STRUCTURAL BIOLOGY

Volume 75 (2019)

Supporting information for article:

Crystal structures and snapshots along the reaction pathway of human phosphoserine phosphatase

Marie Haufroid, Manon Mirgaux, Laurence Leherte and Johan Wouters

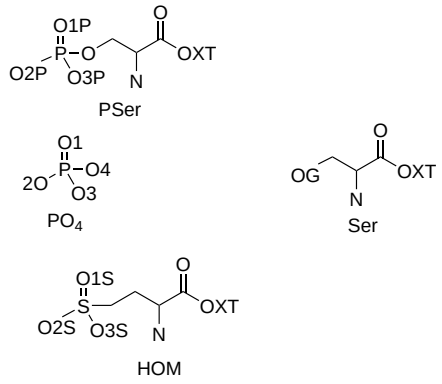

Fig. S1. Ligand structure with PDB code of interacting atom.

S*1.1. Distances tables*

Table S1. *Main interactions and bond lengths (° A) between ligands and hPSP in structure hPSP-PSer*

Ligand	hPSP	d_{X-A} (Å)
H-bond		
P-Ser^{1}O	Thr^{182}OG1 (A)	3.0
P-Ser^{1}O	W182	1.6
P-Ser^{1}O	W96	3.0
P-Ser^{1}N	W184	2.2
P-Ser^{1}OG	W180	2.4
P-Ser^{1}O3P	W180	3.2
P-Ser^{1}O3P	W181	2.0
P-Ser^{1}O3P	W7	2.4
P-Ser^{1}O3P	Gly^{110}N (A)	2.8
P-Ser^{1}O1P	Gly^{110}N (A)	2.9
P-Ser^{1}O1P	W183	2.6
P-Ser^{2}OG	W185	1.9
P-Ser^{2}N	W185	2.4
Salting Bridge		
P-Ser^{1}O1P	Lys^{158}NZ (A)	3.5
Ser^{1}OXT	Arg^{202}NH2 (B)	2.8
P-Ser^{2}O1P	Lys^{168}(B)	3.3
Close contact		
Ser^{1}N	Arg^{202}NH1 (B)	3.4
Ser^{1}OXT	Glu^{29}OE1 (B)	2.6
Ser^{1}O	Glu^{29}OE1 (B)	2.9
π-cation		
P-Ser^{2}N	Phe$^{131}_{cent}$ (B)	3.9

Table S2. *Main interactions and bond lengths between ligands and hPSP in structure hPSP-Pi-Ser*

Ligand	hPSP	d_{X-A} (Å)
H-bond		
$Ser^{1}OG$	$Ser^{109}OG$ (A)	3.0
$Ser^{1}OG$	W347	2.7
$Ser^{1}OG$	$Asp^{22}N$ (A)	3.0
$Ser^{1}OG$	$Val^{21}N$ (A)	3.1
$Ser^{1}OG$	$Asp^{20}OD1$ (A)	2.5
$Ser^{1}O$	W333	2.7
$Ser^{1}O$	W349	3.1
$Ser^{1}O$	$Gly^{110}N$ (A)	3.1
$Ser^{1}OXT$	W166	3.0
$PO4^{A}O1$	W14	2.9
$PO4^{A}O1$	W166	3.4
$PO4^{A}O1$	W191	2.5
$PO4^{A}O1$	W349	2.3
$PO4^{A}O2$	$Thr^{182}OG1$ (A)	2.7
$PO4^{A}O2$	$Thr^{182}N$ (A)	2.9
$PO4^{A}O3$	$Arg^{202B}NE$ (A)	2.8
$PO4^{A}O4$	$Gly^{53}N$ (A)	2.8
$PO4^{A}O4$	W245	2.6
$PO4^{B}O1$	W1	3.1
$PO4^{B}O1$	W2	3.4
$PO4^{B}O1$	W267	3.0
$PO4^{B}O2$	W304	2.5
$PO4^{B}O2$	$Gly^{110}N$ (B)	2.9
$PO4^{B}O3$	W149	2.8
$PO4^{B}O4$	$Ser^{109}OG$ (B)	2.7
$PO4^{B}O4$	$Val^{21}N$ (B)	3.3
$PO4^{B}O4$	$Asp^{22}N$ (B)	2.8
Salt Bridge		
$Ser^{1}OXT$	$Lys^{158}NZ$ (A)	2.4
$Ser^{1}N$	$Asp^{22}OD2$ (A)	2.5
$PO4^{A}O3$	$Arg^{202B}NH1$ (A)	2.4
$PO4^{B}O2$	$Lys^{158}NZ$ (B)	2.7
Metal coordination		
$PO4^{B}O1$	Ca^{2+}	2.4

Table S3. *Main interactions and bond length between ligands and hPSP in structure hPSP-homocysteic acid*

Ligand	hPSP	d_{X-A} (Å)
H-bond		
$HOM^{1}O1S$	$Ser^{109}OG$	3.4
$HOM^{1}O1S$	W93	3.3
$HOM^{1}O1S$	$Asp^{22}N$	2.5
$HOM^{1}O1S$	$Val^{21}N$	3.2
$HOM^{1}O2S$	W66	2.4
$HOM^{1}O2S$	W93	2.8
$HOM^{1}O3S$	W28	3.4
$HOM^{1}O3S$	$Gly^{111}N$	2.2
$HOM^{1}O3S$	$Gly^{110}N$	2.6
$HOM^{1}O3S$	$Ser^{109}OG$	1.9
$HOM^{1}OXT$	W131	2.3
Close-contact		
$HOM^{1}O2S$	$Asp^{22}OD2$	2.3
$HOM^{1}O1S$	$Asp^{20}OD1$	2.6

S*1.2. Structures comparison with other PSP*

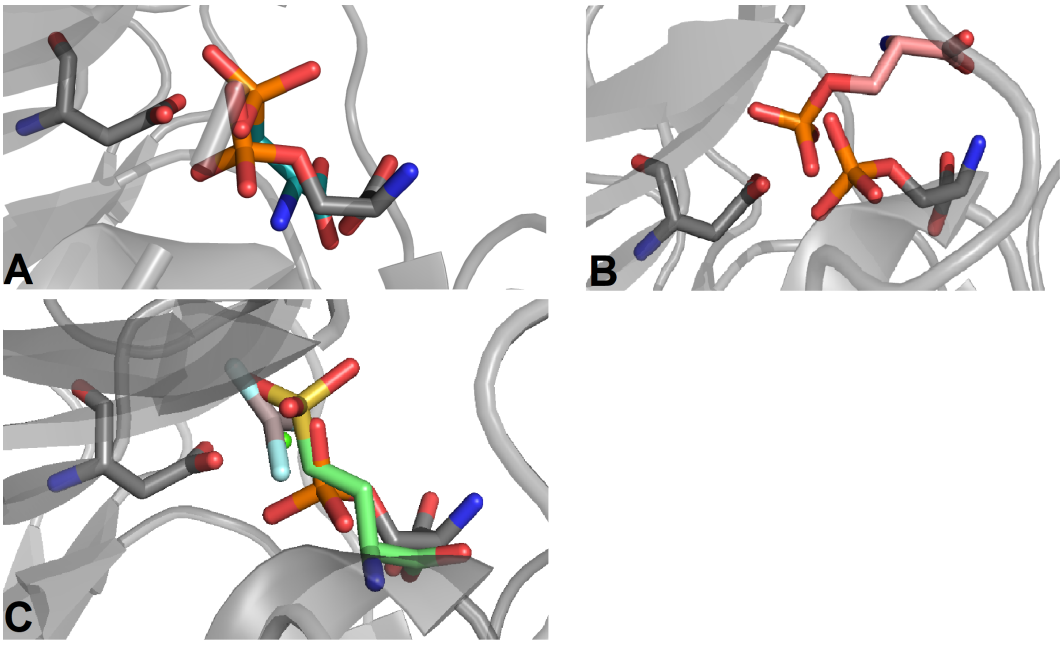

Fig. S2. **A** Superimposition of structure of hPSP with phosphoserine and the one from Kim *et al.* containing 2-amino-3-propionic acid. **B** Superimposition of structure of both PSP from human and *Methanocaldococcus* in complex with phosphoserine. **C** Superimposition of structure of hPSP with phosphoserine and homocysteic acid and the structure of MjPSP with AlF_3.

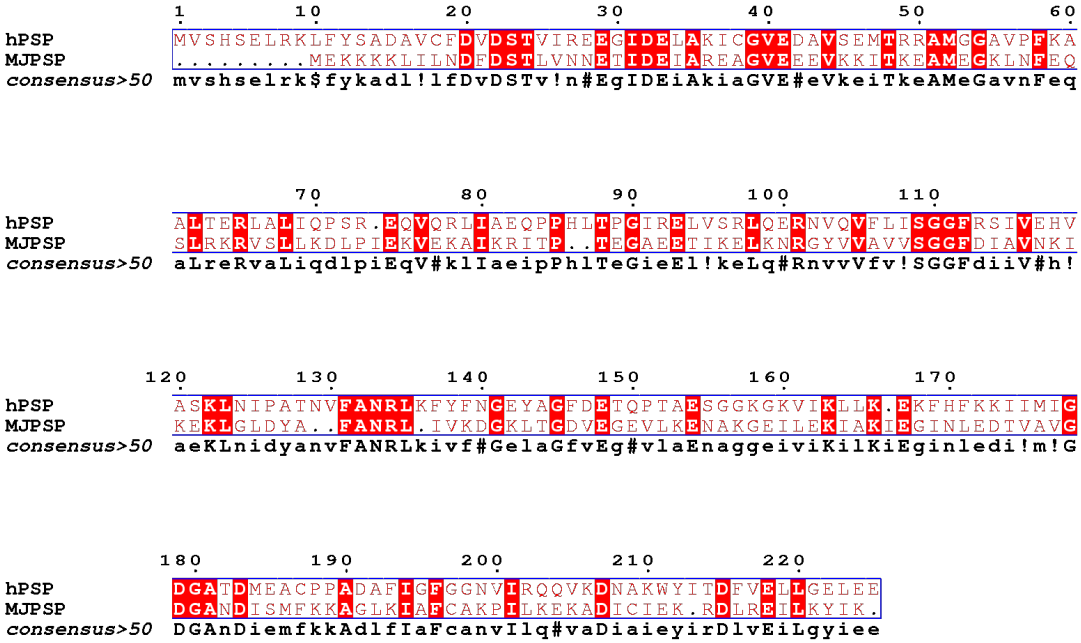

Fig. S3. Sequence alignement of human phosphoserine phosphatase and the one from *Methanocaldococcus jannaschii*.

S*1.3. Technical details regarding the ligand structure building*

Starting from the crystal structure coordinates of phosphoserine (PSer), hydrogen atoms were added with VegaZZ (Pedretti *et al.*, 2004). The resulting structure was then grossly optimized using the steepest descent algorithm of VegaZZ with a tolerance of 1.0 kcal.mol^{-1}.Å^{-1}, a high value selected to prevent any drastic modification of the heavy atoms coordinates. Gromacs coordinates and topology files were generated using Topolbuild (Ray BD, 2018). Regarding the protein, the missing three first amino acid residues, Met^1-Ile^2-Ser^3, were added using the following procedure. First, chain B, which contains the first three residues, was aligned onto chain A. The so-obtained coordinates of Met^1 to Ser^3 were added to the incomplete chain A. The atomic charges of PSer were determined using the Quantum Mechanics program Gaussian (Frisch *et al.*, 2009) at the RHF/6-31G(d) level. A molecular electrostatic potential (MEP)

grid of 250 x 250 x 250 elements and a grid interval of 0.1 Å was generated according to the Merz-Singh-Kollman scheme (Singh & Kollman, 1984; Besler *et al.*, 1990). A modified version of QFIT (Borodin & Smith, 2009) was run to determine the atomic charges (Table 4). The program code was adapted so as to fit electrostatic forces calculated from electrostatic potential grids, as described in Ref. (Leherte, 2016).

Table S4. *RHF 6-31G(d) atomic charges (e-) of PSER obtained using Gaussian (Frisch* et al.*, 2009)*

Atom	Net charge	Atom	Net charge
N	-0.8195	CB	0.1630
H1, H2, H3	0.4145	HB1, HB2	0.0352
CA	0.0919	OG	-0.5441
HA	0.0238	OXT O1P, O2P	-0.9771
C	0.9257	O3P	-0.9615
O, OXT	-0.8257	P	1.4524

From the reference MEP grids, fittings were achieved by considering points located at distances between 1.4 and 2.0 times the van der Waals (vdW) radius of the atoms. These two limiting distance values were selected after the so-called Merz-Singh-Kollman scheme (Singh & Kollman, 1984). Constraints, such as the total molecular charge (-2 e-) and the total dipole moment were applied. Additional constraints were considered so as to force the atomic charges of the atoms H1, H2, and H3, to be equal, as well as for atoms HB1 and HB2, O1P and O2P, and O and OXT.

S*1.4. Description of the Molecular Dynamics calculations*

MD trajectories of the solvated systems were run using the Gromacs4.5.5 program package (Hess *et al.*, 2008; Pronk *et al.*, 2009) with the Amber99sb-ildn FF (Lindorff-Larsen *et al.*, 2010) under particle mesh Ewald periodic boundary conditions (PBC) and a Coulomb cut-off distance of 1.0 nm. The ligand PSer was described using the Generalized Amber Force Field parameters. The Newton equations of motion were numerically integrated using a leap-frog integrator. The van der Waals (vdW) cut-off distance was set equal to 1.4 nm. Long-range dispersion corrections to energy and pressure were applied. The systems were optimized using a steepest descent algorithm

with an initial step size of 0.10 nm. To strongly reduce the calculation time, the hybrid TIP3P/SIRAH water FF was used (Darré *et al.*, 2012; Gonzalez *et al.*, 2013; Leonardo *et al.*, 2015). The initial protein systems (with and without ligand) were solvated so as protein atoms lie at least at a distance of 2.0 nm from the cubic box walls. A shell of 1.0 nm thickness of TIP3P water molecules was defined around the protein complex, and the remaining space of the solvation box was filled with SIRAH water beads (Machado, 2018) where each bead is composed of four interaction sites and represents about 11 water molecules (Table 5).

Table S5. *Description of the protein systems simulated using Gromacs at T = 300 K, and at P = 1 bar in hybrid TIP3P/SIRAH water*

	hPSP	hPSP-PSer
Total no. of atoms	25,175	25,194
No. of all-atom/CG water molecules	3,373/2,880	3,373/2,880
No. of ions	2 K^+, 3 Na^+	3 K^+, 4 Na^+
Final box size (nm)	10.36860	10.36482

K^+ and Na^+ ions were considered to neutralize the electric charge of the protein systems (Table 5). The whole systems were again optimized, using a steepest descent algorithm with an initial step size of 0.10 nm, to eliminate large forces and then heated to 50 K through a 10 ps canonical (NVT) MD, with a time step of 2 fs and LINCS constraints acting on bonds involving H atoms. The trajectory was followed by two successive 20 ps heating stages, at 150 K and at the final temperature of 300 K, under the same conditions. Next, each system was equilibrated during 50 ps in the NPT ensemble, at P = 1 bar, to relax the solvent molecules, and for a further 160 ns MD equilibration run. The 'V-Rescale' and 'Parrinello-Rahman' algorithms were selected to constrain T and P, respectively. A final production run of 200 ns (100 x 10^6 steps) was performed for the evaluation of the structural, energetics, and dynamical properties of each system. Trajectory data were saved every 20 ps.

References

Besler, B. H., Merz Jr, K. M. & Kollman, P. A. (1990). *Journal of Computational Chemistry*, **11**(4), 431–439.s

Borodin, O. & Smith, G. D., (2009). Force field fitting toolkit.
URL: *http://www.eng.utah.edu/ gdsmith/fff.html*

Darré, L., Tek, A., Baaden, M. & Pantano, S. (2012). *Journal of Chemical Theory and Computation*, **8**(10), 3880–3894.

Frisch, M., Trucks, G., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Scalmani, G., Barone, V., Mennucci, B., Petersson, G. *et al.* (2009). *Inc., Wallingford, CT*, **200**.

Gonzalez, H. C., Darré, L. & Pantano, S. (2013). *The Journal of Physical Chemistry B*, **117**(46), 14438–14448.

Hess, B., Kutzner, C., Van Der Spoel, D. & Lindahl, E. (2008). *Journal of Chemical Theory and Computation*, **4**(3), 435–447.

Leherte, L. (2016). *Molecular Simulation*, **42**(4), 289–304.

Leonardo, D., Machado, M. R., Brandner, A. F., Gonzalez, H. C., Ferreira, S. & Pantano, S. (2015). *Journal of Chemical Theory and Computation*, **11**(2), 723–739.

Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J. L., Dror, R. O. & Shaw, D. E. (2010). *Proteins: Structure, Function, and Bioinformatics*, **78**(8), 1950–1958.

Machado, M., (2018). Tutorial 2.SIRAH forcefield in GROMACS, hybrid solvation: Plugging SIRAH solvent to your atomistic system.
URL: *https://training.vi-seem.eu/images/trainingMaterial/LifeSciences/Tutorial_2_sirah4gmx.pdf*

Pedretti, A., Villa, L. & Vistoli, G. (2004). *Journal of Computer-Aided Molecular Design*, **18**(3), 167–173.

Pronk, S., Páll, S., Schulz, R., Larsson, P., Bjelkmar, P., Apostolov, R., Shirts, M. R., Smith, J. C., Kasson, P. M., Van Der Spoel, D. *et al.* (2009). *Bioinformatics*, **29**(7), 845–854.

Ray BD, (2018). Topolbuild 1.3.
URL: *http://www.gromacs.org/Downloads/User_contributions/Other_software*

Singh, U. C. & Kollman, P. A. (1984). *Journal of Computational Chemistry*, **5**(2), 129–145.