Full wwPDB X-ray Structure Validation Report (i)

May 23, 2024 – 09:23 AM EDT

PDB ID	:	2GMW
Title	:	Crystal Structure of D,D-heptose 1.7-bisphosphate phosphatase from E. Coli.
Authors	:	Zhang, K.; DeLeon, G.; Wright, G.D.; Junop, M.S.
Deposited on	:	2006-04-07
Resolution	:	1.50 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (i)) were used in the production of this report:

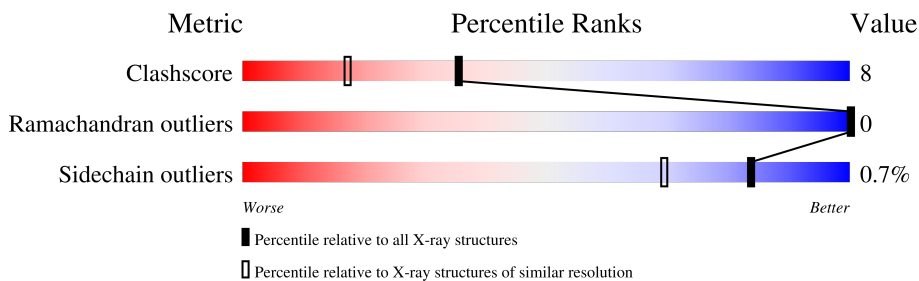|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	FAILED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.50 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.

Metric	Percentile Ranks	Value
Clashscore		8
Ramachandran outliers		0
Sidechain outliers		0.7%

Worse *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	141614	3144 (1.50-1.50)
Ramachandran outliers	138981	3066 (1.50-1.50)
Sidechain outliers	138945	3064 (1.50-1.50)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5%

Note EDS failed to run properly.

Mol	Chain	Length	Quality of chain
1	A	211	81% 5% 14%
1	B	211	75% 12% 13%

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 3600 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called D,D-heptose 1,7-bisphosphate phosphatase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	182	Total	C	N	O	S	0	13	0
			1492	947	249	285	11			
1	B	184	Total	C	N	O	S	0	20	0
			1555	995	255	293	12			

There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	cloning artifact	UNP P63228
A	2	GLY	-	cloning artifact	UNP P63228
A	3	SER	-	cloning artifact	UNP P63228
A	4	SER	-	cloning artifact	UNP P63228
A	5	HIS	-	cloning artifact	UNP P63228
A	6	HIS	-	cloning artifact	UNP P63228
A	7	HIS	-	cloning artifact	UNP P63228
A	8	HIS	-	cloning artifact	UNP P63228
A	9	HIS	-	cloning artifact	UNP P63228
A	10	HIS	-	cloning artifact	UNP P63228
A	11	SER	-	cloning artifact	UNP P63228
A	12	SER	-	cloning artifact	UNP P63228
A	13	GLY	-	cloning artifact	UNP P63228
A	14	LEU	-	cloning artifact	UNP P63228
A	15	VAL	-	cloning artifact	UNP P63228
A	16	PRO	-	cloning artifact	UNP P63228
A	17	ARG	-	cloning artifact	UNP P63228
A	18	GLY	-	cloning artifact	UNP P63228
A	19	SER	-	cloning artifact	UNP P63228
A	20	HIS	-	cloning artifact	UNP P63228
B	1	MET	-	cloning artifact	UNP P63228
B	2	GLY	-	cloning artifact	UNP P63228
B	3	SER	-	cloning artifact	UNP P63228
B	4	SER	-	cloning artifact	UNP P63228
B	5	HIS	-	cloning artifact	UNP P63228

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	6	HIS	-	cloning artifact	UNP P63228
B	7	HIS	-	cloning artifact	UNP P63228
B	8	HIS	-	cloning artifact	UNP P63228
B	9	HIS	-	cloning artifact	UNP P63228
B	10	HIS	-	cloning artifact	UNP P63228
B	11	SER	-	cloning artifact	UNP P63228
B	12	SER	-	cloning artifact	UNP P63228
B	13	GLY	-	cloning artifact	UNP P63228
B	14	LEU	-	cloning artifact	UNP P63228
B	15	VAL	-	cloning artifact	UNP P63228
B	16	PRO	-	cloning artifact	UNP P63228
B	17	ARG	-	cloning artifact	UNP P63228
B	18	GLY	-	cloning artifact	UNP P63228
B	19	SER	-	cloning artifact	UNP P63228
B	20	HIS	-	cloning artifact	UNP P63228

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total 1	Zn 1	0	0
2	B	1	Total 1	Zn 1	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	282	Total 282	O 282	0	0
3	B	269	Total 269	O 269	0	0

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.
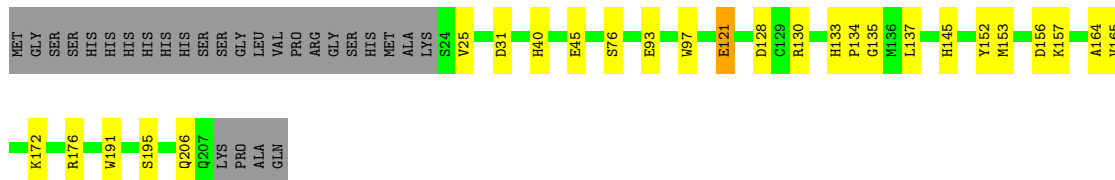
Note EDS failed to run properly.

• Molecule 1: D,D-heptose 1,7-bisphosphate phosphatase

Chain A:

• Molecule 1: D,D-heptose 1,7-bisphosphate phosphatase

Chain B:

4 Data and refinement statistics ⓘ

EDS failed to run properly - this section is therefore incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	51.90Å　　63.98Å　　103.32Å 90.00°　　90.00°　　　90.00°	Depositor
Resolution (Å)	25.00　　–　　1.50	Depositor
% Data completeness (in resolution range)	100.0 (25.00-1.50)	Depositor
R_{merge}	0.03	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.51 (at 1.50Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.163　，　0.212	Depositor
Wilson B-factor (Å^2)	19.0	Xtriage
Anisotropy	0.437	Xtriage
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	3600	wwPDB-VP
Average B, all atoms (Å^2)	23.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 56.52 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 2.7043e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

---

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: ZN

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#$|Z|$>5	RMSZ	#$|Z|$>5
1	A	0.63	0/1562	0.68	0/2118
1	B	0.59	0/1652	0.67	0/2242
All	All	0.61	0/3214	0.68	0/4360

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1492	0	1469	13	0
1	B	1555	0	1549	36	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	282	0	0	4	0
3	B	269	0	0	8	0
All	All	3600	0	3018	49	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 8.

All (49) close contacts within the same asymmetric unit are listed below, sorted by their clash

magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:137[A]:LEU:HD21	1:B:153[A]:MET:CE	1.41	1.50
1:B:137[A]:LEU:HD21	1:B:153[A]:MET:HE2	1.30	1.12
1:B:93[A]:GLU:HG2	1:B:97[A]:TRP:CZ3	1.87	1.08
1:B:137[A]:LEU:CD2	1:B:153[A]:MET:CE	2.32	1.08
1:B:137[A]:LEU:HD21	1:B:153[A]:MET:HE1	1.05	1.02
1:B:137[A]:LEU:CD2	1:B:153[A]:MET:HE1	1.91	0.98
1:A:61:GLU:HG3	3:A:497:HOH:O	1.62	0.96
1:B:128:ASP:O	1:B:133[B]:HIS:CD2	2.27	0.87
1:B:45:GLU:HG3	3:B:503:HOH:O	1.74	0.86
1:B:121[A]:GLU:H	1:B:121[A]:GLU:CD	1.80	0.82
1:B:93[A]:GLU:CG	1:B:97[A]:TRP:CZ3	2.65	0.80
1:B:137[A]:LEU:CD2	1:B:153[A]:MET:HE2	2.04	0.78
1:B:165[A]:VAL:CG1	3:B:563:HOH:O	2.36	0.73
1:A:137[B]:LEU:HD21	1:A:153[B]:MET:CE	2.21	0.70
1:B:165[A]:VAL:HG11	3:B:563:HOH:O	1.89	0.70
1:B:93[A]:GLU:HG2	1:B:97[A]:TRP:HZ3	1.52	0.69
1:A:137[B]:LEU:HD21	1:A:153[B]:MET:HE1	1.75	0.68
1:B:133[B]:HIS:HD2	1:B:135:GLY:H	1.40	0.67
1:B:93[A]:GLU:CG	1:B:97[A]:TRP:HZ3	2.05	0.67
1:B:76[A]:SER:OG	1:B:130:ARG:NH1	2.29	0.65
1:A:137[B]:LEU:HD11	1:A:153[B]:MET:CE	2.30	0.61
1:A:165:VAL:HG12	1:A:172:LYS:HD3	1.83	0.61
1:A:157:LYS:NZ	3:A:400:HOH:O	2.33	0.59
1:A:80:ARG:HG2	1:A:82:LYS:HE2	1.86	0.58
1:B:40:HIS:HD2	3:B:504:HOH:O	1.86	0.58
1:B:145:HIS:HE1	3:B:545:HOH:O	1.86	0.58
1:A:176[B]:ARG:NH2	3:A:312:HOH:O	2.40	0.55
1:A:137[B]:LEU:HD11	1:A:153[B]:MET:HE1	1.89	0.53
1:B:31:ASP:OD2	1:B:157:LYS:NZ	2.42	0.51
1:A:165:VAL:HG12	1:A:172:LYS:CD	2.43	0.49
1:B:133[B]:HIS:HD2	1:B:135:GLY:N	2.09	0.49
1:B:152:TYR:OH	1:B:206:GLN:NE2	2.47	0.47
1:B:25:VAL:HG21	3:B:519:HOH:O	2.13	0.47
1:B:191:TRP:CH2	1:B:206:GLN:NE2	2.83	0.47
1:B:128:ASP:O	1:B:133[B]:HIS:HD2	1.93	0.46
1:B:191:TRP:HH2	1:B:206:GLN:NE2	2.13	0.45
1:A:137[B]:LEU:HD21	1:A:153[B]:MET:HE2	1.95	0.45
1:B:93[A]:GLU:HG2	1:B:97[A]:TRP:CE3	2.48	0.45
1:B:165[A]:VAL:HG12	1:B:172:LYS:HE3	2.00	0.44
1:B:133[B]:HIS:CD2	1:B:134:PRO:HD2	2.54	0.43
1:B:145:HIS:CE1	3:B:545:HOH:O	2.65	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:153[A]:MET:HE3	1:B:164:ALA:HB2	2.01	0.42
1:B:156:ASP:OD1	1:B:157:LYS:NZ	2.51	0.42
1:A:56:ILE:HD13	1:A:102[A]:ARG:HD3	2.02	0.41
1:B:165[A]:VAL:HG12	1:B:172:LYS:HD3	2.02	0.41
1:B:176[B]:ARG:NH2	3:B:349:HOH:O	2.53	0.41
1:A:102[A]:ARG:NH1	3:A:577:HOH:O	2.53	0.41
1:B:165[A]:VAL:HG12	1:B:172:LYS:CD	2.51	0.41

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	193/211 (92%)	189 (98%)	4 (2%)	0	100	100
1	B	203/211 (96%)	199 (98%)	4 (2%)	0	100	100
All	All	396/422 (94%)	388 (98%)	8 (2%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	164/176 (93%)	164 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	B	174/176 (99%)	171 (98%)	3 (2%)	60	33
All	All	338/352 (96%)	335 (99%)	3 (1%)	84	61

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	121[A]	GLU
1	B	121[B]	GLU
1	B	195	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	145	HIS
1	A	201	GLN
1	B	206	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6　Fit of model and data ⓘ

6.1　Protein, DNA and RNA chains ⓘ

EDS failed to run properly - this section is therefore empty.

6.2　Non-standard residues in protein, DNA, RNA chains ⓘ

EDS failed to run properly - this section is therefore empty.

6.3　Carbohydrates ⓘ

EDS failed to run properly - this section is therefore empty.

6.4　Ligands ⓘ

EDS failed to run properly - this section is therefore empty.

6.5　Other polymers ⓘ

EDS failed to run properly - this section is therefore empty.