Supporting Information:

Divergence of Structure and Function in the Haloacid Dehalogenase Enzyme Superfamily:

***Bacteroides Thetaiotaomicron* BT2127 is an Inorganic Pyrophosphatase**[+]

Hua Huang[1], Yury Patskovsky [2], Rafael Toro[2], Jeremiah D. Farelli[3], Chetanya Pandya[4], Steve Almo[2*], Karen N. Allen[3]*, and Debra Dunaway-Mariano[1]*

[1]*Department of Chemistry and Chemical Biology, University of New Mexico, Albuquerque, New Mexico 87131;* [2] *Department of Biochemistry, Albert Einstein College of Medicine, Jack and Pearl Resnick Campus, 1300 Morris Park Avenue, Bronx, NY 10461;* [3]*Department of Chemistry, Boston University 590 Commonwealth Avenue, Boston, MA 02215-2521,* [4]*Bioinformatics Graduate Program, Boston University, 24 Cummington Street Boston, MA 02215-2521*

*To whom correspondence should be addressed regarding X-ray structure determination (S.A. or K.N.A.) or kinetic analyses (D.D.-M.)

Contents: 20 pages total, which contain Tables S1-S8 and Figures S1-S9.

Table S1. Crystallization conditions used for wild-type and mutant BT2127.

BT2127	PDB code	Ligand	Crystallization conditions
WT	3QU2	citrate, Cl^-, glycerol, Mg^{2+}	0.1 M sodium citrate, pH 5.6, 30% PEG4000, 0.2 M ammonium acetate, 5 mM, $MgCl_2$, 294 K
WT	3QXG	acetate, Ca^{2+}, Mg^{2+}, PO_4^{2-}, tartrate	0.1 M sodium acetate, 0.2 M calcium acetate, 30% PEG400, 5 mM $MgCl_2$, 294 K
WT	3QUQ	Cl^-, formate, Mg^{2+}, unknown	0.1 M Bis-Tris, pH 6.5, 25% PEG3350, 0.2 M $MgCl_2$, 294 K
WT	3QX7	Mg^{2+}, PO_4^{2-}	0.056 M sodium phosphate monobasic, pH 8.2, 1.34 M potassium phosphate dibasic, 5 mM $MgCl_2$, 294 K
D11N	3QU5	Cl^-	0.1 M Bis-Tris, pH 6.5, 25% PEG3350, 0.2 M $MgCl_2$, 294 K
D13A	3QU4	Acetate, Cl^-, Mg^{2+}	0.1 M MES, pH 6.0, 20% PEG8000, 0.2 M calcium acetate, 5 mM $MgCl_2$, 291 K
D13N	3QU9	Cl^-, glycerol, Mg^{2+}, tartrate	0.1 M potassium sodium tartrate, pH 8.0, 0.1 M imidazole, 0.2 M NaCl, 5 mM MgCl2, 294 K
D13N	3QU7	Acetate, Ca^{2+}, PO_4^{2-}	0.1 M sodium acetate, pH 4.5, 30% PEG400, 0.2 M calcium acetate, 5 mM $MgCl_2$, 294 K
D13N	3QUT	Cl^-, Mg^{2+}, malate	0.15 M Malic acid, pH 7, 20% PEG3350, 5 mM $MgCl_2$, 294 K
E47D	3R9K	SO_4^{2-}	0.1 M Tris-HCl, pH 8.5, 30% PEG4000, 0.2 M $LiSO_4$, 5 mM $MgCl_2$, 294 K
E47N	3QYP	Ca^{2+}, Cl^-, glycerol, PO_4^{2-}, unknown	0.1 M MES, pH 6.0, 20% PEG8000, 0.2 M calcium acetate, 5 mM $MgCl_2$, 294 K
E47N	3QUC	SO_4^{2-}	0.1 M Tris-HCl, pH 8.5, 30% PEG4000, 0.2 M $LiSO_4$, 5 mM $MgCl_2$, 294 K
E47A	3QUB	SO_4^{2-}	0.1 M Tris-HCl, pH 8.5, 30% PEG4000, 0.2 M $LiSO_4$, 5 mM $MgCl_2$, 294 K

Table S2. Data collection and refinement statistics for wild-type BT2127. Numbers in parenthesis are those for the high resolution shell.

PDB ID	3QUQ	3QU2	3QX7	3QXG
Cap orientation	Opened	Closed	Closed	Closed
Space group	P212121	C2	P3121	C2
Unit cell dimensions (Å)	-	-	-	-
-a	45.20	137.60	84.58	114.07
-b	65.34	71.76	84.58	70.59
-c	75.62	114.53	79.27	76.32
Cell angles (degrees)	-	-	-	-
-alpha	90.00	90.00	90.00	90.00
-beta	90.00	105.80	90.00	119.77
-gamma	90.00	90.00	120.00	90.00
Molecules per asymmetric unit	1	4	1	2
Solvent content	38.42	51.25	61.05	50.18
Matthew's Coefficient	2.00	2.52	3.16	2.47
Ligands	Cl^-, formate, Mg^{2+}, unknown	Citrate, Cl^-, glycerol, Mg^{2+}	Mg^{2+}, PO_4^{2-}	Acetate, Ca^{2+}, Mg^{2+}, $PO4^{2-}$, tartrate
X-ray source	NSLS X29A	NSLS X29A	NSLS X29A	NSLS X29A
Wavelength	1.0750	1.0809	0.9791	0.9791
Resolution	49.45-1.65	40.00-1.94	20.00-2.00	50.00-1.24
Reflections	28,023	79,445	22,599	147,352
Completeness (%)	99.2 (92.1)	99.3 (93.1)	95.00	99.0 (86.0)
I/sigma (I)	6.10 (3.00)	3.40 (1.20)	7.40 (1.00)	7.40 (0.90)
R_{sym}	0.092 (0.530)	0.124 (0.800)	0.078 (0.900)	0.056 (0.920)
R_{work} (R_{free})	16.4 (17.8)	23.4 (27.5)	19.1 (24.3)	15.2 (18.2)
R_{free} reflections (%)	683 (3.2%)	2349 (3.0%)	683 (3.2%)	4418 (3.0%)
Average B factor (overall)	23.69	20.31	58.11	22.23
-amino acid residues	22.54	19.78	58.15	20.26
-waters	33.12	25.21	56.95	35.98
-ligands	26.74, 13.70, 47.62, 13.19	43.90, 32.70, 32.49, 13.95	51.78, 58.89	19.22, 18.43, 24.61, 21.26, 21.78
Root-mean square deviation	-	-	-	-
-bond lengths	0.011	0.010	0.011	0.012
-bond angles	1.293	1.273	1.293	1.411
Ramachandran favored/allowed (%)	99.1 (0.9)	99.1 (0.9)	98.7 (1.3)	98.8 (1.2)
Number of solvent molecules	98	649	98	525

Table S3. Data collection and refinement statistics for BT2127 D11N. Numbers in parenthesis are those for the high resolution shell

PDB ID	3QU5
Mutation	D11N
Cap orientation	Closed
Space group	P21
Unit cell dimensions (Å)	-
-a	42.80
-b	77.21
-c	70.76
Cell angles (degrees)	-
-alpha	90.00
-beta	90.41
-gamma	90.00
Molecules per asymmetric unit	2
Solvent content	45.44
Matthew's Coefficient	2.25
Ligands	Cl^-
X-ray source	NSLS X29A
Wavelength	0.9786
Resolution	40.00-1.24
Reflections	129,987
Completeness (%)	96.3 (68.4)
I/sigma (I)	8.00 (1.40)
R_{sym}	0.043 (0.550)
R_{work} (R_{free})	13.0 (16.0)
R_{free} reflections (%)	3775 (3.0%)
Average B factor (overall)	22.35
-amino acid residues	19.62
-waters	35.36
-ligands	23.42
Root-mean square deviation	-
-bond lengths	0.011
-bond angles	1.407
Ramachandran favored/allowed (%)	99.4 (0.6)
Number of solvent molecules	764

Table S4. Data collection and refinement statistics for BT2127 D13. Numbers in parenthesis are those for the high resolution shell

PDB ID	3QU4	3QUT	3QU7	3QU9
Mutation	D13A	D13N	D13N	D13N
Cap orientation	Closed	Opened	Closed	Closed
Space group	P1	P212121	C2221	P3121
Unit cell dimensions (Å)	-	-	-	-
-a	69.94	45.47	72.55	84.17
-b	76.22	65.40	136.87	84.17
-c	95.59	76.17	114.83	79.36
Cell angles (degrees)	-	-	-	-
-alpha	89.97	90.00	90.00	90.00
-beta	89.97	90.00	90.00	90.00
-gamma	90.00	90.00	90.00	120.00
Molecules per asymmetric unit	8	1	2	1
Solvent content	47.95	38.42	53.48	59.60
Matthew's Coefficient	2.36	2.00	2.64	3.04
Ligands	Acetate, Cl^-, Mg^{2+}	Cl^-, Mg^{2+}, malate	Acetate, Ca^{2+}, PO_4^{2-}	Cl^-, glycerol, Mg^{2+}, tartrate
X-ray source	Rotating Anode/Rigaku	NSLS X29A	NSLS X29A	NSLS X29A
Wavelength	1.5418	0.9791	0.9792	1.0750
Resolution	50.00-2.00	50.00-1.50	50.00-1.90	40.00-1.87
Reflections	133,986	37,187	45,325	27,292
Completeness (%)	93.2 (83.9)	100.0 (100.0)	99.8 (99.5)	99.1 (92.8)
I/sigma (I)	4.50 (0.60)	7.40 (3.00)	7.70 (2.00)	8.10 (1.00)
R_{sym}	0.101 (0.730)	0.067 (0.600)	0.053 (0.680)	0.069 (0.820)
R_{work} (R_{free})	21.8 (28.6)	14.5 (17.2)	19.0 (22.6)	20.9 (25.2)
R_{free} reflections (%)	3211 (3.0%)	1152 (3.1%)	1408 (3.1%)	826 (3.2%)
Average B factor (overall)	36.17	25.13	47.74	51.00
-amino acid residues	36.24	23.78	47.63	50.92
-waters	35.32	35.55	49.74	52.02
-ligands	42.94, 32.50, 9.15	27.87, 14.21, 16.22	38.21, 37.62, 37.76	52.69, 80.66, 39.26, 40.96
Root-mean square deviation	-	-	-	-
-bond lengths	0.009	0.011	0.009	0.011
-bond angles	1.177	1.387	1.202	1.352
Ramachandran favored/allowed (%)	98.3 (1.7)	98.7 (1.3)	98.7 (1.3)	97.0 (3.0)
Number of solvent molecules	982	243	275	110

Table S5. Data collection and refinement statistics for BT2127 E47.

PDB ID	3QUB	3R9K	3QUC	3QYP
Mutation	E47A	E47D	E47N	E47N
Cap orientation	Closed	Closed	Closed	Closed
Space group	P3121	P3121	P3121	P21
Unit cell dimensions (Å)	-	-	-	-
-a	85.38	84.60	84.72	42.94
-b	85.38	84.60	84.72	77.18
-c	78.92	79.24	79.21	70.80
Cell angles (degrees)	-	-	-	-
-alpha	90.00	90.00	90.00	90.00
-beta	90.00	90.00	90.00	90.21
-gamma	120.00	120.00	120.00	90.00
Molecules per asymmetric unit	1	1	1	2
Solvent content	59.60	59.60	59.60	43.50
Matthew's Coefficient	3.04	3.04	3.04	2.17
Ligands	SO_4^{2-}	SO_4^{2-}	SO_4^{2-}	Ca^{2+}, Cl^-, glycerol, PO_4^{2-}, unknown
X-ray source	NSLS X29A	NSLS X29A	NSLS X29A	NSLS X29A
Wavelength	1.0750	1.0750	1.0750	0.9792
Resolution	40.00-1.90	50.00-1.80	50.00-1.87	50.00-1.60
Reflections	27,369	30,588	27,575	61,148
Completeness (%)	99.1 (99.6)	99.7 (100.0)	97.2 (95.3)	95.3 (69.0)
I/sigma (I)	4.60 (1.00)	8.50 (1.20)	4.30 (1.20)	9.50 (1.00)
R_{sym}	0.137 (0.770)	0.071 (0.800)	0.076 (.760)	0.065 (0.800)
R_{work} (R_{free})	19.8 (24.3)	19.0 (22.7)	21.6 (28.3)	18.5 (21.7)
R_{free} reflections (%)	846 (3.2%)	966 (3.2%)	703 (3.2%)	1776 (3.1%)
Average B factor (overall)	43.06	44.14	57.09	29.12
-amino acid residues	42.58	43.57	57.04	28.16
-waters	47.80	49.74	58.06	36.31
-ligands	32.25	38.55	58.69	23.05, 48.68, 56.89, 23.89, 41.27
Root-mean square deviation	-	-	-	-
-bond lengths	0.010	0.011	0.010	0.012
-bond angles	1.253	1.225	1.256	1.343
Ramachandran favored/allowed (%)	99.1 (0.9)	98.3 (1.7)	98.7 (1.3)	99.6 (0.4)
Number of solvent molecules	191	192	127	467

Table S6. The apparent first order rate constants for BT2127 catalyzed hydrolysis of phosphate esters and anhydrides at pH 7.5 and 25 ^{o}C. Reaction solutions initially contained 0.3 mM substrate, 8.4 μM BT2127, 1 mM $MgCl_2$, 1.0 unit/mL purine nucleoside phosphorylase and 0.2 mM MESG in 50 mM Tris (pH 7.5). The k_{obs} value was calculated by dividing the initial velocity of the reaction by the enzyme concentration.

Substrate	**k_{obs} (min^{-1})**	**Substrate**	**k_{obs} (min^{-1})**
Pyrophosphate	19	β-fucose-1-phosphate	0.10
Imidodiphosphate	0.06	GMP	0.38
Glyceraldehyde 3-phosphate	0.17	UDP	0.17
Glycerate-3-phosphate	0.31	ATP	0.20
D-erythrose-4-phosphate	0.20	dATP	0.18
Dihydroxyacetone phosphate	0.30	Farnesylpyrophosphate	0.059
Glycerol-2-phosphate	0.49	PLP	NA[a]
Glucose-1-phosphate	0.12	CTP	NA[a]
Glucose-6-phosphate	0.19	Serine-3-phosphate	NA[a]
Fructose-6-phosphate	0.32	Threonine phosphate	NA[a]
α-fucose-6-phosphate	0.21		

[a] NA represents no detectable activity.

Table S7. The apparent first order rate (k_{obs} (min^{-1})) constants for BT2127 catalyzed hydrolysis of phosphate esters and anhydrides at pH 7.5 and 25 ^{o}C. Reaction solutions initially contained 300 μM substrate, 8.4 μM mutant BT2127, 1 mM $MgCl_2$, 1.0 unit/mL purine nucleoside phosphorylase and 0.2 mM MESG in 50 mM Tris (pH 7.5). The k_{obs} value was calculated by dividing the initial velocity of the reaction by the enzyme concentration.

Substrate	wild-type	E47A	E47N	M20A	M20L	M20K
pyrophosphate	1.9×10^{1}	4.7×10^{-4}	5.5×10^{-4}	8.4×10^{-4}	3.1×10^{-3}	6.5×10^{-3}
D-ribose-5-P	2.0×10^{-1}	2.5×10^{-4}	7.3×10^{-4}	1.3×10^{-3}	2.1×10^{-3}	2.3×10^{-4}
glycerol-1-P	3.5×10^{-1}	4.4×10^{-4}	4.6×10^{-4}	4.8×10^{-3}	4.2×10^{-3}	4.0×10^{-3}
P-nitrophenylphosphate	5.8×10^{-3}	4.2×10^{-4}	3.8×10^{-4}	2.1×10^{-5}	2.6×10^{-3}	4.6×10^{-3}
fructose-6-P	5.3×10^{-3}	2.5×10^{-4}	9.5×10^{-5}	4.4×10^{-4}	1.6×10^{-3}	3.6×10^{-4}
UMP	1.5×10^{-1}	2.16×10^{-4}	2.6×10^{-4}	3.16×10^{-4}	1.6×10^{-3}	3.7×10^{-3}
imidodiphosphate	1.03×10^{-3}	6.2×10^{-5}	2.2×10^{-4}	6.0×10^{-3}	3.1×10^{-3}	4.4×10^{-3}
glucose-1-P	2.0×10^{-3}	1.1×10^{-4}	9.5×10^{-5}	8.4×10^{-4}	1.3×10^{-3}	2.9×10^{-3}
glucose-6-P	3.2×10^{-3}	1.9×10^{-4}	2.9×10^{-4}	2.8×10^{-4}	6.2×10^{-4}	3.7×10^{-3}
β-glucose-1,6-P	4.1×10^{-3}	3.0×10^{-5}	2.8×10^{-5}	6.3×10^{-4}	6.7×10^{-4}	1.5×10^{-4}
α-fucose 1-P	3.5×10^{-3}	2.3×10^{-4}	9.2×10^{-4}	2.5×10^{-4}	2.1×10^{-3}	1.6×10^{-3}
β-fucose 1-P	1.7×10^{-3}	5.5×10^{-4}	3.2×10^{-4}	2.3×10^{-4}	1.9×10^{-3}	9.1×10^{-4}
glycerol-2-P	8.16×10^{-3}	2.1×10^{-4}	2.7×10^{-4}	1.6×10^{-4}	1.6×10^{-3}	3.7×10^{-3}
serine-3-P	3.3×10^{-6}	1.4×10^{-4}	4.1×10^{-4}	3.5×10^{-4}	1.0×10^{-3}	3.5×10^{-3}

Table S8. Structures of HADSF proteins that have cap domains similar to that of BT2127.

PDB	Z-score	Organism	Name
3DV9	14.7	Bacteriodes vulgatis	putative BPGM
3NAS	10.6	Bacillus subtilis	putative BPGM
1LVH	10.2	Lactoccocus lactis	BPGM
3D6J	10.2	Bacteriodes fragilis	putative HAD member
2HSZ	9.5	Haemophilus somnus	predicted phosphatase
2HDO	9.2	Lactobacillus plantarum	putative phosphoglycolate phosphatase
3MC1	8.8	Clostridium acetobutylicum	predicted phosphatase
2AH5	8.7	Streptococcus pneumoniae	putative HAD member
2GO7	8.3	Streptococcus pneumoniae	putative HAD member
1RDF	7.8	Bacillus cereus	phosphonoacetaldehyde hydrolase
3KBB	7.4	Thermotoga maritima	putative BPGM
2FI1	7.4	Streptococcus pneumoniae	putative hydrolase
2HI0	7.4	Lactobacillus delbrueckii	putative phosphoglycolate phosphatase
1TE2	7.3	Escherichia coli K12	putative phosphatase Ynic
2HCF	7.3	Chlorobium tepidum	putative HAD member
3L5K	6.8	Homo sapiens	putative HAD member
2NYV	6.2	Aquifex aeolicus	putative phosphoglycolate phosphatase
2HOQ	5.1	Pyrococcus Horikoshii	putative HAD member
3E58-	5	Streptococcus thermophilus	putative BPGM
2FDR	4.5	Agrobacterium tumefaciens	putative HAD member
3ED5	4.5	Bacillus subtilis	putative HAD member
3IRU	4	Oleispira antarctica	putative phosphonoacetaldehyde hydrolase

Figure S1. Commassie blue stained SDS-PAGE gel of purified BT2127 (Q8A5V9).

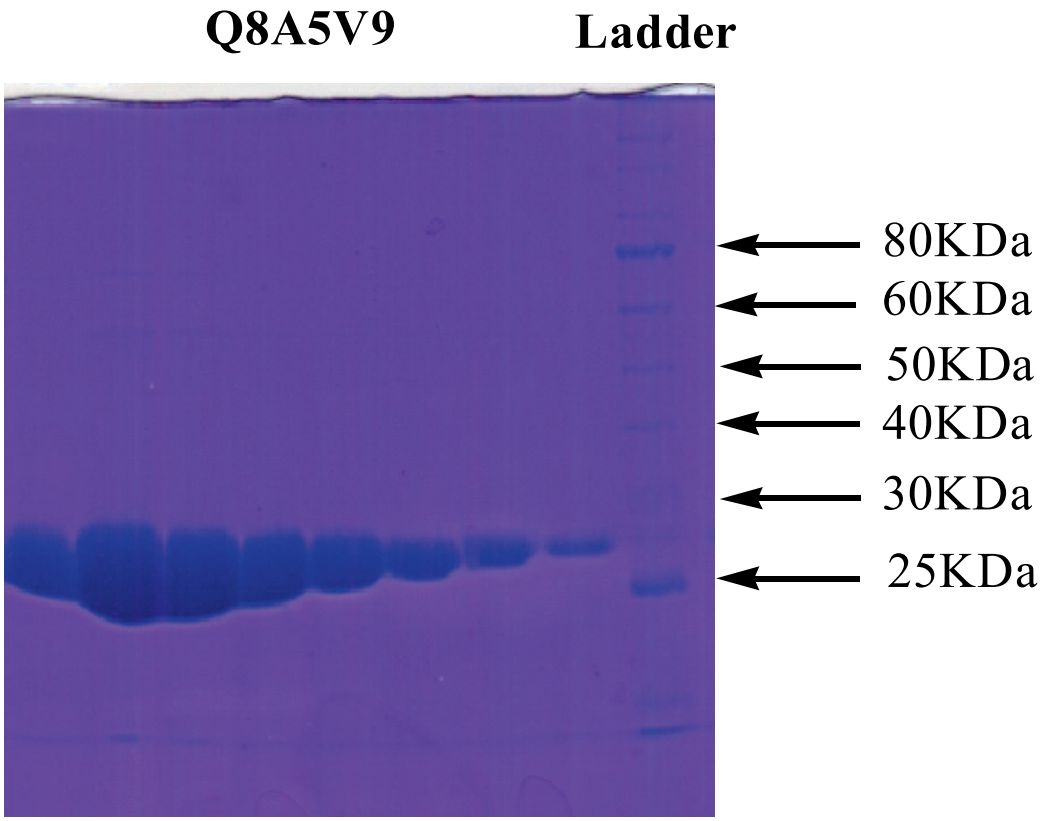

Figure S2. The pH-rate profiles measured for BT2127 catalyzed hydrolysis of pyrophosphate. (See Materials and Methods for details).

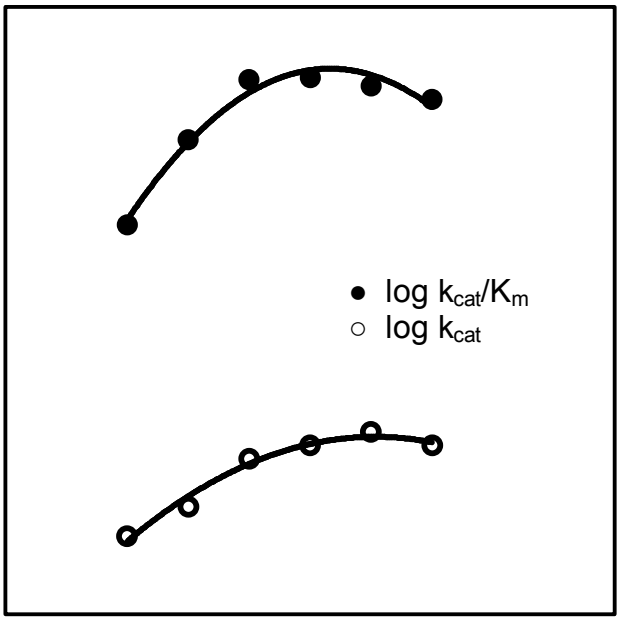

Figure S3. Double reciprocal plot of the initial velocities BT2127 catalyzed hydrolysis of inorganic pyrophosphate at varying concentration (3, 5, 8, 10, 15, 20 μM) measured as a function of imidodiphosphate concentration (0 μM, 10 μM, 35 μM, 50 μM). Reaction solutions contained 1 mM $MgCl_2$, 0.2 mM MESG and 1.0 unit/mL purine nucleoside phosphorylase in 50 mM Tris (pH 7.5; 25 ^{o}C). Data were fitted to equation 2.

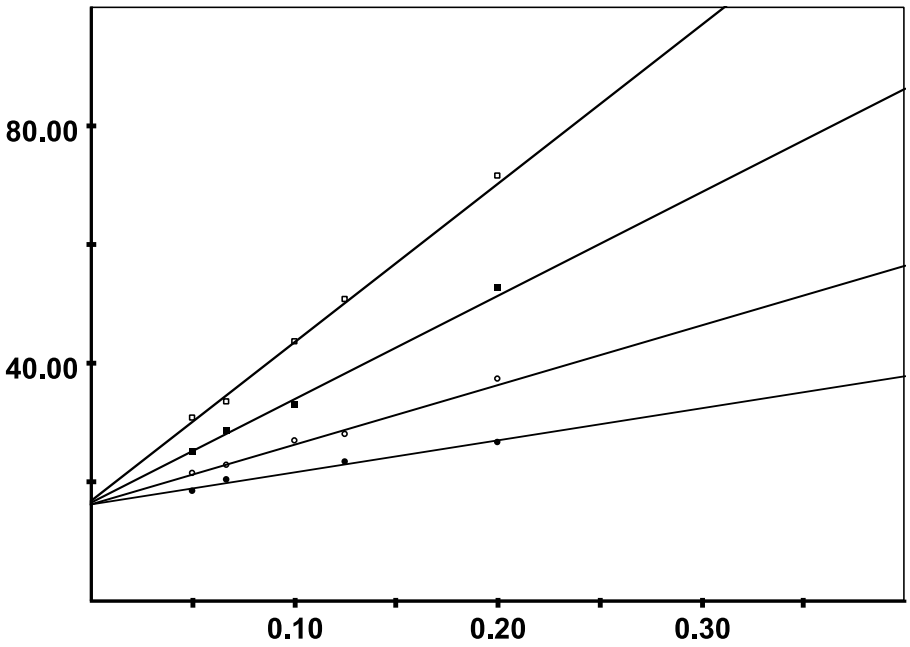

Figure S4. (Top) Superposition of BT2127 (gray ribbon) bound to phosphate and Mg^{2+} (magenta sphere) (PDB ID 3QX7) and β-PGM (cyan ribbon) bound to Mg^{2+} and β-glucose 1,6-(bis)phosphate (PDB ID 1O08). (Bottom) The active site of the overlayed structures.

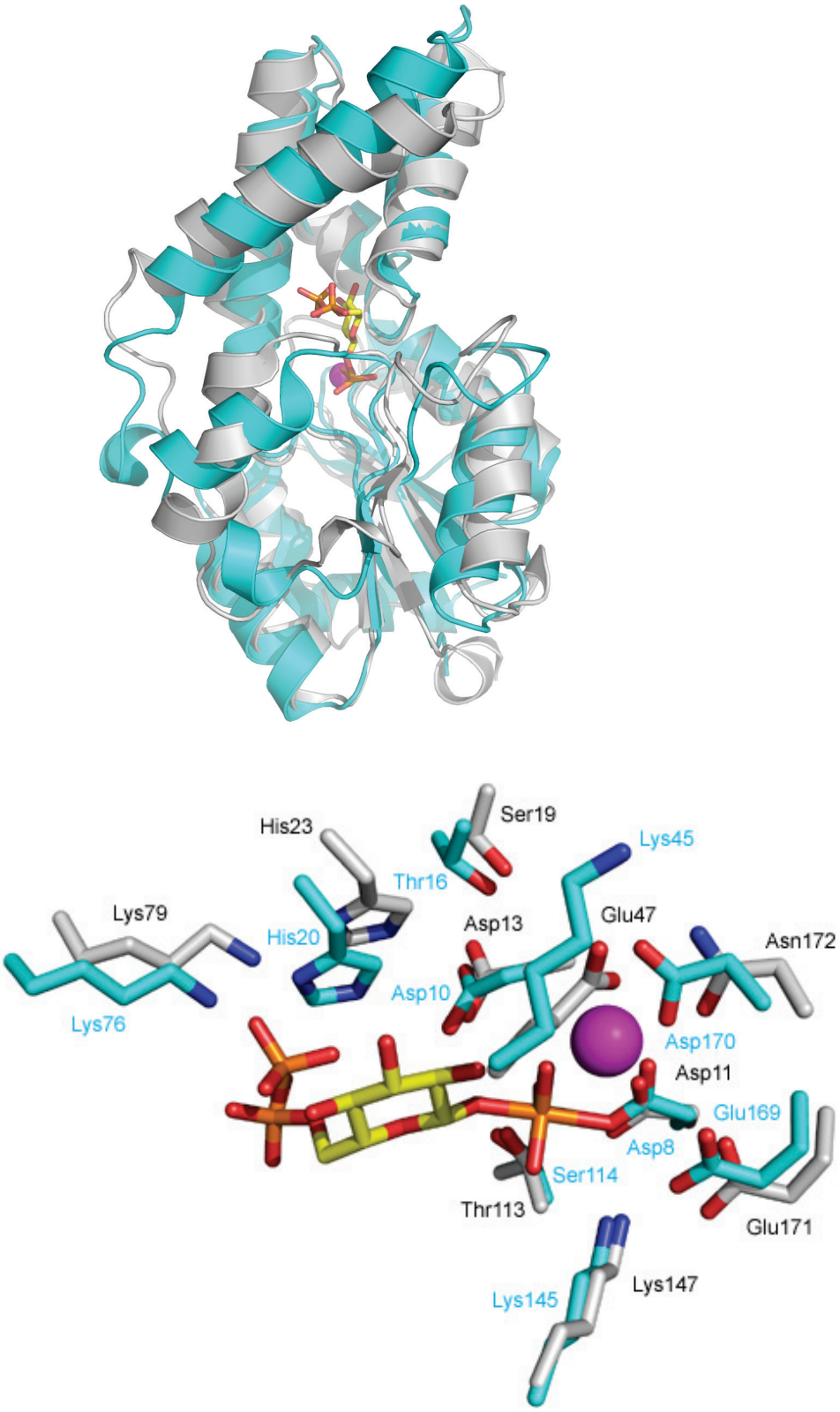

Figure S5. Multiple alignment of the amino acid sequences of BT2127 orthologues.

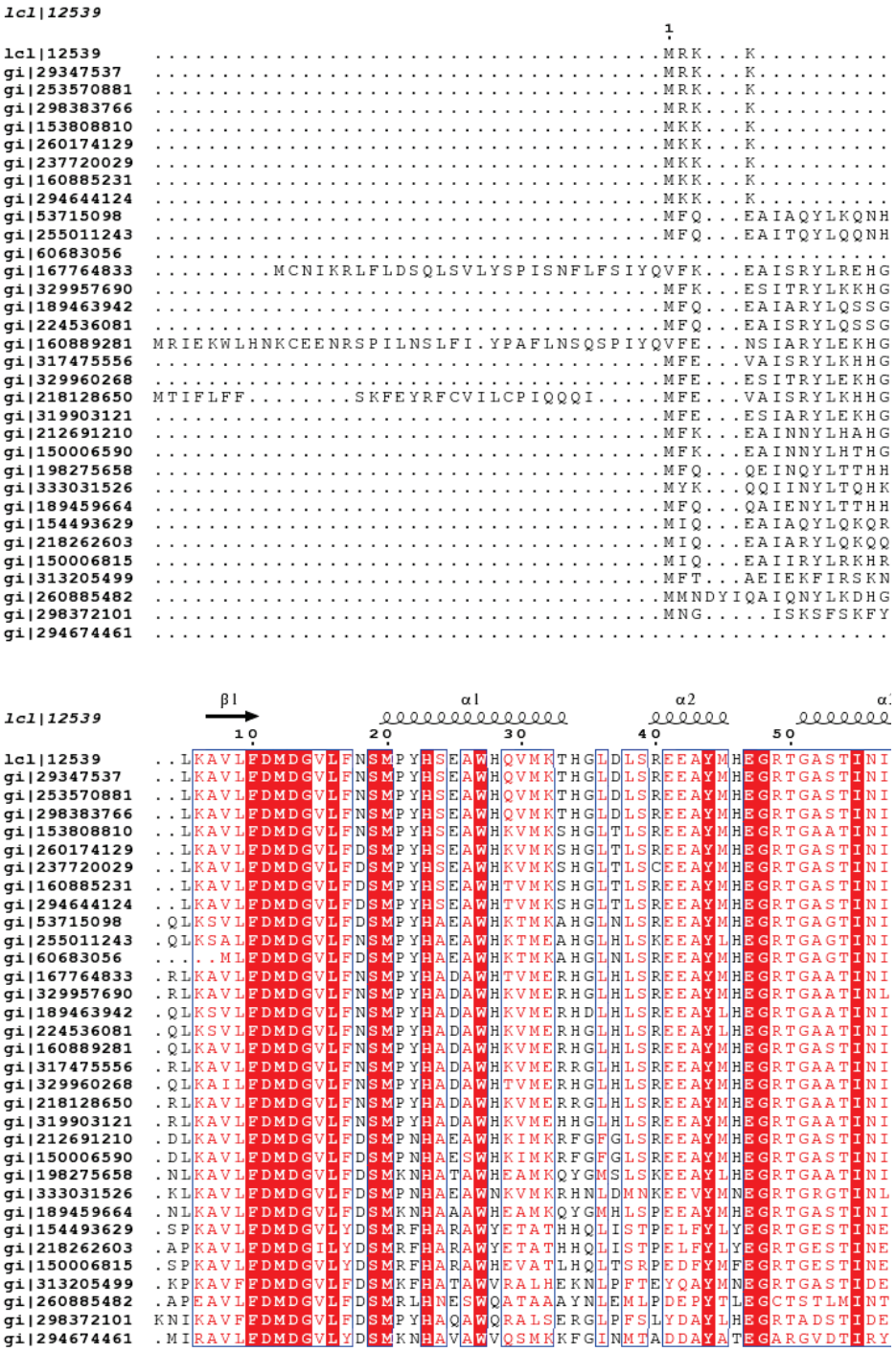

Figure S5 Continued

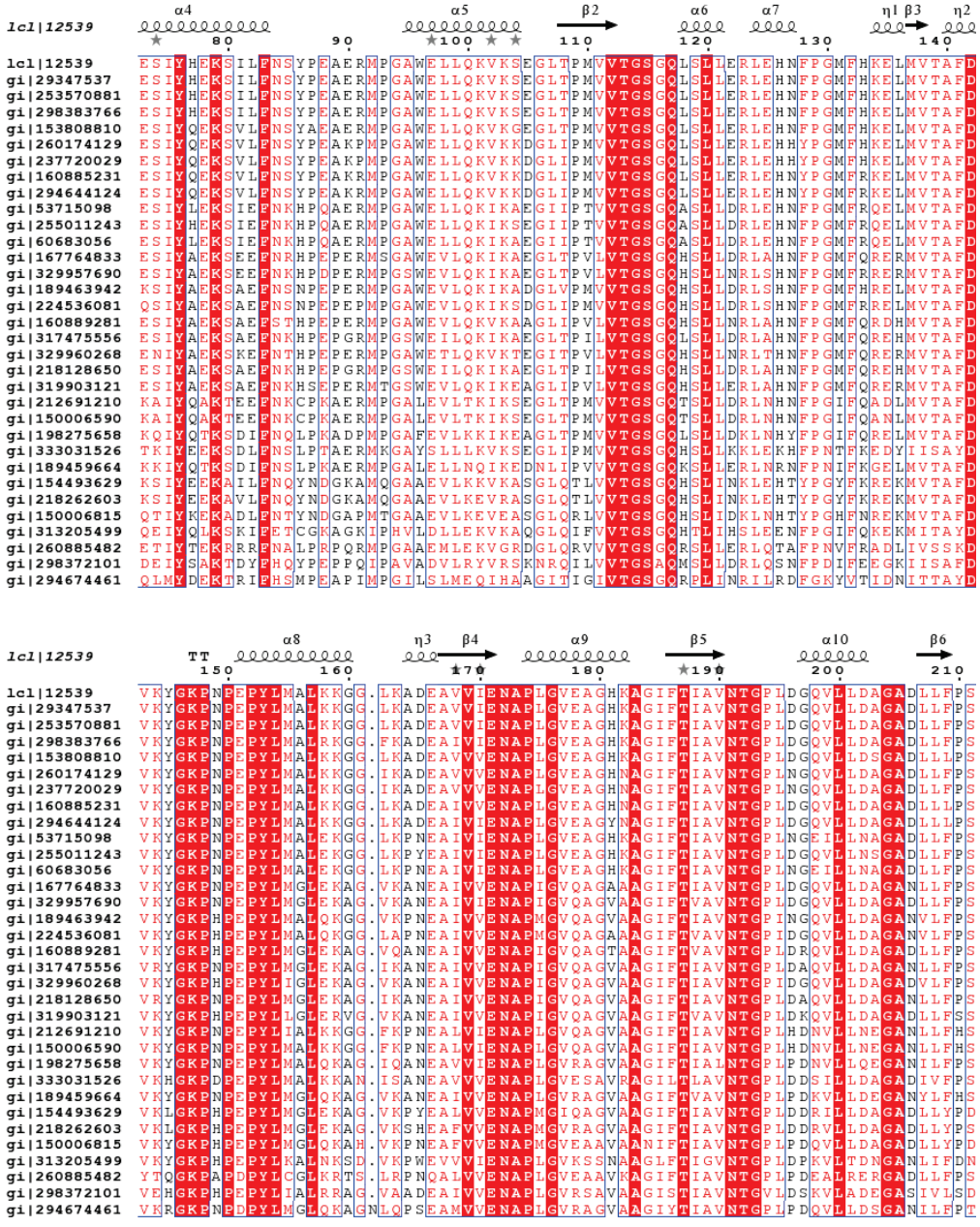

Figure S5 Continued

Figure S6. Interactions at the BT2127 cap domain (yellow)-catalytic domain (gray) interface.

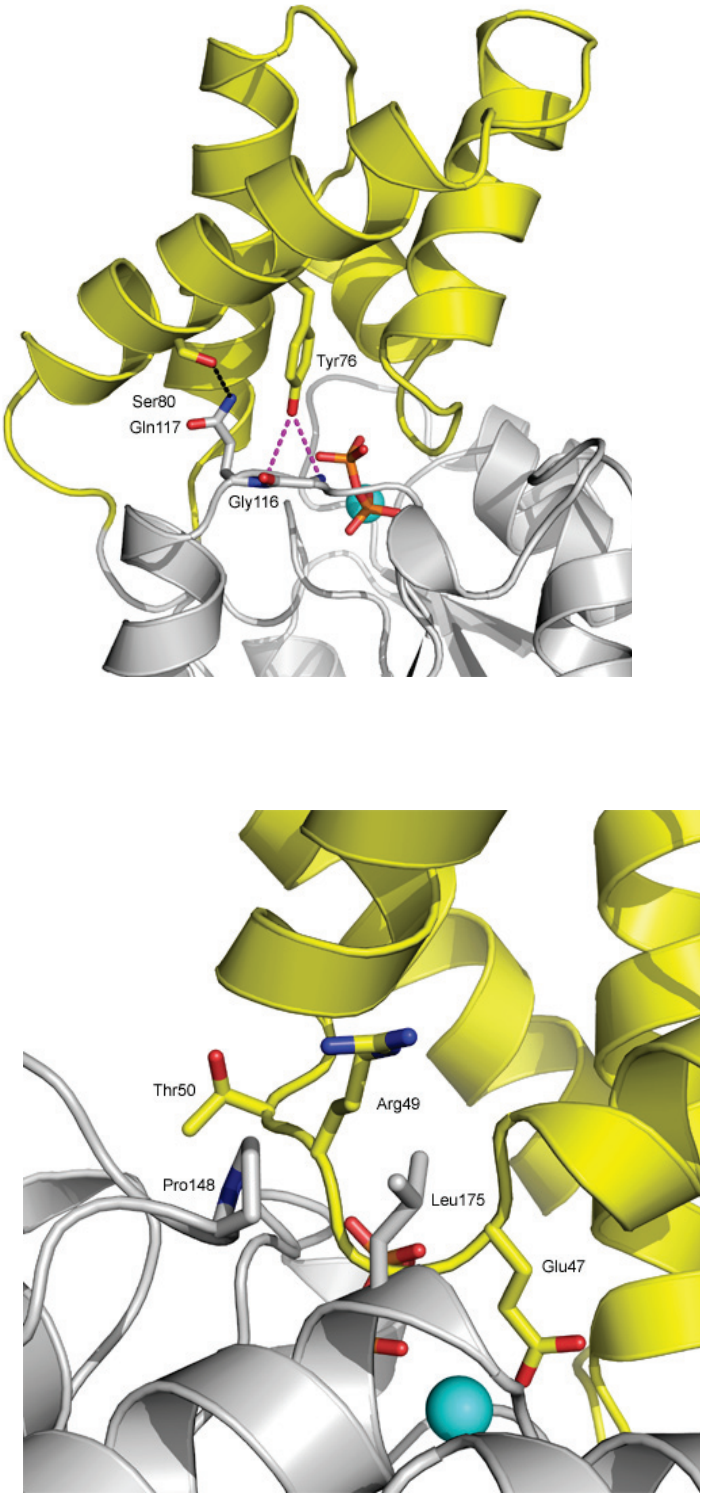

Figure S7. Alignment of the BT2127 sequence with the sequences of more distantly related homologues that do not conserve the marker residue Asn172.

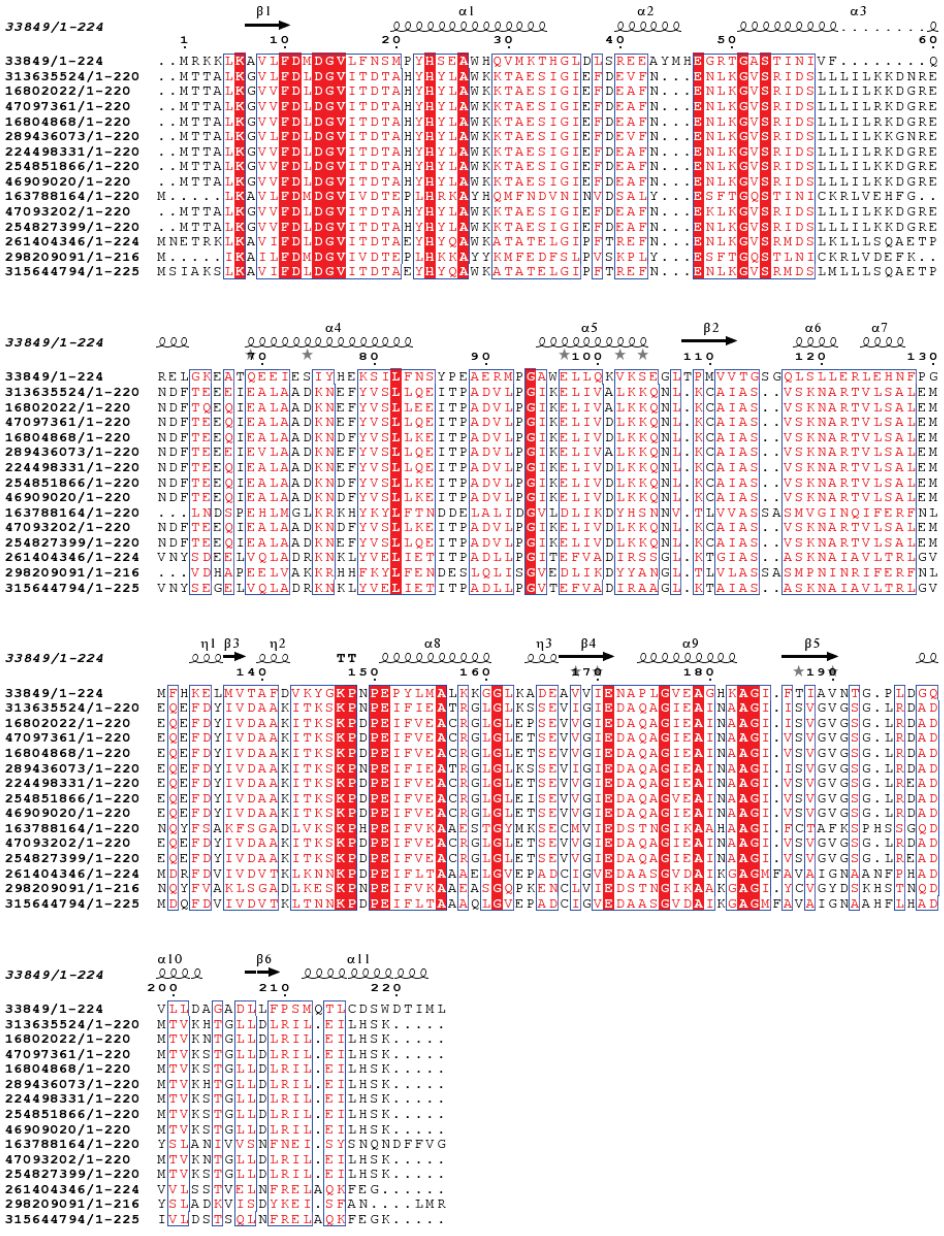

Figure S8. Left: The hydrogen bond network (black dashed line) of *L. lactis* β-PGM with Mg^{2+} (cyan sphere) and β-glucose 1,6-(bis)phosphate as ball and stick (phosphorus orange). Center: The placement of Asp10 in the cap-open conformation (arrows indicate the direction of movement of the Thr16 backbone and Asp10 side chain in going to the cap-closed conformation. Right: The placement of Asp10 in the cap-closed conformation.

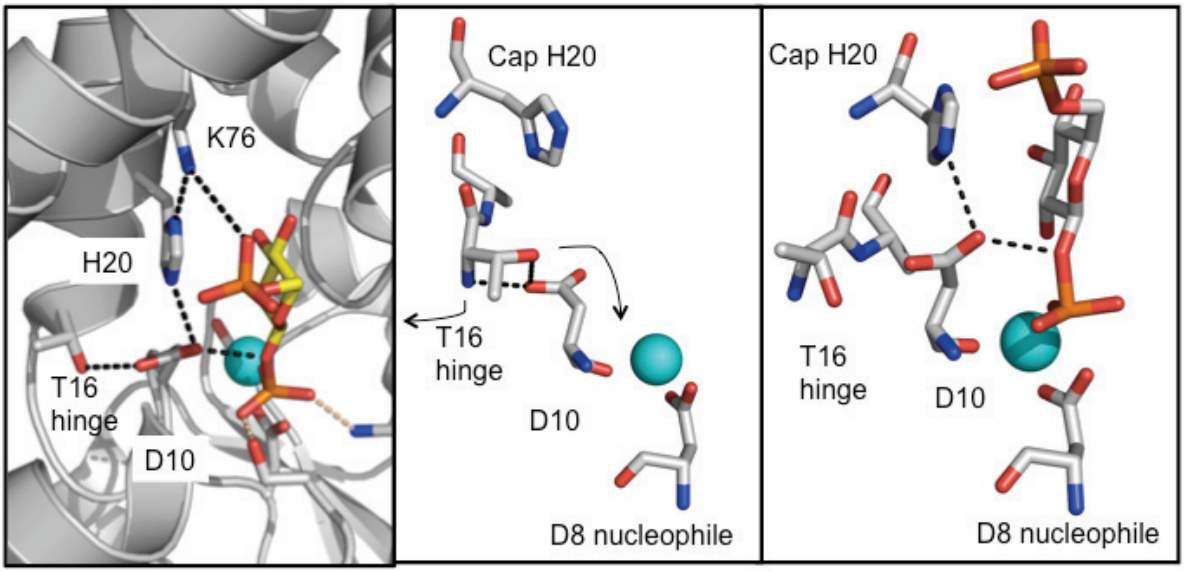

Figure S9. Multiple sequence alignment of the amino acid sequences of TON0002 orthologues.