



wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 23, 2023 – 03:11 am GMT

PDB ID : 8A8F
Title : Crystal structure of Glc7 phosphatase in complex with the regulatory region of Ref2
Authors : Carminati, M.; Manav, C.M.; Bellini, D.; Passmore, L.A.
Deposited on : 2022-06-22
Resolution : 1.85 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 ⓘ 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 ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.8.4, CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.36
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0158
CCP4 : 7.0.044 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.36

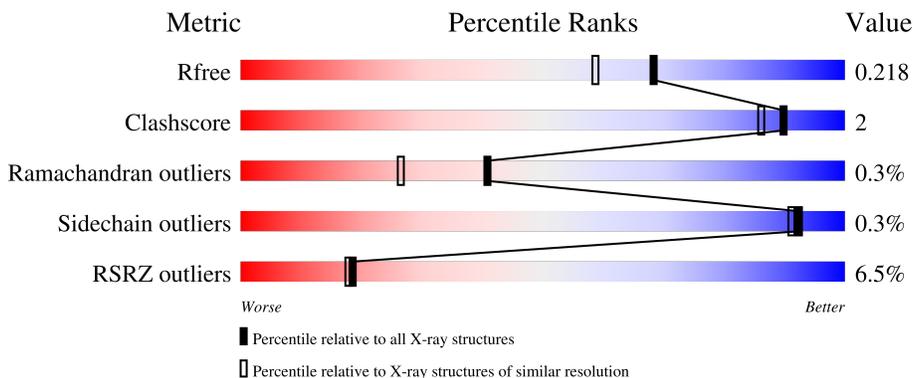
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.85 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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2469 (1.86-1.86)
Clashscore	141614	2625 (1.86-1.86)
Ramachandran outliers	138981	2592 (1.86-1.86)
Sidechain outliers	138945	2592 (1.86-1.86)
RSRZ outliers	127900	2436 (1.86-1.86)

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 $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	313	 3% 87% 6% • 6%
2	B	80	 19% 66% 8% 26%

2 Entry composition i

There are 5 unique types of molecules in this entry. The entry contains 2986 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 Serine/threonine-protein phosphatase PP1-2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	295	2397	1537	397	443	20	0	2	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	SER	-	expression tag	UNP P32598

- Molecule 2 is a protein called RNA end formation protein 2.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	B	59	461	292	82	87	0	0	0

There are 21 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	407	GLY	-	expression tag	UNP P42073
B	408	SER	-	expression tag	UNP P42073
B	409	GLY	-	expression tag	UNP P42073
B	410	SER	-	expression tag	UNP P42073
B	411	GLY	-	expression tag	UNP P42073
B	412	SER	-	expression tag	UNP P42073
B	413	GLY	-	expression tag	UNP P42073
B	414	SER	-	expression tag	UNP P42073
B	415	PRO	-	expression tag	UNP P42073
B	416	SER	-	expression tag	UNP P42073
B	417	TYR	-	expression tag	UNP P42073
B	418	SER	-	expression tag	UNP P42073
B	419	PRO	-	expression tag	UNP P42073
B	420	THR	-	expression tag	UNP P42073
B	421	SER	-	expression tag	UNP P42073

Continued on next page...

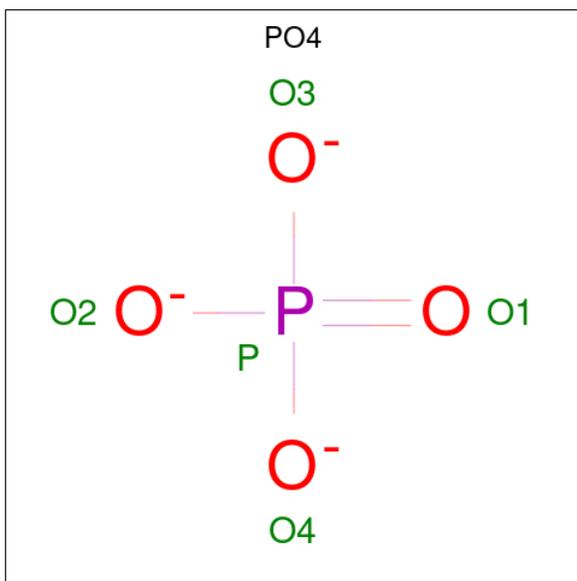
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	422	PRO	-	expression tag	UNP P42073
B	423	SER	-	expression tag	UNP P42073
B	424	TYR	-	expression tag	UNP P42073
B	425	SER	-	expression tag	UNP P42073
B	426	PRO	-	expression tag	UNP P42073
B	427	THR	-	expression tag	UNP P42073

- Molecule 3 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	2	Total Mn 2 2	0	0

- Molecule 4 is PHOSPHATE ION (three-letter code: PO4) (formula: O₄P).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total O P 5 4 1	0	0

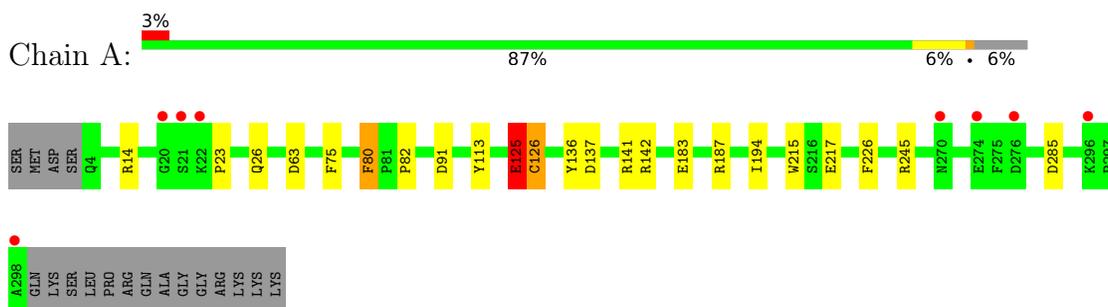
- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	116	Total O 116 116	0	0
5	B	5	Total O 5 5	0	0

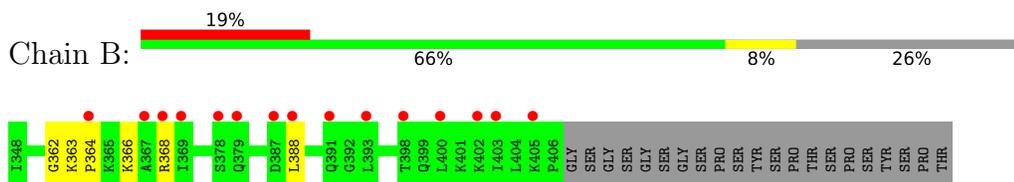
3 Residue-property plots [i](#)

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 and electron density. 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. A red dot above a residue indicates a poor fit to the electron density ($RSRZ > 2$). 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.

- Molecule 1: Serine/threonine-protein phosphatase PP1-2



- Molecule 2: RNA end formation protein 2



4 Data and refinement statistics

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α , β , γ	90.23Å 90.23Å 101.14Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	45.12 – 1.85 45.12 – 1.85	Depositor EDS
% Data completeness (in resolution range)	99.7 (45.12-1.85) 99.7 (45.12-1.85)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.68 (at 1.86Å)	Xtrriage
Refinement program	PHENIX 1.17.1_3660	Depositor
R, R_{free}	0.192 , 0.221 0.192 , 0.218	Depositor DCC
R_{free} test set	2090 reflections (5.10%)	wwPDB-VP
Wilson B-factor (Å ²)	43.2	Xtrriage
Anisotropy	0.400	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 56.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.026 for -h,-k,l	Xtrriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	2986	wwPDB-VP
Average B, all atoms (Å ²)	54.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.52% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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: PO4, MN

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	0.86	4/2458 (0.2%)	0.95	10/3322 (0.3%)
2	B	0.56	0/466	0.73	0/626
All	All	0.82	4/2924 (0.1%)	0.92	10/3948 (0.3%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	2
2	B	0	1
All	All	0	3

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	126	CYS	CB-SG	-8.29	1.68	1.82
1	A	217	GLU	CB-CG	-7.53	1.37	1.52
1	A	245	ARG	CG-CD	-6.39	1.35	1.51
1	A	136	TYR	CD1-CE1	5.91	1.48	1.39

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	63	ASP	CB-CG-OD1	9.35	126.72	118.30
1	A	125	GLU	OE1-CD-OE2	-7.73	114.03	123.30
1	A	245	ARG	NE-CZ-NH2	-6.86	116.87	120.30
1	A	245	ARG	CA-CB-CG	-5.73	100.80	113.40
1	A	285	ASP	CB-CG-OD1	5.61	123.35	118.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	125	GLU	Sidechain
1	A	80	PHE	Sidechain
2	B	362	GLY	Peptide

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	2397	0	2360	6	0
2	B	461	0	496	3	0
3	A	2	0	0	0	0
4	A	5	0	0	1	0
5	A	116	0	0	1	0
5	B	5	0	0	0	0
All	All	2986	0	2856	10	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 2.

The worst 5 of 10 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:A:26:GLN:HE22	1:A:141:ARG:HH12	1.33	0.76
1:A:126:CYS:HB2	1:A:194:ILE:HD12	1.70	0.71
2:B:363:LYS:HB3	2:B:364:PRO:HD2	1.88	0.55
4:A:403:PO4:P	5:A:507:HOH:O	2.71	0.49
2:B:366:LYS:HE2	2:B:368:ARG:HB2	1.97	0.45

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	295/313 (94%)	280 (95%)	14 (5%)	1 (0%)	41	26
2	B	57/80 (71%)	56 (98%)	1 (2%)	0	100	100
All	All	352/393 (90%)	336 (96%)	15 (4%)	1 (0%)	41	26

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	23	PRO

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	267/280 (95%)	266 (100%)	1 (0%)	91	89
2	B	55/72 (76%)	55 (100%)	0	100	100
All	All	322/352 (92%)	321 (100%)	1 (0%)	92	91

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	125	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	26	GLN
2	B	394	GLN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 3 ligands modelled in this entry, 2 are monoatomic - leaving 1 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
4	PO4	A	403	3	4,4,4	1.12	0	6,6,6	1.93	3 (50%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	403	PO4	O4-P-O2	2.56	116.18	107.97
4	A	403	PO4	O4-P-O1	2.43	119.80	110.89
4	A	403	PO4	O3-P-O2	-2.09	101.25	107.97

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	403	PO4	1	0

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 [i](#)

6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
1	A	295/313 (94%)	0.03	8 (2%) 54 53	37, 48, 73, 101	0
2	B	59/80 (73%)	1.39	15 (25%) 0 0	46, 72, 96, 108	0
All	All	354/393 (90%)	0.26	23 (6%) 18 18	37, 50, 86, 108	0

The worst 5 of 23 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	368	ARG	8.0
1	A	20	GLY	6.6
2	B	402	LYS	5.6
2	B	387	ASP	5.0
2	B	369	ILE	3.7

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
4	PO4	A	403	5/5	0.98	0.12	35,36,39,39	5
3	MN	A	401	1/1	0.99	0.11	38,38,38,38	1
3	MN	A	402	1/1	1.00	0.14	37,37,37,37	0

6.5 Other polymers [i](#)

There are no such residues in this entry.