signaling proteins are controlled via specific reaction. Thus, the phosphorylation state is controlled by a very dynamic way to avoid severe malfunction of cell [5].

PTPs can act as tumor suppressor by inhibiting cell growth. Functionally two types of PTPs sequences are conserved and well distinguished by structure comparison. The one known as classical PTPs are specific for tyrosine residues

and other with the dual-specificity phosphates (DSPS) are essential for serine and threonine dephosphorylation. The active site (C-(X)$_5$-R) of these PTPs contains several conserved cysteine and arginine, important for catalyzing phosphorylation processes, and thus plays a vital role in regulation of signal transduction. All these acid phosphatase enzymes share little sequence homology, different range of molecular weight (18 kDa or above), but exhibit same catalytic mechanism [6,7]. The structural features of low molecular weight PTPs comprises relatively different fold comparing from fission yeast to mammals [8,9]. The overall three dimensional structural features contains four B sheets at center and surrounded by a helices [6,10,11]. However, several similarities can be seen in structural features and binding side pockets of PTPs to Mr-PTPs [10–15]. Importantly, the conserved the p-loop stabilized by complex hydrogen network and favor the phosphate trigonal bipyramidal transition state geometry [16–19]. Thus all low and higher Mr-PTPs and PTPs shares identical catalytic mechanism [20]. The enzymatic reaction is triggered by the first cysteine, where the substrate binding at active site is stabilized by the p-loop residues via hydrogen bonding and three anionic oxygen atoms. These transient interaction orient phosphorous atom in feasible position for nucleophilic attack and favors the enzymatic phosphorylation reaction [20–22]. The nucleophilic attack of thiol (Sγ) group takes place in the presence of proton donor aspartic acid and thus phosphor-cysteine intermediate is formed [23–25]. The formation of phosphor-cysteine intermediate is also favored and stabilized by the presence of p-loop where several residues are involved in binding and lowering the activation energy [20]. In the subsequent step, hydrolysis of phosphor-enzyme intermediate complex via attack of water takesplace, resulting in the liberation of inorganic phosphate (Pi). The enzymatic phosphorylation via hydrolysis works well at wide range of pH 5.5-7.5 from substrate.

The structural details of proteins are important parameters to understand the reactivity and stability of proteins. Several advance techniques like X-ray crystallography, nuclear magnetic resonance (NMR) and Electron Microscopy are frequently used to determine the structure of proteins. However, theoretical approaches like comparative modeling often used as a useful alternative to other biophysical and analytical techniques by providing insights into structural and functional aspects of proteins. In the current project, the comparative modeling technique has been used for the 3D structural prediction of sequence emerging from Primo-2 of *Drosophila melanogaster* (fruit fly) classified as low molecular weight PTPs family (LMW-PTPs). The present work is designed to elaborate the prediction of evolutionary context of the sequence homological ancestors, structural aspects and active site conformational states of *Drosophila Melanogaster* PTPs (Dr-PTPs) and spatial geometry formation of the active site. Dr-PTPs shares

46% amino acid sequence identity with that of Bh-PTPs (PDB: 1DG9) particularly in active site regions.

2. Results and discussion

2.1. Sequence Analysis

Two novel pairs of proteins tyrosine phosphatases were identified in the drosophila pupil retina. It was found that the primary sequences were approximately 50% identical and characterized as low molecular weight protein phosphatase (Fig. 1&2). The first sequence was incorporated at primo-1 (155 amino acid) and the other encoded by primo-2 (164 amino acid).

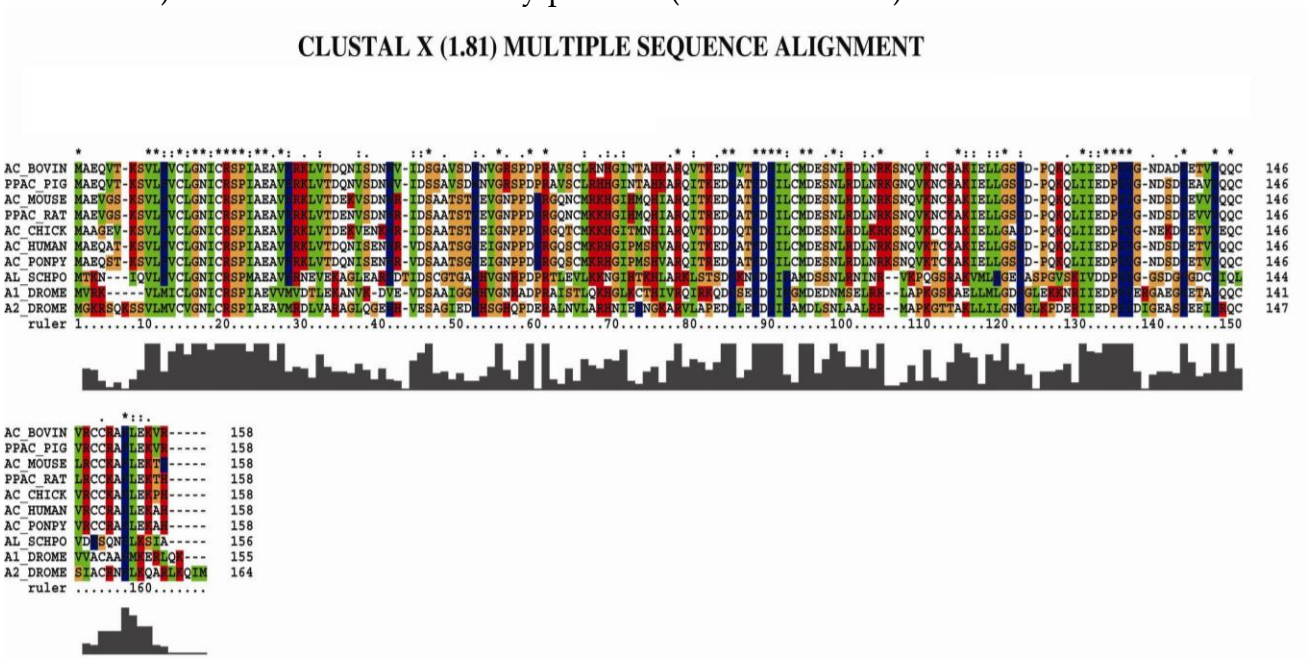


Figure 1. Multiple sequence alignment of 10 PTPs sequences. Sequences are named as Swiss-Prot entery first letter represents gene, the second part represents the biological source of gene. The symbol ''*'' represents strongly conserved '' .'' represents weakly conserved'' :''represents identical residues

```
                10        20        30        40        50        60
1DG9    AEQV-TKSVLFVCLGNICRSPIAEAVFRKLVTDQNISDNWVIDSGAVSDWNVGRSPDPRAVSCLRNHG
PTPs    GKRSQKSSVLMVCVGNLCRSPIAEAVMRDLVARAGLQGEWHVESAGIEDWHSGHQPDERALNVLARHN
               *** ** ** ********* * **         *   *    **  *  ** **   *  *

          70        80        90        100       110       120       130
1DG9    INTAHKARQVTKEDFVTFDYILCMDESNLRDLNRKSNQVKNCRAKIELLGSYDPQ-KQLIIEDPYYG-
PTPs    IEYNGKARVLAPEDFLEFDYIFAMDLSNLAALRRMAP--KGTTAKLLILGNFGLKPDERIIEDPYYDI
        *    ***    ***  ****  ** ***  * *     *   **   **          *******

          140       150       160
1DG9    NDADFETVYQQCVRCCRAFLEKVR-----
PTPs    GEASFEEIYRQCSIACRNFLKQARLKQIM
          * **  * **   ** **   *
```

Figure 2. Pairwise sequence alignment used for building model of Drosophila phosphatase. Target sequence represented by PTPs. based on the structure of Bovin heart phosphatase template structure represented by1DG9.

123

124 2.2. Structure Topology

125 The structure of Dr-PTPs (primo-2) comprises a fold containing four central β
126 parallel sheets gathered by a-helices: a right-handed β-*α*-β motif. The conserved
127 sequence known as active site C—(X)$_5$—R(S/T) of acid phosphatases was present
128 as sequence CVGNLCRS in Dr-PTPs. The active site was present in the form of
129 loop extending from between β 1 and helix *α*1 (Fig. 4).

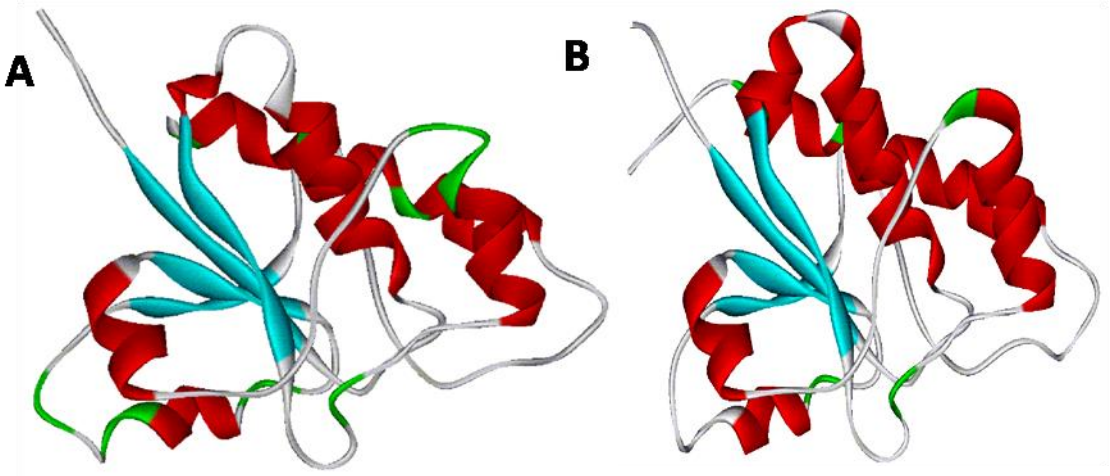


Figure 3. A complete Ribbon diagram of low molecular weight phosphotyrosine protein phosphatase. Showing *α*/β Proteins, *α* helix (Red), β Sheets (blue).(A) 1BVH(template), (B) Drosophila(target).

130

131 The crystal structure (PDB: IDG9) shows a complex of Bh-PTPs with HEPES,
132 where the active site residues were interacting with the sulfonate moiety,
133 accumulating phosphate binding manner in the pocket. All the oxygen atoms of

sulfonate group was involved in complex hydrogen bonding network, offered by N, H and S of the backbone residues of the active site loop and conserved arginine (Arg 18). The structural features of sulfonate complex were found similar to phosphates bounded at active sites. In our case, the highest sequence similarity of Dr-PTPs (Primo-2) primary sequence at substrate attracting site C-(X5)-R, tyrosine phosphorylation site RIIEDPYY was found. The 3D structure of Dr-PTPs (Primo-2) was superimposed at Bh-PTPs (Pdb: 1DG9). It was found the both structures aligned well for several motifs. However, regions like Gly 1, Lys 6, Pro 105, Thr 108, Ile 134, Glu 136, Lys 121, Asp 123 and Leu 159, Met 163 of the Dr-PTPs structural orientation was different from the target. Furthermore, the sequence analysis on the basis of multiple sequence alignment and the construction of phylogenetic tree (PHYLIP package) shows that both Dr-PTPs and Bh-PTPs belong to common ancestors. They are homologous sequences and members of same subfamily as per evolutionary context (Fig. 4).

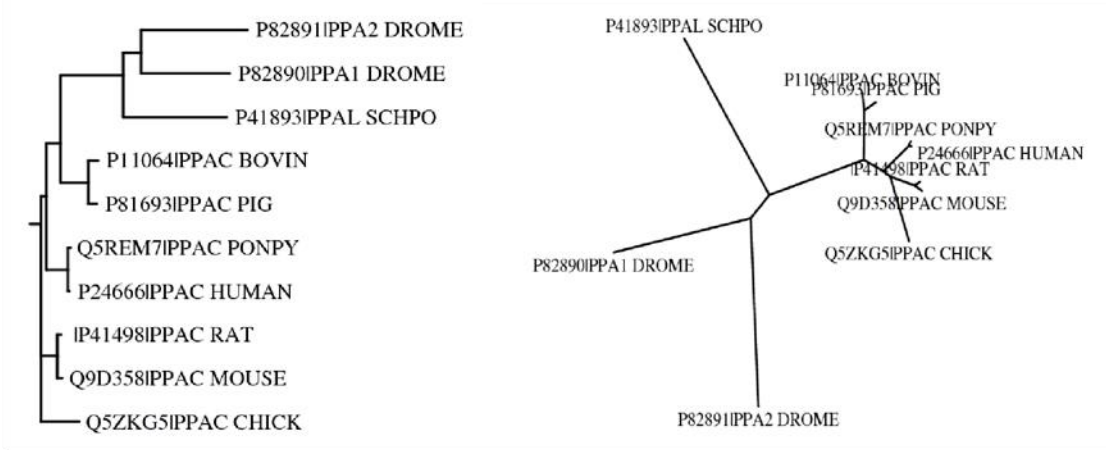


Figure 4. Dendrogram of PPA2 *Drosophila Melanogaster* and related proteins made by PHYLIP & depicts relationship among various forms of PPAC,PPAI, PPA2 genes. The first part of the code represents gene where as the second part of the code represents the source of protein.

2.3. Structure Topology

The overall secondary and tertiary structure of the Drosophila phosphatase strongly resembles to templates 1BVH, 1DG9, 1PNT (Bovin heart) and other low molecular weight phosphotyrosine proteins phosphatase (Fig 3). The 3D structure was characterized with the active site end at the α1 and situated close to the N terminal region followed by the P-loop and β1 strand. The active site emerged as deep groove encompassing aromatic residues ((Trp-49, Tyr-131, Tyr-132) appears like claws while the conserved residues Asp-56 and Arg 58 of variable loop provides a network of hydrogen bonds with other adjacent α5-helice and β4-strand. All these residues work together for the target recognition

and sets a proper orientation of ligand for catalytically important residues Cys12, Cys17 and Arg18, whether Asp129 (β4 extend loop) on the other end of the pocket to facilitate the protonation of phosphorylated intermediate together with Tyr131 and Tyr132. The hydrophobicity of active site groove is maintained by several buried hydrophobic residues like Leu 9, Val11, Phe82, Ile 88, Leu 99 and Lys 102.

2.4. Quality of model

The different aspects of structure of Dr-PTPs were validated using different tools. The stereochemical outcomes were analyzed using software PROCHECK where the restraints obtained were compared to the stereochemical properties of Bh-PTPs. The degree of violation of secondary structure elements were evaluated using ramachandran plot where 94% regions were found in allowed region and no dihedral region in disallowed regions confirms the validity of the model.

The 10 best structures of the Dr-PTPs were compared with Bh-PTPs crystal structure, both in free and complexed with HEPES. The RMSD based on backbone (α-carbon) were found 0.26 Å and 0.49 Å in the presence and absence of complexation (Table 1), respectively. These observations further confirmed the validity of the model beside the higher sequence identity. However, the conformational variability can be seen in several regions (1-6, 105-108, 134-136 and 121-123) due to presence of different amino acids inducing different orientation (Fig. 7a).

Table 1.

Template	Target	DRMS
Bovin heart PTPs (1BVH) with out inhibitor	Drosophila PTPs	0.4840
Bovin heart PTPs (1DG9) with inhibitor (HEPES)	Drosophila PTPs	0.2549
Bovin heart PTPs (1PNT) with inhibitor (PO4)	Drosophila PTPs	0.1536

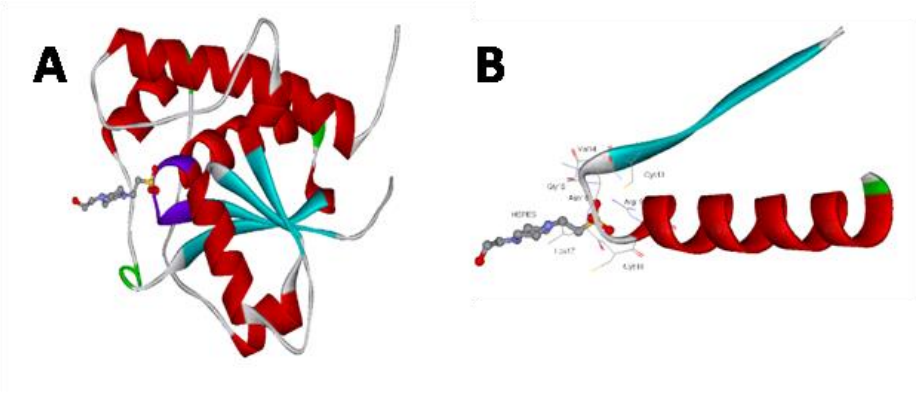


Figure 5. (A). Schematic diagram of Bovin heart PTPs (1DG9)complexed with HEPES (B). Active site residues complexed with HEPES in 1DG9 (Template)

184

185 The fold energetic Dr-PTPs was calculated by program Prosa using template Bh-
186 PTPs crystal structure. The comparison of energy was explored as shown by
187 energy graph (Fig. 6).

188

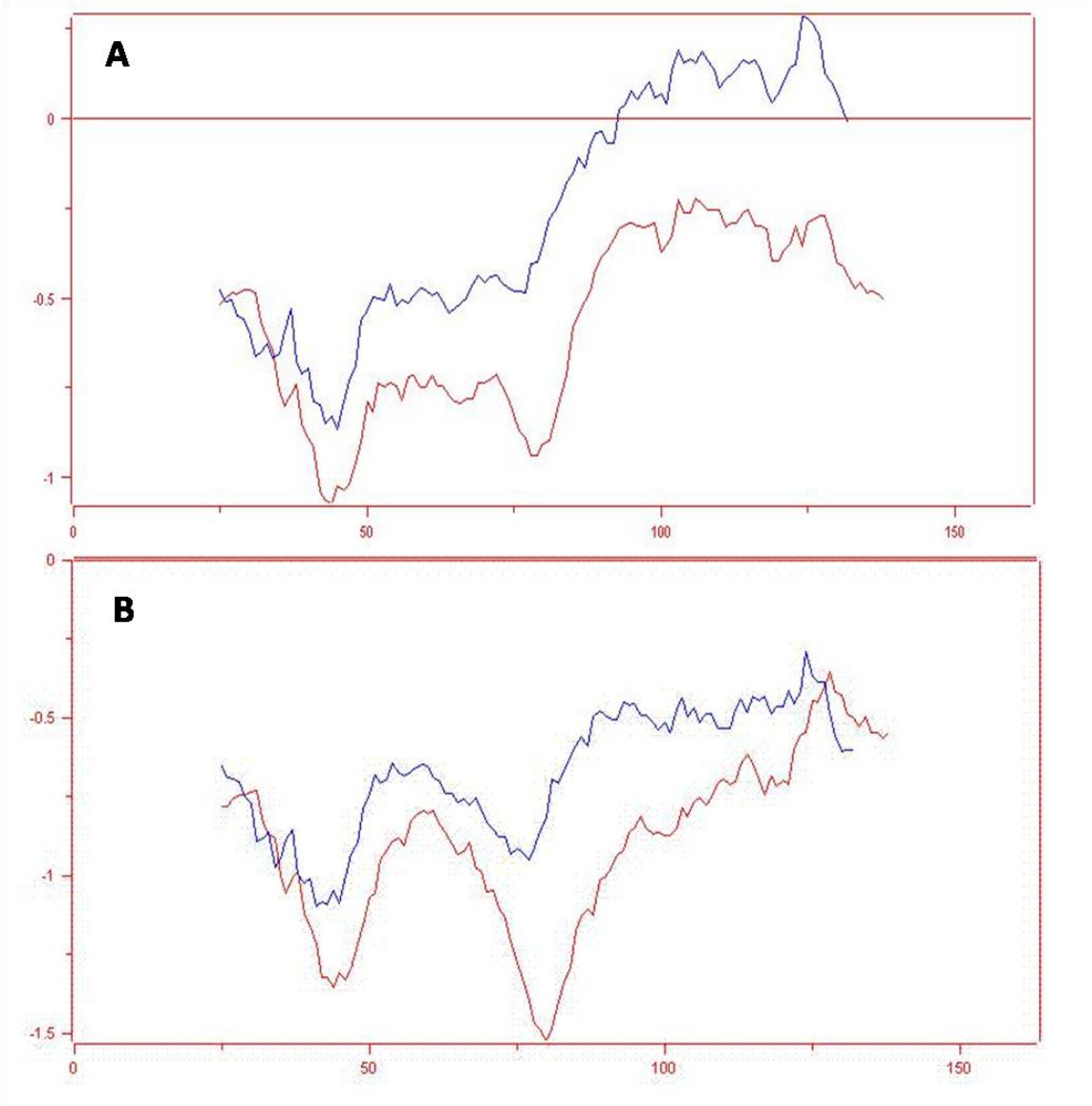


Figure 6. Comparison of PROSA II combined surface and pairing energy plots examined as a function of residue b/w template (blue) and the model (blue) for Bh-PTPs (A), 1DG9-PTPs, 1PNT-PTPs.

2.5. Active site and protein inhibitor interactions

The structural similarity of Dr-PTPs and Bh-PTPs and conformational resemblance of active site demonstrates the identity of catalytic mechanism. The catalytic reaction is triggered by the nucleophilic attack by Cys 3. The phosphate group at tyrosine group of the ligand or substrate bound at active site in orientation such that all the three oxygen atoms form hydrogen bonding with p-loop residues, making feasible the nucleophilic attack at phospho group by Cys3, Cys14, Cys16 and Cys18 where the Asp 129 works as proton donor resulting in

the formation of phosphoenzyme intermediate (17, 19). The hydrolytic cleavage of intermediate takesplace resulting in the formation of inorganic phosphate.

The complexation of HEPES at active site of enzyme is stabilized by hydrophilic interactions like hydrogen bonding and hydrophobic electrostatic interactions and thus acts as a potent inhibitor. In Bh-PTPs (PDB: IDG9), the inhibitor HEPES is stabilized by forming seven hydrogen bonds with active site residues of enzyme. In similar fashion, nine hydrogen bonds were observed in Dr-PTPase. In case of Bh-PTPs, residues like Leu 13, Gly 14, Ile 16, Cys 17 and Arg 18 are involved in hydrogen bonding with inhibitor HEPES. Only three oxygens of HEPES involved in hydrogen bonding with active site residues of PTPs (Table 2). Residues like Val 14, Gly 15, Leu 17, Cys 18, Arg 19 participate in hydrogen bonding of Dr-PTPs (Target). Two nitrogen atoms and three oxygen atoms of HEPES (inhibitor) are involved in hydrogen bonding with active site residues of PTPase (Fig 7b).

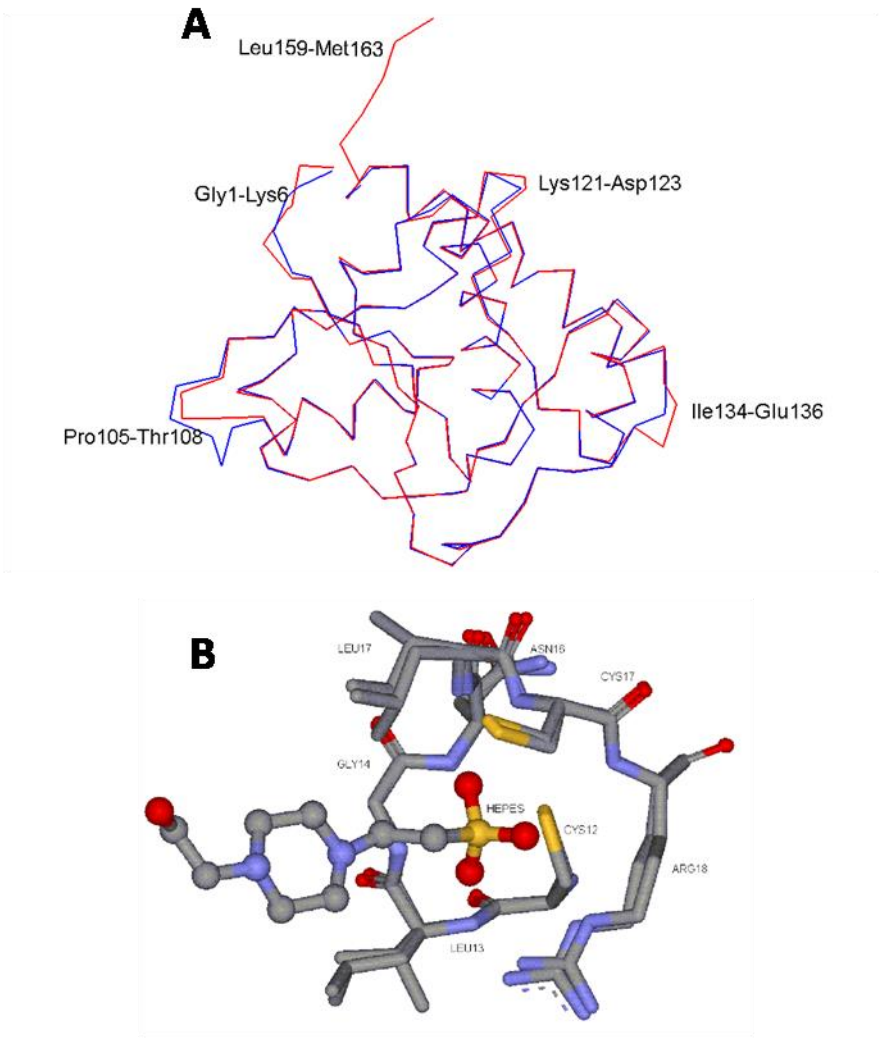


Figure 7. The active site is 100 % super imposed in template (Bh-PTPs) and Dr-PTPs. Active site is represented in sticks whereas inhibitor (HEPES) represents in ball and sticks style.

The stabilization of HEPES for complexation with Bh-PTPs is favored by hydrophobic interaction by residues Ile 16 and Tyr 131, while in Dr-PTPs, Leu 17, His 51 and Tyr 131 are involved in hydrophobic interactions.

Table 2. Ligand oxygen (HEPES)-Protein Nitrogen Bond length in PTPs complex

Bovine Heart PTPs Complex with HEPES			Drosophila PTPs Complex with HEPES		
1DG9	Amino acids	Distance (A^0)	Model PTPs	Amino acids	Distance (A^0)
EPE201:O1S	LEU13: N	3.17	EPE164:O1S	VA14: N	3.18
EPE201:O1S	GLY14: N	2.98	EPE164:O1S	GLY15: N	3.02
EPE201:O2S	ILE16: N	3.04	EPE164: C2	LEU17: N	3.06
EPE201:O2S	CYS17: N	2.88	EPE164:O2S	CYS18: N	2.88
EPE201:O1S	ARG 18: NH2	2.93	EPE164:O2S	ARG19: N	3.26
EPE201:O3S	ARG18: N	3.2	EPE164:O1S	ARG19: NH2	3.1
EPE201:O3S	ARG18: NE	2.94	EPE164:O3S	ARG19: NE	2.9
			EPE164: N1	EPE164: N4	2.99
			EPE164: N4	EPE164: N1	2.99

Table 3. Ligand oxygen (PO_4)-Protein Nitrogen Bond length in PTPs complex

Bovine Heart PTPs Complex with PO_4			Drosophila PTPs Complex with PO_4		
1PNT	Aminoacids	Distance((A)	PPD10	Aminoacids	Distance(A)
PO4158:O3	LEU13:N	3.24	PO4164:O3	VAL14:N	3.24
PO4158:O3	GLY14:N	2.87	PO4164:O3	GLY15:N	2.91
PO4158:O2	ILE16:N	2.87	PO4164:O2	ASN16:N	3.28
PO4158:O2	CYS17:N	2.73	PO4164:O2	LEU17:N	2.9
PO4158:O3	ARG18:NH2	2.92	PO4164:O2	CYS18:N	2.74
PO4158:O4	ARG18:N	3.29	PO4164:O3	ARG19:NH2	3.06
PO4158:O4	ARG18:NE	2.78	PO4164:O4	ARG19:NE	2.94
			PO4164:O1	TYR131:OH	3.17
			PO4164:O3	PO4164:O4	2.51
			PO4164:O4	PO4164:O3	2.51

3. Materials and Methods

3.1. Sequence analysis:

Sequence analysis of Dr–PTPs was obtained from SWISSPROT data base [26]. The sequence homology from the protein data bank was obtained from BLAST [27,28], where modeller was used for target template alignment [29] for Dr–PTPs. The program Cluster was used for analysis of multiple sequences and adjustment of parameters made where necessary and finally, phylogenetic lineage was established with program phylip [30,31].

3.2. Model building and refinement:

The three-dimensional model of Dr-PTPs was constructed using modeller ($9V_2$) using Bh-PTPs crystal structure (PDB: 1DG9) as template model. The program was allowed to satisfy all dihedral angle, bond and spatial restraints and distances automatically as per default parameters. The input files consist of Dr-PTPs and Bh-PTPs aligned sequence. Several runs of calculations were performed to get more reliable and plausible model. The homology modeling was performed using standard parameters of calculations and known 3D structure models from protein data bank. The secondary structure elements of model were visualized using pymol and molmol [16,31] and other structural statistics was performed using psvs site (https://montelionelab.chem.rpi.edu/PSVS). The interaction of several ligands was analyzed using program Ligand Explorer (http://users.sdsc.edu/~q2hang/ligand) [32].

3.3. Inhibitor modeling

The identification of hotspot residues, important for the target recognition and interaction was performed. The Dr-PTPs complexed with [N-(2-hydroxy ethyl) piperazine-N-2-ethanesulfonic acid sodium salt] (HEPES) was constructed using program modeller where the crystal structure 1DG9 was used as template. All these ligands known as potential inhibitors were models for the active sites for Dr-PTPs. The geometrical analysis, stereochemical analysis and all energies of bonds and dihedral restraints were analyzed. The homology models was subjected to program PROCHECK and ProSA for reliability of the model, secondary structure elements, backbone and energetic architecture and fold [33–36].

4. Conclusion

The sequence, secondary and tertiary structural similarities were studied in proteins Dr-PTPs and Bh-PTPs. The sequences of both proteins were found homologous for overall motifs and more especially for the active site motif represented by conserved residues C-$(X)_5$-R. The comparative analysis of

sequences is evident on the fact that this strong signature (CXXXXXR) at active site is the characteristic of low Molecular weight phosphotyrosine protein phosphatases. It was found that residues in 10-27, 81-88, and 127-130 regions were highly conserved with low Molecular weight PTPs. The structure obtained for this novel sequence were found of reliable and valid based on analysis performed by various protein structure validation tools as shown by the ramachandran plot, PROCHECK, energy of the fold and comparative analysis with other template crystal structures. The complexation profile of Dr-PTPs was also established based on the potent inhibitor HEPES. The overall stabilization factors important for the inhibitor complexation were studied and compared with known literature. We found that strong conformational similarity of Dr-PTPs with other homologous PTPs may shares same types of inhibitors exclusively considered to inhibit low Molecular weight phosphotyrosine protein PTPS. All these structural details obtained from the model are important for scheming further specific inhibitors. It can be used as an additional probe to decipher the discrete biological role of the low Molecular weight phosphotyrosine PTPS family and to explore the potential use of these macromolecular species as therapeutic targets.

Data and Software Availability

The data acquired in this manuscript is available for method validation and reproducibility of the results. We used softwares Clustalx [30,37] , Modeller[29], Procheck[33] for make sequence alignments and modeling studies. The data obtained is available in supporting information or otherwise mentioned in the articles.

Acknowledgment: The project was financed by Lucian Blaga University of Sibiu and Hasso Plattner Foundation research grants LBUS-IRG-2021-07.

References

(1) Kumar, A.; Rana, D.; Rana, R.; Bhatia, R. Protein Tyrosine Phosphatase (PTP1B): A Promising Drug Target Against Life-Threatening Ailments. *Current Molecular Pharmacology* **2020**, *13* (1), 17–30. https://doi.org/10.2174/1874467212666190724150723.

(2) Tonks, N. K. Protein Tyrosine Phosphatases: From Genes, to Function, to Disease. *Nature Reviews Molecular Cell Biology*. Nature Publishing Group November 2006, pp 833–846. https://doi.org/10.1038/nrm2039.

(3) Sa, N.; Rawat, R.; Thornburg, C.; Walker, K. D.; Roje, S. Identification and Characterization of the Missing Phosphatase on the Riboflavin Biosynthesis Pathway in *Arabidopsis Thaliana*. *The Plant Journal* **2016**, *88* (5),

705–716. https://doi.org/10.1111/tpj.13291.

(4) Niu, R. J.; Zheng, Q. C.; Zhang, H. X. Molecular Basis of the Recognition of FMN by a HAD Phosphatase TON_0338. *Journal of Molecular Graphics and Modelling* **2016**, *69*, 17–25. https://doi.org/10.1016/j.jmgm.2016.08.006.

(5) Bull, H.; Murray, P. G.; Thomas, D.; Fraser, A. M.; Nelson, P. N. Acid Phosphatases. *Journal of Clinical Pathology - Molecular Pathology*. BMJ Publishing Group 2002, pp 65–72. https://doi.org/10.1136/mp.55.2.65.

(6) Logan, T. M.; Zhou, M. M.; Nettesheim, D. G.; Meadows, R. P.; Fesik, S. W.; Van Etten, R. L. Solution Structure of a Low Molecular Weight Protein Tyrosine Phosphatase. *Biochemistry* **1994**, *33* (37), 11087–11096. https://doi.org/10.1021/bi00203a005.

(7) Zhang, M.; Van Etten, R. L.; Stauffacher, C. V. Crystal Structure of Bovine Heart Phosphotyrosyl Phosphatase at 2.2-Å Resolution. *Biochemistry* **1994**, *33* (37), 11097–11105. https://doi.org/10.1021/bi00203a006.

(8) Zhang, Z. Y. Protein-Tyrosine Phosphatases: Biological Function, Structural Characteristics, and Mechanism of Catalysis. *Critical Reviews in Biochemistry and Molecular Biology*. Crit Rev Biochem Mol Biol 1998, pp 1–52. https://doi.org/10.1080/10409239891204161.

(9) Granjeiro, J. M.; Miranda, M. A.; Maia, M. D. G. S. T.; Ferreira, C. V.; Taga, E. M.; Aoyama, H.; Volpe, P. L. O. Effect of Homologous Series of N-Alkyl Sulfates and N-Alkyl Trimethylammonium Bromides on Low Molecular Mass Protein Tyrosine Phosphatase Activity. *Molecular and Cellular Biochemistry* **2004**, *265* (1–2), 133–140. https://doi.org/10.1023/B:MCBI.0000044390.18530.39.

(10) Evans, B.; Tishmack, P. A.; Pokalsky, C.; Zhang, M.; Van Etten, R. L. Site-Directed Mutagenesis, Kinetic, and Spectroscopic Studies of the P- Loop Residues in a Low Molecular Weight Protein Tyrosine Phosphatase. *Biochemistry* **1996**, *35* (42), 13609–13617. https://doi.org/10.1021/bi9605651.

(11) Tailor, P.; Gilman, J.; Williams, S.; Couture, C.; Mustelin, T. Regulation of the Low Molecular Weight Phosphotyrosine Phosphatase by Phosphorylation at Tyrosines 131 and 132. *Journal of Biological Chemistry* **1997**, *272* (9), 5371–5374. https://doi.org/10.1074/jbc.272.9.5371.

(12) Yuvaniyama, J.; Denu, J. M.; Dixon, J. E.; Saper, M. A. Crystal Structure of the Dual Specificity Protein Phosphatase VHR. *Science* **1996**, *272* (5266), 1328–1331. https://doi.org/10.1126/science.272.5266.1328.

(13) Nam, H.-J.; Poy, F.; Krueger, N. X.; Saito, H.; Frederick, C. A. Crystal Structure of the Tandem Phosphatase Domains of RPTP LAR. *Cell* **1999**, *97* (4), 449–457. https://doi.org/10.1016/S0092-8674(00)80755-2.

(14) Kennelly, P. J. Life among the Primitives: Protein O-Phosphatases in Prokaryotes. *Frontiers in Bioscience* **1999**, *4* (1–3), d372.

https://doi.org/10.2741/kennelly.

(15) Hoffmann, K. M. V.; Tonks, N. K.; Barford, D. The Crystal Structure of Domain 1 of Receptor Protein-Tyrosine Phosphatase μ. *Journal of Biological Chemistry* **1997**, *272* (44), 27505–27508. https://doi.org/10.1074/jbc.272.44.27505.

(16) Glover, N. R.; Tracey, A. S. The Phosphatase Domains of LAR, CD45, and PTP1B: Structural Correlations with Peptide-Based Inhibitors 1 . *Biochemistry and Cell Biology* **2000**, *78* (1), 39–50. https://doi.org/10.1139/o99-069.

(17) Zhang, M.; Stauffacher, C. V.; Lin, D.; Van Etten, R. L. Crystal Structure of a Human Low Molecular Weight Phosphotyrosyl Phosphatase. *Journal of Biological Chemistry* **1998**, *273* (34), 21714–21720. https://doi.org/10.1074/jbc.273.34.21714.

(18) Wang, S.; Tabernero, L.; Zhang, M.; Harms, E.; Van Etten, R. L.; Stauffacher, C. V. Crystal Structures of a Low-Molecular Weight Protein Tyrosine Phosphatase from Saccharomyces Cerevisiae and Its Complex with the Substrate P-Nitrophenyl Phosphate. *Biochemistry* **2000**, *39* (8), 1903–1914. https://doi.org/10.1021/bi991348d.

(19) Zhang, M.; Zhou, M.; Van Etten, R. L.; Stauffacher, C. V. Crystal Structure of Bovine Low Molecular Weight Phosphotyrosyl Phosphatase Complexed with the Transition State Analog Vanadate. *Biochemistry* **1997**, *36* (1), 15–23. https://doi.org/10.1021/bi961804n.

(20) Jia, Z.; Barford, D.; Flint, A. J.; Tonks, N. K. Structural Basis for Phosphotyrosine Peptide Recognition by Protein Tyrosine Phosphatase 1B. *Science* **1995**, *268* (5218), 1754–1758. https://doi.org/10.1126/science.7540771.

(21) Wo, Y. Y. P.; McCormack, A. L.; Shabanowitz, J.; Hunt, D. F.; Davis, J. P.; Mitchell, G. L.; Van Etten, R. L. Sequencing, Cloning, and Expression of Human Red Cell-Type Acid Phosphatase, a Cytoplasmic Phosphotyrosyl Protein Phosphatase. *Journal of Biological Chemistry* **1992**, *267* (15), 10856–10865. https://doi.org/10.1016/s0021-9258(19)50097-7.

(22) Zhang, Z.; Harms, E.; Van Etten, R. L. Asp129 of Low Molecular Weight Protein Tyrosine Phosphatase Is Involved in Leaving Group Protonation. *Journal of Biological Chemistry* **1994**, *269* (42), 25947–25950. https://doi.org/10.1016/s0021-9258(18)47139-6.

(23) Lau, K. H. W.; Farley, J. R.; Baylink, D. J. Phosphotyrosyl Protein Phosphatases. *Biochemical Journal*. 1989, pp 23–36. https://doi.org/10.1042/bj2570023.

(24) Wu, L.; Zhang, Z. Y. Probing the Function of Asp128 in the Low Molecular Weight Protein-Tyrosine Phosphatase-Catalyzed Reaction. A Pre-Steady-State and Steady-State Kinetic Investigation. *Biochemistry* **1996**, *35* (17),

5426–5434. https://doi.org/10.1021/bi952885a.

(25) Zhang, Z. Y.; Van Etten, R. L. Leaving Group Dependence and Proton Inventory Studies of the Phosphorylation of a Cytoplasmic Phosphotyrosyl Protein Phosphatase from Bovine Heart. *Biochemistry* **1991**, *30* (37), 8954–8959. https://doi.org/10.1021/bi00101a006.

(26) Bairoch, A.; Apweiler, R. The SWISS-PROT Protein Sequence Database and Its Supplement TrEMBL in 2000. *Nucleic Acids Research*. Oxford University Press January 1, 2000, pp 45–48. https://doi.org/10.1093/nar/28.1.45.

(27) Boratyn, G. M.; Camacho, C.; Cooper, P. S.; Coulouris, G.; Fong, A.; Ma, N.; Madden, T. L.; Matten, W. T.; McGinnis, S. D.; Merezhuk, Y.; Raytselis, Y.; Sayers, E. W.; Tao, T.; Ye, J.; Zaretskaya, I. BLAST: A More Efficient Report with Usability Improvements. *Nucleic acids research* **2013**, *41* (Web Server issue). https://doi.org/10.1093/nar/gkt282.

(28) Morgulis, A.; Coulouris, G.; Raytselis, Y.; Madden, T. L.; Agarwala, R.; Schäffer, A. A. Database Indexing for Production MegaBLAST Searches. In *Bioinformatics*; Bioinformatics, 2008; Vol. 24, pp 1757–1764. https://doi.org/10.1093/bioinformatics/btn322.

(29) Eswar, N.; Webb, B.; Marti-Renom, M. A.; Madhusudhan, M. S.; Eramian, D.; Shen, M.; Pieper, U.; Sali, A. Comparative Protein Structure Modeling Using Modeller. *Current Protocols in Bioinformatics* **2006**, *15* (1), Unit. https://doi.org/10.1002/0471250953.bi0506s15.

(30) Wang, Y.; Wu, H.; Cai, Y. A Benchmark Study of Sequence Alignment Methods for Protein Clustering. *BMC Bioinformatics* **2018**, *19* (S19), 529. https://doi.org/10.1186/s12859-018-2524-4.

(31) Shimada, M. K.; Nishida, T. A Modification of the PHYLIP Program: A Solution for the Redundant Cluster Problem, and an Implementation of an Automatic Bootstrapping on Trees Inferred from Original Data. *Molecular Phylogenetics and Evolution* **2017**, *109*, 409–414. https://doi.org/10.1016/j.ympev.2017.02.012.

(32) Moreland, J. L.; Gramada, A.; Buzko, O. V.; Zhang, Q.; Bourne, P. E. The Molecular Biology Toolkit (MBT): A Modular Platform Fro Developing Molecular Visualization Applications. *BMC Bioinformatics* **2005**, *6* (1), 21. https://doi.org/10.1186/1471-2105-6-21.

(33) Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M. PROCHECK: A Program to Check the Stereochemical Quality of Protein Structures. *Journal of Applied Crystallography* **1993**, *26* (2), 283–291. https://doi.org/10.1107/s0021889892009944.

(34) Laskowski, R. A.; Rullmann, J. A. C.; MacArthur, M. W.; Kaptein, R.; Thornton, J. M. AQUA and PROCHECK-NMR: Programs for Checking the Quality of Protein Structures Solved by NMR. *Journal of Biomolecular NMR*

1996, *8* (4), 477–486. https://doi.org/10.1007/BF00228148.

(35) Furey, W.; Cowtan, K. D.; Zhang, K. Y. J.; Main, P.; Brunger, A. T.; Adams, P. D.; DeLano, W. L.; Gros, P.; Grosse-Kunstleve, R. W.; Jiang, J.-S.; Pannu, N. S.; Read, R. J.; Rice, L. M.; Simonson, T.; Tronrud, D. E.; Ten Eyck, L. F.; Lamzin, V. S.; Perrakis, A.; Wilson, K. S.; Laskowski, R. A.; MacArthur, M. W.; Thornton, J. M.; Kraulis, P. J.; Richardson, D. C.; Richardson, J. S.; Kabsch, W.; Sheldrick, G. M. Programs and Program Systems in Wide Use. In *International Tables for Crystallography*; International Union of Crystallography, 2006; pp 695–743. https://doi.org/10.1107/97809553602060000724.

(36) Wiederstein, M.; Sippl, M. J. ProSA-Web: Interactive Web Service for the Recognition of Errors in Three-Dimensional Structures of Proteins. *Nucleic Acids Research* **2007**, *35* (SUPPL.2), W407. https://doi.org/10.1093/nar/gkm290.

(37) Chenna, R.; Sugawara, H.; Koike, T.; Lopez, R.; Gibson, T. J.; Higgins, D. G.; Thompson, J. D. Multiple Sequence Alignment with the Clustal Series of Programs. *Nucleic Acids Research* **2003**, *31* (13), 3497–3500. https://doi.org/10.1093/nar/gkg500.

1

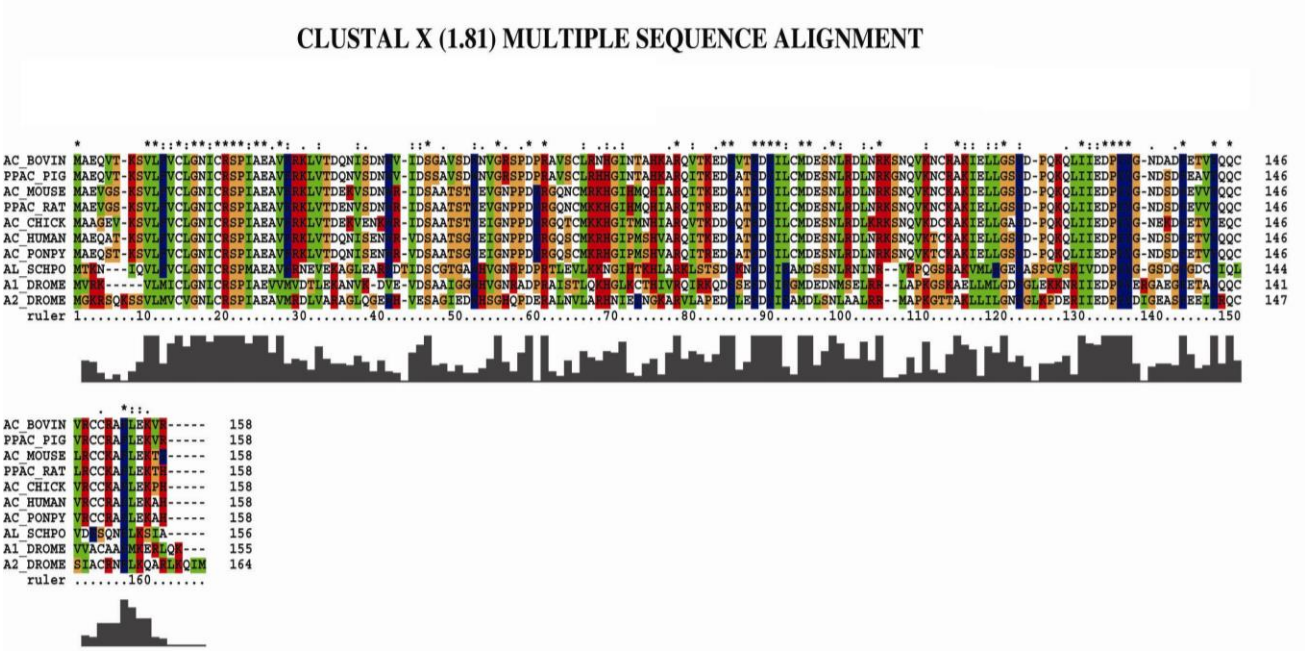


Figure 1. Multiple sequence alignment of 10 PTPs sequences. Sequences are named as Swiss-Prot entery first letter represents gene, the second part represents the biological source of gene. The symbol ''*'' represents strongly conserved '' .'' represents weakly conserved'' :''represents identical residues

2

3

```
                    10         20         30         40         50         60
1DG9        AEQV-TKSVLFVCLGNICRSPIAEAVFRKLVTDQNISDNWVIDSGAVSDWNVGRSPDPRAVSCLRNHG
PTPs        GKRSQKSSVLMVCVGNLCRSPIAEAVMRDLVARAGLQGEWHVESAGIEDWHSGHQPDERALNVLARHN
                   *** ** ** ********* * **         *   *     **  *  ** **   *  *

            70         80         90         100        110        120        130
1DG9        INTAHKARQVTKEDFVTFDYILCMDESNLRDLNRKSNQVKNCRAKIELLGSYDPQ-KQLIIEDPYYG-
PTPs        IEYNGKARVLAPEDFLEFDYIFAMDLSNLAALRRMAP--KGTTAKLLILGNFGLKPDERIIEDPYYDI
            *    ***     ***  ****  ** ***  * *      *   **    **          *******

             140        150        160
1DG9        NDADFETVYQQCVRCCRAFLEKVR-----
PTPs        GEASFEEIYRQCSIACRNFLKQARLKQIM
              * **  * **   ** **   *
```

Figure 2. Pairwise sequence alignment used for building model of Drosophila phosphatase. Target sequence represented by PTPs. based on the structure of Bovin heart phosphatase template structure represented by1DG9.

4

5

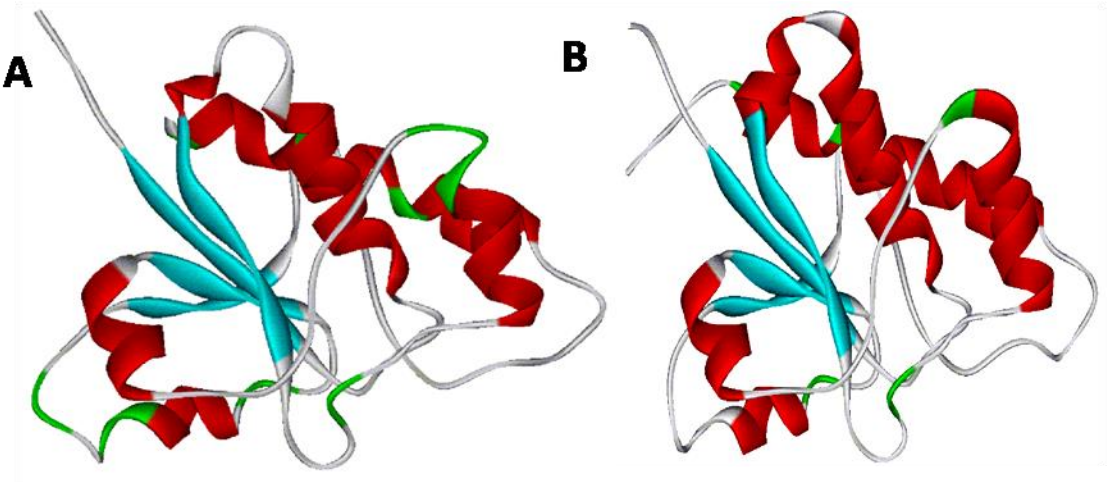


Figure 3. A complete Ribbon diagram of low molecular weight phosphotyrosine protein phosphatase. Showing α/β Proteins, α helix (Red), β Sheets (blue).(A) 1BVH(template), (B) Drosophila(target).

6

7

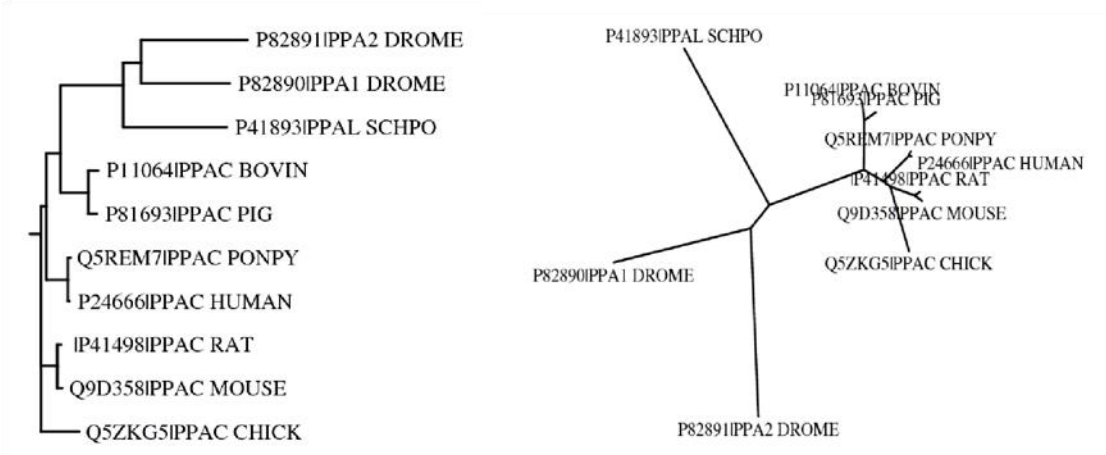


Figure 4. Dendrogram of PPA2 *Drosophila Melanogaster* and related proteins made by PHYLIP & depicts relationship among various forms of PPAC,PPAI, PPA2 genes. The first part of the code represents gene where as the second part of the code represents the source of protein.

8

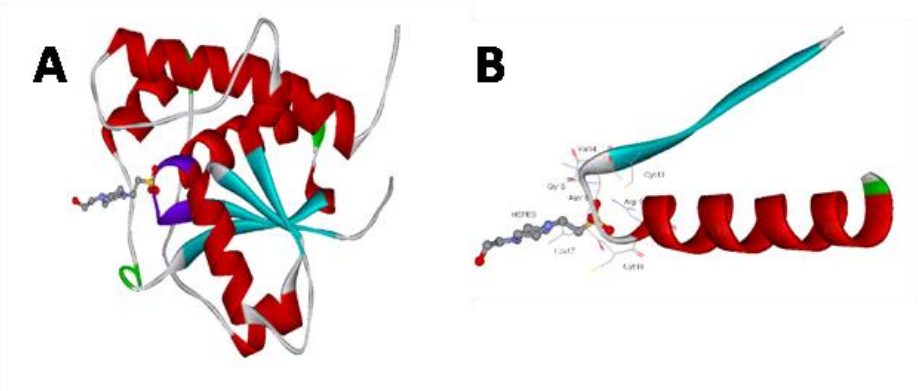
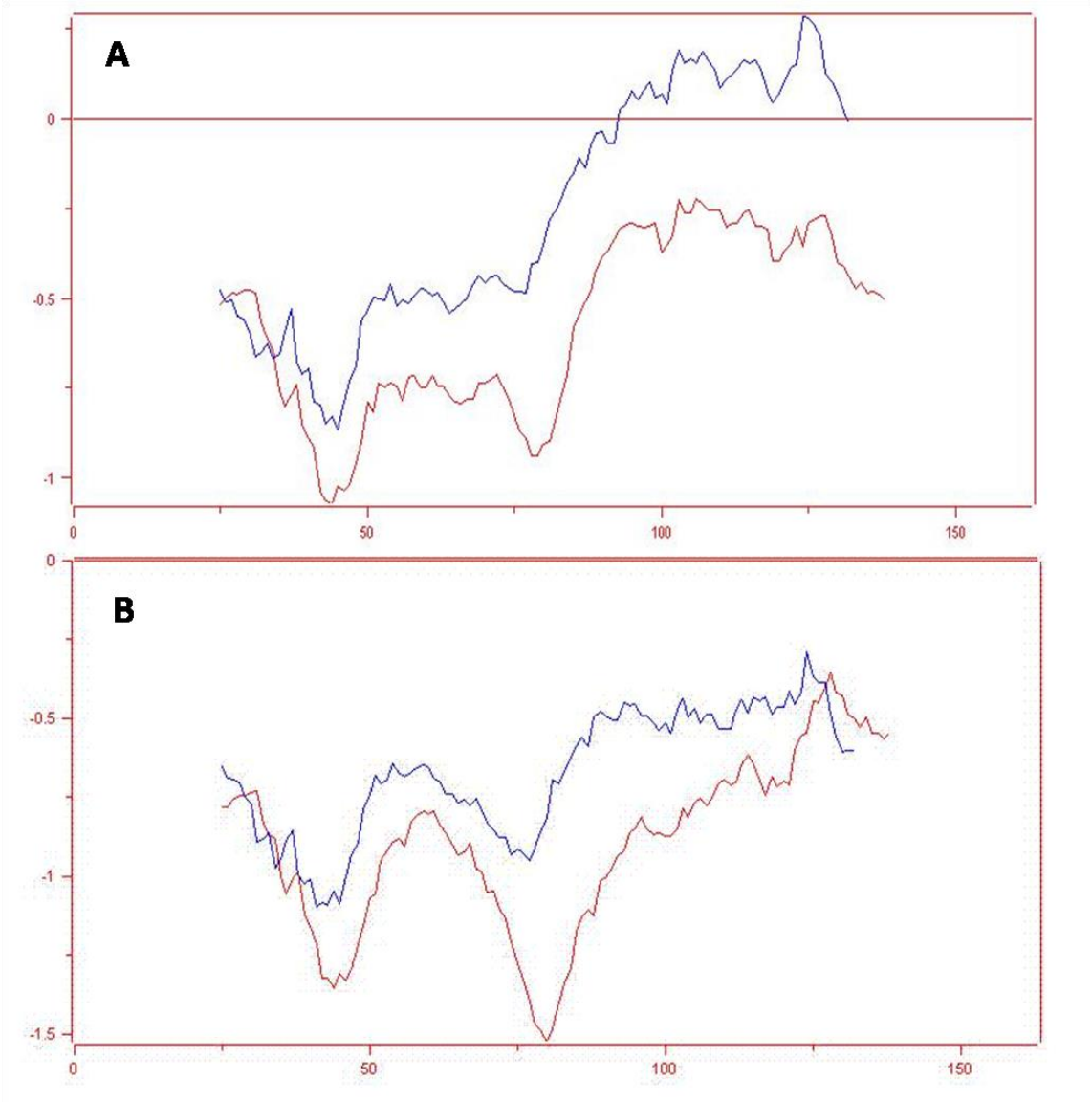


Figure 5. (A). Schematic diagram of Bovin heart PTPs (1DG9) complexed with HEPES (B). Active site residues complexed with HEPES in 1DG9 (Template)

9



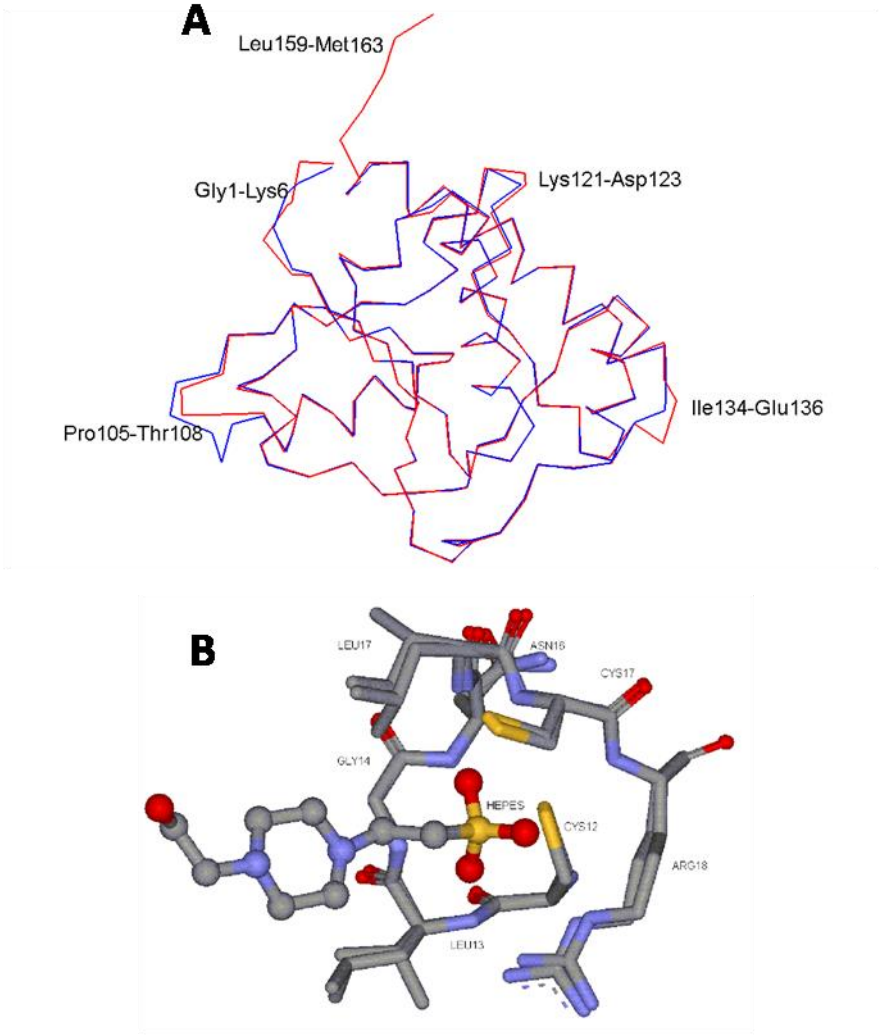


Figure 7. The active site is 100 % super imposed in template (Bh-PTPs) and Dr-PTPs. Active site is represented in sticks whereas inhibitor (HEPES) represents in ball and sticks style.