



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 8, 2026 – 02:46 PM UTC

PDB ID : 9R16 / pdb_00009r16
Title : Structure of mutant SHP2
Authors : Karunaraj, P.; Scheele, R.; Song, W.; Vasan, N.
Deposited on : 2025-04-25
Resolution : 2.63 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	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

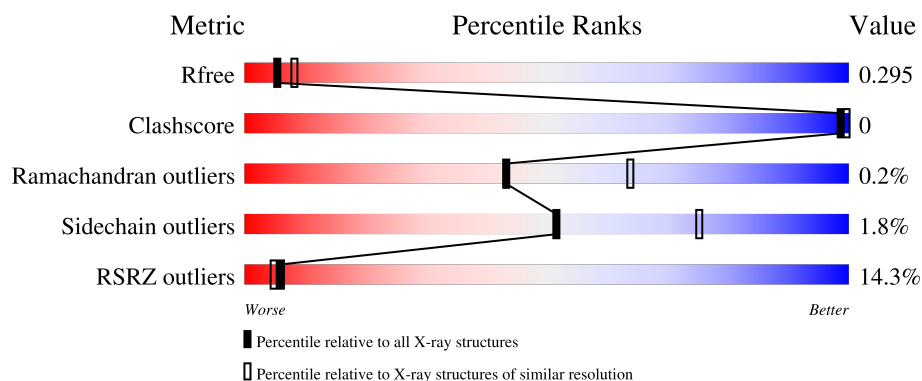
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.63 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}	180053	4951 (2.64-2.60)
Clashscore	190562	5303 (2.64-2.60)
Ramachandran outliers	187476	5217 (2.64-2.60)
Sidechain outliers	187428	5217 (2.64-2.60)
RSRZ outliers	180081	4950 (2.64-2.60)

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	536	13% 92% 5%
1	B	536	14% 92% 5%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 16635 atoms, of which 8119 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 11.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	510	Total	C	H	N	O	S	4070	3	0
			8209	2597	4070	743	780	19			
1	B	509	Total	C	H	N	O	S	4049	2	0
			8174	2586	4049	741	779	19			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	62	ASP	TYR	engineered mutation	UNP Q06124
A	529	LEU	-	expression tag	UNP Q06124
A	530	GLU	-	expression tag	UNP Q06124
A	531	HIS	-	expression tag	UNP Q06124
A	532	HIS	-	expression tag	UNP Q06124
A	533	HIS	-	expression tag	UNP Q06124
A	534	HIS	-	expression tag	UNP Q06124
A	535	HIS	-	expression tag	UNP Q06124
A	536	HIS	-	expression tag	UNP Q06124
B	62	ASP	TYR	engineered mutation	UNP Q06124
B	529	LEU	-	expression tag	UNP Q06124
B	530	GLU	-	expression tag	UNP Q06124
B	531	HIS	-	expression tag	UNP Q06124
B	532	HIS	-	expression tag	UNP Q06124
B	533	HIS	-	expression tag	UNP Q06124
B	534	HIS	-	expression tag	UNP Q06124
B	535	HIS	-	expression tag	UNP Q06124
B	536	HIS	-	expression tag	UNP Q06124

- Molecule 2 is water.

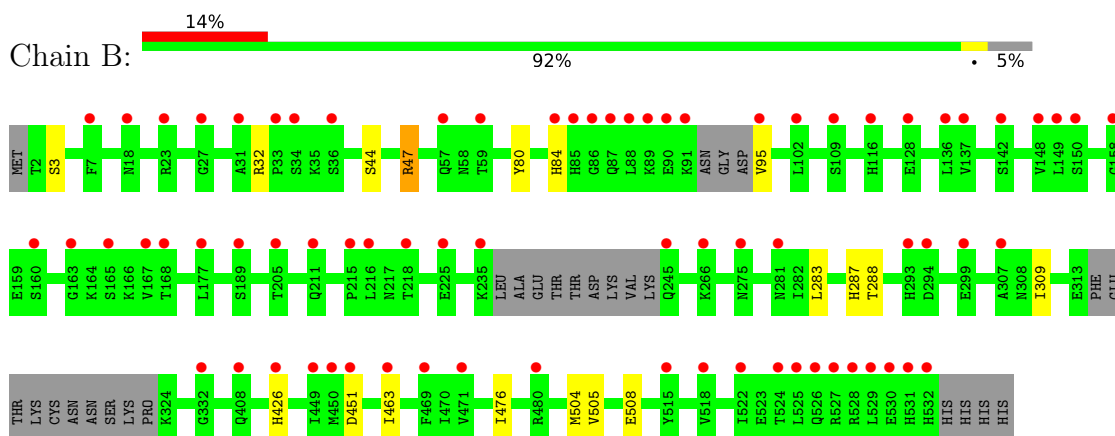
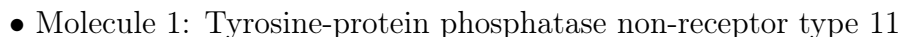
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	119	Total	O	0	0
			119	119		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	133	Total 133	O 133	0	0

- Molecule 1: Tyrosine-protein phosphatase non-receptor type 11



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	45.18Å 212.16Å 54.92Å 90.00° 96.46° 90.00°	Depositor
Resolution (Å)	27.78 – 2.63 27.78 – 2.63	Depositor EDS
% Data completeness (in resolution range)	69.1 (27.78-2.63) 69.0 (27.78-2.63)	Depositor EDS
R_{merge}	0.25	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.11 (at 2.64Å)	Xtriage
Refinement program	BUSTER 2.11.8	Depositor
R, R_{free}	0.280 , 0.304 0.270 , 0.295	Depositor DCC
R_{free} test set	1036 reflections (4.92%)	wwPDB-VP
Wilson B-factor (Å ²)	34.4	Xtriage
Anisotropy	0.034	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 46.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.83	EDS
Total number of atoms	16635	wwPDB-VP
Average B, all atoms (Å ²)	50.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.84% 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

5.1 Standard geometry

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	0.51	0/4236	0.77	0/5712
1	B	0.52	0/4219	0.77	0/5688
All	All	0.52	0/8455	0.77	0/11400

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

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	4139	4070	4056	3	0
1	B	4125	4049	4035	5	0
2	A	119	0	0	0	0
2	B	133	0	0	0	0
All	All	8516	8119	8091	8	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 0.

The worst 5 of 8 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:47:ARG:NH2	1:B:95:VAL:O	2.05	0.88

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:32:ARG:NH1	1:B:44:SER:OG	2.50	0.44
1:A:80:TYR:O	1:A:84:HIS:HB3	2.18	0.44
1:B:80:TYR:O	1:B:84:HIS:HB3	2.18	0.43
1:A:153:THR:O	1:A:165:SER:HA	2.19	0.43

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	505/536 (94%)	492 (97%)	12 (2%)	1 (0%)	43	64
1	B	503/536 (94%)	492 (98%)	10 (2%)	1 (0%)	43	64
All	All	1008/1072 (94%)	984 (98%)	22 (2%)	2 (0%)	43	64

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	505	VAL
1	B	505	VAL

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.

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	457/478 (96%)	449 (98%)	8 (2%)	51 75
1	B	455/478 (95%)	447 (98%)	8 (2%)	51 75
All	All	912/956 (95%)	896 (98%)	16 (2%)	51 75

5 of 16 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	476	ILE
1	B	451	ASP
1	B	3	SER
1	B	426	HIS
1	A	508	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	277	ASN
1	B	273	ASN
1	B	87	GLN
1	A	458	HIS
1	B	132	HIS

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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

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	510/536 (95%)	1.10	72 (14%) 6 5	7, 23, 41, 66	2 (0%)
1	B	509/536 (94%)	1.11	74 (14%) 6 5	10, 24, 42, 62	1 (0%)
All	All	1019/1072 (95%)	1.10	146 (14%) 6 5	7, 23, 41, 66	3 (0%)

The worst 5 of 146 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	150	SER	5.4
1	B	88	LEU	5.3
1	A	86	GLY	5.2
1	B	529	LEU	4.8
1	B	86	GLY	4.5

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 oligosaccharides 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.