



# Full wwPDB X-ray Structure Validation Report ⓘ

May 26, 2020 – 08:47 am BST

PDB ID : 6UZT  
Title : Crystal Structure of RPTP alpha  
Authors : Santelli, E.; Wen, Y.; Yang, S.; Svensson, M.N.D.; Stanford, S.M.; Bottini, N.  
Deposited on : 2019-11-15  
Resolution : 1.80 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](mailto: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 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	:	2.11
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)	:	2.11

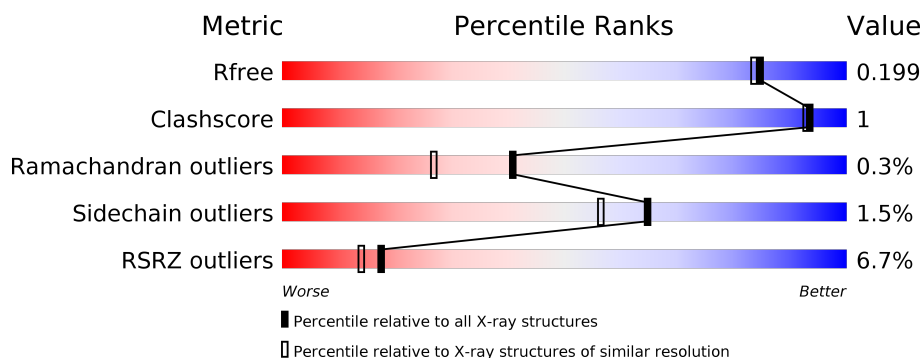
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.80 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	5950 (1.80-1.80)
Clashscore	141614	6793 (1.80-1.80)
Ramachandran outliers	138981	6697 (1.80-1.80)
Sidechain outliers	138945	6696 (1.80-1.80)
RSRZ outliers	127900	5850 (1.80-1.80)

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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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	601	<div> <div>4%</div> <div>94%</div> <div>• •</div> </div>
1	B	601	<div> <div>9%</div> <div>90%</div> <div>• • 5%</div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 11211 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 Receptor-type tyrosine-protein phosphatase alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	587	Total	C	N	O	S	0	20	0
			4879	3098	841	906	34			
1	B	572	Total	C	N	O	S	0	21	0
			4803	3046	832	890	35			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	201	MET	-	initiating methionine	UNP P18433
A	794	LEU	-	expression tag	UNP P18433
A	795	GLU	-	expression tag	UNP P18433
A	796	HIS	-	expression tag	UNP P18433
A	797	HIS	-	expression tag	UNP P18433
A	798	HIS	-	expression tag	UNP P18433
A	799	HIS	-	expression tag	UNP P18433
A	800	HIS	-	expression tag	UNP P18433
A	801	HIS	-	expression tag	UNP P18433
B	201	MET	-	initiating methionine	UNP P18433
B	794	LEU	-	expression tag	UNP P18433
B	795	GLU	-	expression tag	UNP P18433
B	796	HIS	-	expression tag	UNP P18433
B	797	HIS	-	expression tag	UNP P18433
B	798	HIS	-	expression tag	UNP P18433
B	799	HIS	-	expression tag	UNP P18433
B	800	HIS	-	expression tag	UNP P18433
B	801	HIS	-	expression tag	UNP P18433

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	825	Total	O	0	0
			825	825		

*Continued on next page...*

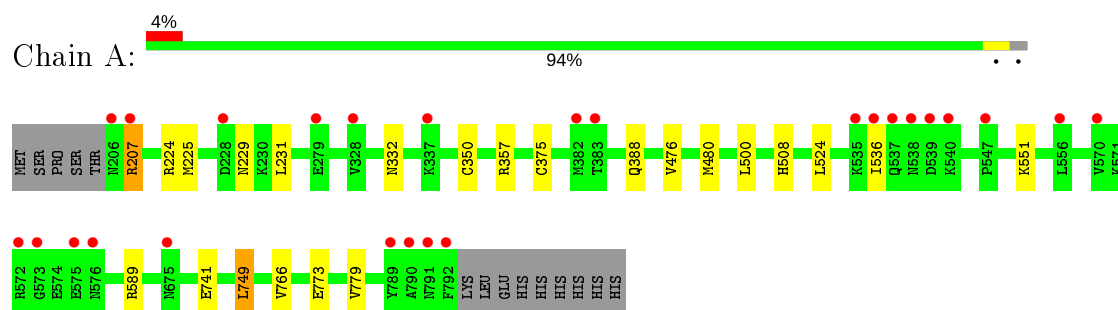
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	704	Total 704	O 704	0	2

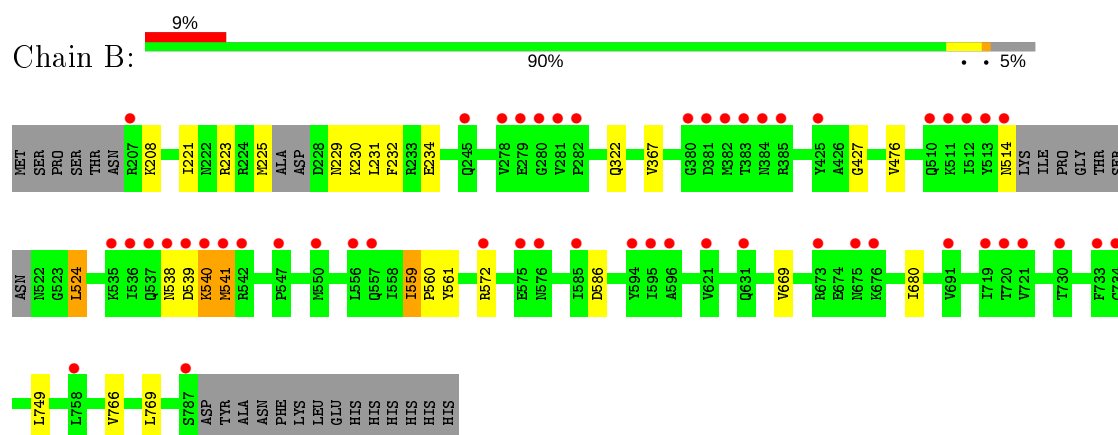
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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: Receptor-type tyrosine-protein phosphatase alpha



- Molecule 1: Receptor-type tyrosine-protein phosphatase alpha



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	104.72Å 112.11Å 136.37Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 1.80 29.94 – 1.80	Depositor EDS
% Data completeness (in resolution range)	99.9 (30.00-1.80) 99.9 (29.94-1.80)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.03 (at 1.80Å)	Xtriage
Refinement program	REFMAC 5.8.0257	Depositor
R, $R_{free}$	0.155 , 0.191 0.166 , 0.199	Depositor DCC
$R_{free}$ test set	7392 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	26.4	Xtriage
Anisotropy	0.453	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 44.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	11211	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	38.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.46% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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](#)

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	0.64	0/5041	0.78	1/6817 (0.0%)
1	B	0.65	0/4920	0.80	0/6645
All	All	0.65	0/9961	0.79	1/13462 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	207	ARG	NE-CZ-NH2	5.08	122.84	120.30

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	4879	0	4852	10	0
1	B	4803	0	4753	17	0
2	A	825	0	0	2	0
2	B	704	0	0	1	0
All	All	11211	0	9605	27	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 1.

All (27) 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:225[B]:MET:HG3	1:B:229[B]:ASN:OD1	1.98	0.64
1:B:541:MET:HB3	1:B:560:PRO:O	2.03	0.59
1:A:749:LEU:HD13	1:A:779:VAL:HG11	1.85	0.59
1:A:589[A]:ARG:NH1	1:A:741:GLU:OE2	2.36	0.59
1:B:230[B]:LYS:HD2	1:B:231[B]:LEU:HD12	1.85	0.56
1:A:225:MET:HA	1:A:229:ASN:HD22	1.72	0.55
1:B:541:MET:HE2	1:B:559:ILE:HG23	1.89	0.54
1:B:230[A]:LYS:HG2	1:B:231[A]:LEU:HD23	1.90	0.53
1:B:322:GLN:O	1:B:427:GLY:HA3	2.10	0.51
1:B:541:MET:CE	1:B:559:ILE:HG23	2.41	0.50
1:B:541:MET:HG3	1:B:561:TYR:HD1	1.78	0.48
1:B:221[B]:ILE:O	1:B:225[B]:MET:HB2	2.14	0.47
1:A:350[B]:CYS:SG	1:A:357:ARG:CZ	3.04	0.46
1:B:540:LYS:HG2	2:B:1142:HOH:O	2.15	0.46
1:A:375[B]:CYS:SG	1:A:388:GLN:CD	2.94	0.45
1:B:221[A]:ILE:O	1:B:225[A]:MET:HG3	2.16	0.44
1:A:224:ARG:HA	1:A:231:LEU:HB2	2.00	0.44
1:A:524:LEU:HD21	1:A:773:GLU:HB2	2.01	0.43
1:B:367:VAL:HG21	1:B:586:ASP:HB3	2.01	0.42
1:A:500:LEU:HD11	1:A:508:HIS:CG	2.55	0.42
1:B:230[A]:LYS:O	1:B:234:GLU:HB2	2.19	0.42
1:A:332:ASN:ND2	2:A:913:HOH:O	2.52	0.42
1:B:221[B]:ILE:O	1:B:225[B]:MET:N	2.46	0.42
1:B:541:MET:HG3	1:B:561:TYR:CD1	2.55	0.42
1:A:480[B]:MET:HE1	2:A:1308:HOH:O	2.20	0.41
1:B:669:VAL:HG21	1:B:680:ILE:HD12	2.03	0.41
1:B:524:LEU:HD12	1:B:769:LEU:HD11	2.03	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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	605/601 (101%)	583 (96%)	20 (3%)	2 (0%)	41	27
1	B	585/601 (97%)	566 (97%)	17 (3%)	2 (0%)	41	27
All	All	1190/1202 (99%)	1149 (97%)	37 (3%)	4 (0%)	41	27

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	476	VAL
1	B	476	VAL
1	B	766	VAL
1	A	766	VAL

### 5.3.2 Protein sidechains ⓘ

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	545/539 (101%)	541 (99%)	4 (1%)	84	81
1	B	533/539 (99%)	519 (97%)	14 (3%)	46	32
All	All	1078/1078 (100%)	1060 (98%)	18 (2%)	65	51

All (18) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	207	ARG
1	A	536	ILE
1	A	551	LYS
1	A	749	LEU
1	B	208	LYS
1	B	223[A]	ARG
1	B	223[B]	ARG
1	B	232[A]	PHE
1	B	232[B]	PHE
1	B	514	ASN
1	B	524	LEU
1	B	538	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	539	ASP
1	B	540	LYS
1	B	541	MET
1	B	559	ILE
1	B	572	ARG
1	B	749	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (8) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	229	ASN
1	A	332	ASN
1	A	474	GLN
1	A	487	GLN
1	A	538	ASN
1	B	355	ASN
1	B	388	GLN
1	B	514	ASN

### 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 carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

There are no ligands in this entry.

## 5.7 Other polymers ⓘ

There are no such residues in this entry.

## 5.8 Polymer linkage issues ⓘ

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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	587/601 (97%)	-0.06	26 (4%) 34 28	18, 30, 64, 142	0
1	B	572/601 (95%)	0.23	52 (9%) 9 7	19, 33, 76, 157	1 (0%)
All	All	1159/1202 (96%)	0.08	78 (6%) 17 14	18, 31, 73, 157	1 (0%)

All (78) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	382	MET	7.3
1	B	383	THR	7.2
1	B	536	ILE	7.0
1	B	538	ASN	6.6
1	B	207	ARG	6.4
1	A	572	ARG	6.2
1	B	281	VAL	5.8
1	B	381	ASP	5.4
1	B	380	GLY	5.4
1	B	384	ASN	5.4
1	B	572	ARG	4.5
1	B	425	TYR	4.5
1	A	573	GLY	4.5
1	B	539	ASP	4.4
1	B	541	MET	4.4
1	B	721	VAL	4.4
1	B	278	VAL	4.3
1	A	382	MET	4.3
1	B	537	GLN	4.2
1	A	207	ARG	4.2
1	B	535	LYS	4.2
1	A	536	ILE	4.0
1	A	575	GLU	4.0
1	B	514	ASN	3.9

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	539	ASP	3.9
1	B	279	GLU	3.9
1	B	512	ILE	3.9
1	B	556	LEU	3.7
1	B	513	TYR	3.6
1	A	538	ASN	3.5
1	B	540	LYS	3.5
1	A	792	PHE	3.4
1	A	537	GLN	3.4
1	B	282	PRO	3.4
1	B	675	ASN	3.4
1	A	789	TYR	3.3
1	B	280	GLY	3.3
1	B	734	CYS	3.2
1	A	570	VAL	3.2
1	A	540	LYS	3.2
1	A	790	ALA	3.1
1	B	245	GLN	3.1
1	B	510	GLN	3.1
1	B	595	ILE	3.0
1	B	385	ARG	3.0
1	B	511	LYS	3.0
1	B	787	SER	3.0
1	A	279	GLU	2.9
1	B	547	PRO	2.9
1	B	621	VAL	2.8
1	B	550	MET	2.8
1	A	535	LYS	2.8
1	B	585	ILE	2.8
1	B	730	THR	2.7
1	A	337	LYS	2.7
1	B	575	GLU	2.7
1	A	328	VAL	2.6
1	B	758	LEU	2.6
1	B	691	VAL	2.5
1	A	383	THR	2.4
1	B	733	PHE	2.4
1	B	673	ARG	2.4
1	B	631	GLN	2.4
1	A	206	ASN	2.4
1	A	675	ASN	2.3
1	B	596	ALA	2.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	720	THR	2.3
1	B	557	GLN	2.2
1	B	719	ILE	2.2
1	B	542	ARG	2.2
1	A	791	ASN	2.1
1	B	594	TYR	2.1
1	A	576	ASN	2.1
1	B	576	ASN	2.1
1	A	556	LEU	2.1
1	A	547	PRO	2.1
1	B	676	LYS	2.0
1	A	228	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 carbohydrates in this entry.

## 6.4 Ligands [i](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.