Full wwPDB X-ray Structure Validation Report ⓘ

Apr 22, 2024 – 04:11 PM EDT

PDB ID	:	8U1E
Title	:	Apo protein tyrosine phosphatase 1B (PTP1B) at high resolution (1.43 A) in space group P43212 with two distinctly ordered chains
Authors	:	Sharma, S.; Mehlman, S.T.; Keedy, D.A.
Deposited on	:	2023-08-31
Resolution	:	1.43 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 ⓘ 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:

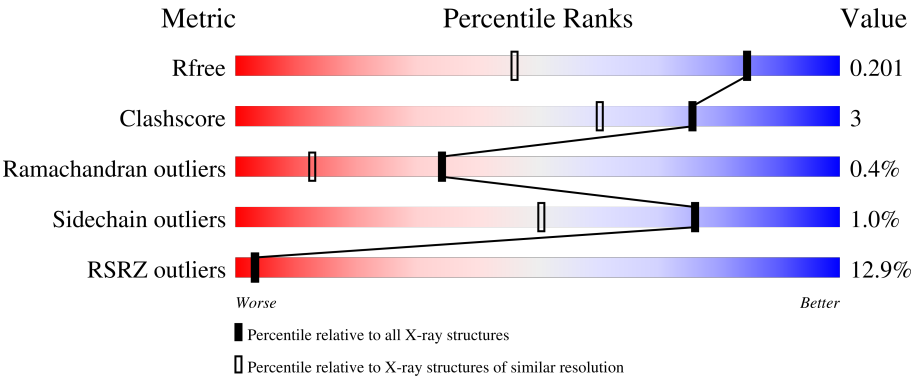
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 1.43 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	2021 (1.46-1.42)
Clashscore	141614	2086 (1.46-1.42)
Ramachandran outliers	138981	2047 (1.46-1.42)
Sidechain outliers	138945	2047 (1.46-1.42)
RSRZ outliers	127900	1993 (1.46-1.42)

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% 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	330	14% 76% 9% 15%
1	B	330	8% 80% 5% 14%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 10520 atoms, of which 5078 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 Tyrosine-protein phosphatase non-receptor type 1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	281	Total	C	H	N	O	S	0	49	0
			5077	1607	2545	438	466	21			
1	B	284	Total	C	H	N	O	S	0	44	0
			5052	1603	2533	430	463	23			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	initiating methionine	UNP P18031
A	322	LEU	-	expression tag	UNP P18031
A	323	GLU	-	expression tag	UNP P18031
A	324	HIS	-	expression tag	UNP P18031
A	325	HIS	-	expression tag	UNP P18031
A	326	HIS	-	expression tag	UNP P18031
A	327	HIS	-	expression tag	UNP P18031
A	328	HIS	-	expression tag	UNP P18031
A	329	HIS	-	expression tag	UNP P18031
B	0	MET	-	initiating methionine	UNP P18031
B	322	LEU	-	expression tag	UNP P18031
B	323	GLU	-	expression tag	UNP P18031
B	324	HIS	-	expression tag	UNP P18031
B	325	HIS	-	expression tag	UNP P18031
B	326	HIS	-	expression tag	UNP P18031
B	327	HIS	-	expression tag	UNP P18031
B	328	HIS	-	expression tag	UNP P18031
B	329	HIS	-	expression tag	UNP P18031

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Mg	0	0
			1	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	1	Total Mg 1 1	0	0

- Molecule 3 is water.

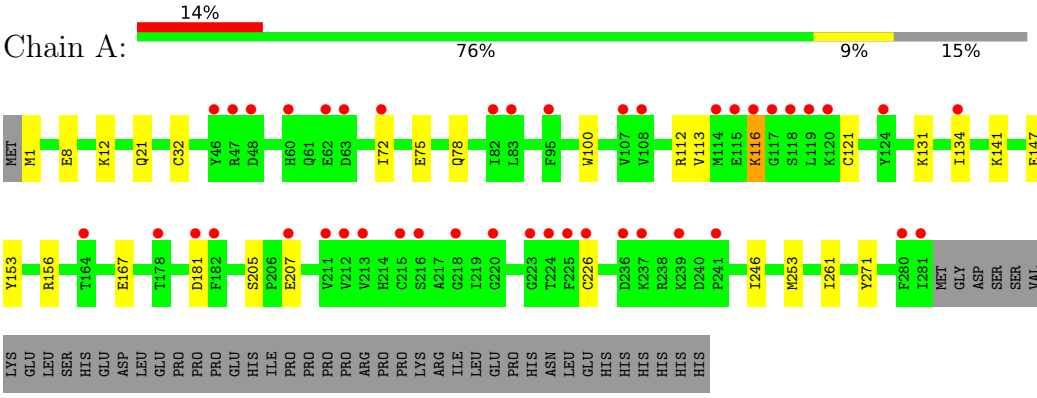
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	156	Total O 156 156	0	2
3	B	233	Total O 233 233	0	2

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.

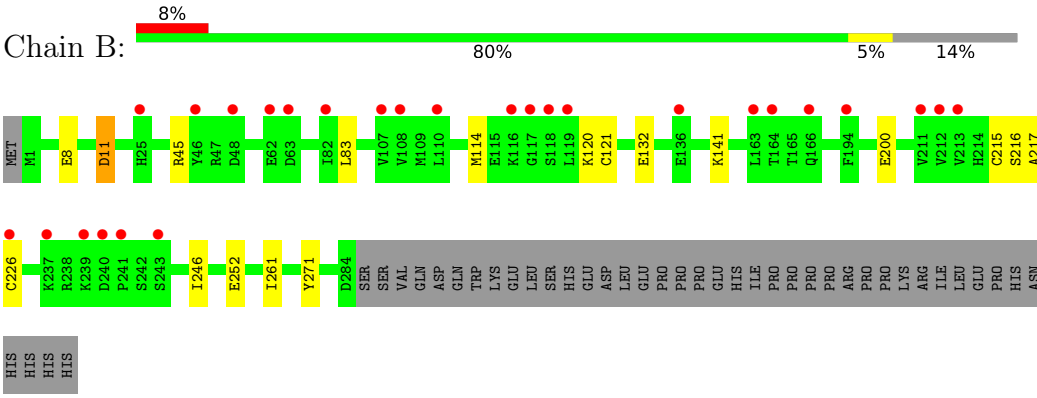
• Molecule 1: Tyrosine-protein phosphatase non-receptor type 1

Chain A:



• Molecule 1: Tyrosine-protein phosphatase non-receptor type 1

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α, β, γ	88.41Å 88.41Å 163.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.59 – 1.43 30.59 – 1.43	Depositor EDS
% Data completeness (in resolution range)	98.0 (30.59-1.43) 98.0 (30.59-1.43)	Depositor EDS
R_{merge}	0.15	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	0.98 (at 1.43Å)	Xtriage
Refinement program	PHENIX 1.19.2_4158	Depositor
R, R_{free}	0.150 , 0.203 0.149 , 0.201	Depositor DCC
R_{free} test set	5855 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	24.8	Xtriage
Anisotropy	0.085	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 47.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.98	EDS
Total number of atoms	10520	wwPDB-VP
Average B, all atoms (Å^2)	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.53% of the height of the origin peak. No significant pseudotranslation is detected.*

[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

5.1 Standard geometry

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

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	0.45	0/2764	0.65	0/3718
1	B	0.55	0/2735	0.72	2/3678 (0.1%)
All	All	0.50	0/5499	0.69	2/7396 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	11	ASP	CB-CG-OD1	6.20	123.88	118.30
1	B	11	ASP	CB-CG-OD2	-5.38	113.46	118.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

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	2532	2545	2352	22	0
1	B	2519	2533	2387	9	1
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	156	0	0	2	0
3	B	233	0	0	1	1
All	All	5442	5078	4739	31	2

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 3.

All (31) 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:83:LEU:HD11	1:B:226[B]:CYS:SG	2.27	0.74
1:B:252[B]:GLU:OE2	3:B:601:HOH:O	2.10	0.69
1:A:75[B]:GLU:O	1:A:78:GLN:NE2	2.30	0.60
1:A:150[B]:LYS:HD2	1:A:153[B]:TYR:CE1	2.37	0.60
1:A:226[B]:CYS:SG	1:A:253[B]:MET:CE	2.91	0.59
1:B:215[B]:CYS:SG	1:B:221:ARG:HD2	2.45	0.56
1:A:134[A]:ILE:HD11	1:A:141[A]:LYS:HD3	1.86	0.55
1:A:75[A]:GLU:O	1:A:78:GLN:NE2	2.30	0.53
1:A:150[B]:LYS:HD2	1:A:153[B]:TYR:CZ	2.43	0.53
1:A:21[A]:GLN:NE2	3:A:501[A]:HOH:O	2.41	0.53
1:B:246[B]:ILE:CG2	1:B:271:TYR:CZ	2.93	0.52
1:A:116:LYS:O	1:A:116:LYS:HD2	2.11	0.51
1:A:147:GLU:HG3	1:A:156:ARG:HG2	1.93	0.50
1:A:32[B]:CYS:SG	3:A:656:HOH:O	2.60	0.50
1:A:226[B]:CYS:SG	1:A:253[B]:MET:HE1	2.52	0.49
1:A:152[B]:TYR:CZ	1:A:153[B]:TYR:CE1	3.01	0.48
1:A:8:GLU:O	1:A:12:LYS:HG3	2.14	0.48
1:B:120:LYS:HA	1:B:120:LYS:HD2	1.70	0.45
1:B:45:ARG:NH2	1:B:121[B]:CYS:HA	2.32	0.45
1:A:113:VAL:HG13	1:A:121:CYS:O	2.17	0.44
1:A:246:ILE:HG21	1:A:271:TYR:CE1	2.53	0.44
1:A:152[B]:TYR:N	1:A:152[B]:TYR:CD1	2.86	0.44
1:A:100:TRP:CZ2	1:A:167[A]:GLU:HG3	2.54	0.43
1:B:114:MET:HE2	1:B:114:MET:HB2	1.91	0.43
1:A:112:ARG:NH2	1:A:181:ASP:OD1	2.51	0.42
1:B:132:GLU:HG3	1:B:141[A]:LYS:HE3	2.00	0.42
1:A:72[B]:ILE:N	1:A:72[B]:ILE:HD12	2.34	0.42
1:B:216[B]:SER:OG	1:B:217:ALA:N	2.53	0.41
1:A:116:LYS:O	1:A:116:LYS:CG	2.67	0.41
1:A:205[A]:SER:OG	1:A:207:GLU:CD	2.59	0.41

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:200:GLU:OE2	1:B:200:GLU:OE2[7_555]	2.02	0.18

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:B:773:HOH:O	3:B:775:HOH:O[6_445]	2.06	0.14

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	328/330 (99%)	319 (97%)	7 (2%)	2 (1%)	25 7
1	B	326/330 (99%)	316 (97%)	9 (3%)	1 (0%)	41 19
All	All	654/660 (99%)	635 (97%)	16 (2%)	3 (0%)	34 8

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	261[A]	ILE
1	A	261[B]	ILE
1	B	261	ILE

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	303/303 (100%)	300 (99%)	3 (1%)	76 50
1	B	299/303 (99%)	297 (99%)	2 (1%)	84 64
All	All	602/606 (99%)	597 (99%)	5 (1%)	76 61

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1	MET
1	A	116	LYS
1	A	131	LYS
1	B	8	GLU
1	B	11	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 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

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, 95^{th} 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	281/330 (85%)	0.80	45 (16%)	1	2	24, 35, 65, 104	2 (0%)
1	B	284/330 (86%)	0.39	28 (9%)	7	7	17, 25, 57, 85	1 (0%)
All	All	565/660 (85%)	0.60	73 (12%)	3	3	17, 31, 61, 104	3 (0%)

All (73) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	116	LYS	9.8
1	A	182	PHE	9.7
1	A	281	ILE	7.9
1	A	119	LEU	7.5
1	A	117	GLY	7.2
1	A	239	LYS	6.0
1	A	114	MET	5.5
1	A	152[A]	TYR	5.5
1	A	47	ARG	5.3
1	A	62	GLU	5.0
1	B	119	LEU	4.6
1	B	118	SER	4.5
1	B	117	GLY	4.5
1	A	215[A]	CYS	4.5
1	B	237	LYS	4.5
1	A	108	VAL	4.4
1	B	239	LYS	4.2
1	A	120	LYS	4.2
1	A	118	SER	4.1
1	A	213	VAL	4.1
1	A	83[A]	LEU	4.0
1	B	116	LYS	3.9
1	A	212	VAL	3.9
1	B	164	THR	3.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	280	PHE	3.7
1	A	95	PHE	3.6
1	A	82	ILE	3.6
1	B	63	ASP	3.5
1	B	108	VAL	3.4
1	A	46	TYR	3.4
1	A	48	ASP	3.3
1	B	62	GLU	3.3
1	A	181	ASP	3.0
1	B	166	GLN	3.0
1	A	225[A]	PHE	3.0
1	A	178	THR	2.9
1	B	240	ASP	2.8
1	A	107	VAL	2.8
1	B	25[A]	HIS	2.8
1	B	225	PHE	2.8
1	B	213	VAL	2.8
1	A	134[A]	ILE	2.7
1	A	224[A]	THR	2.7
1	A	151[A]	SER	2.7
1	B	107	VAL	2.6
1	A	115	GLU	2.6
1	A	60	HIS	2.6
1	A	226[A]	CYS	2.5
1	B	82	ILE	2.5
1	A	236	ASP	2.5
1	B	226[A]	CYS	2.5
1	B	163	LEU	2.5
1	B	243	SER	2.5
1	A	63	ASP	2.3
1	B	212	VAL	2.3
1	A	223	GLY	2.3
1	A	72[A]	ILE	2.2
1	B	241	PRO	2.2
1	B	46	TYR	2.2
1	A	207	GLU	2.2
1	A	218	GLY	2.1
1	A	220	GLY	2.1
1	A	237	LYS	2.1
1	B	211	VAL	2.1
1	A	164	THR	2.1
1	B	110	LEU	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	136	GLU	2.1
1	A	124	TYR	2.1
1	A	241	PRO	2.1
1	A	216[A]	SER	2.1
1	B	194	PHE	2.1
1	A	211	VAL	2.0
1	B	48	ASP	2.0

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, 95^{th} 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(Å^2)	Q<0.9
2	MG	A	401	1/1	0.85	0.07	70,70,70,70	0
2	MG	B	501	1/1	0.95	0.07	41,41,41,41	0

6.5 Other polymers (i)

There are no such residues in this entry.