



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 14, 2018 – 04:40 am GMT

PDB ID : 5XZR
Title : The atomic structure of SHP2 E76A mutant in complex with allosteric inhibitor 9b
Authors : Li, D.; Xie, J.; Zhu, J.; Liu, C.
Deposited on : 2017-07-13
Resolution : 2.80 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

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
Mogul : 1.7.3 (157068), CSD as539be (2018)
Xtriage (Phenix) : 1.13
EDS : trunk31020
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac : 5.8.0158
CCP4 : 7.0 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk31020

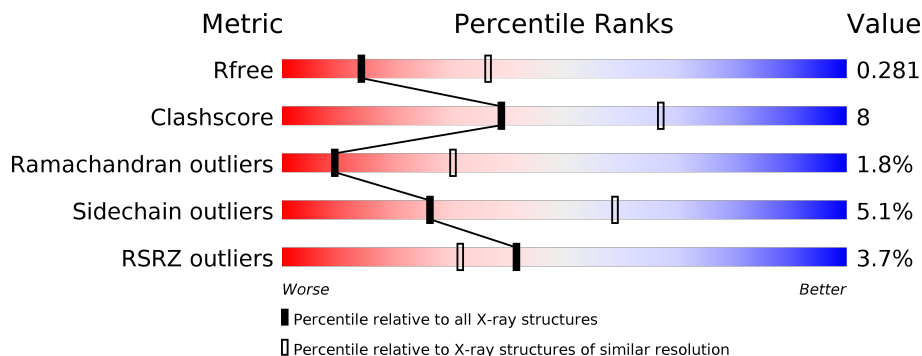
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.80 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}	111664	2792 (2.80-2.80)
Clashscore	122126	3209 (2.80-2.80)
Ramachandran outliers	120053	3158 (2.80-2.80)
Sidechain outliers	120020	3160 (2.80-2.80)
RSRZ outliers	108989	2726 (2.80-2.80)

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 ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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	557	

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 4124 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.

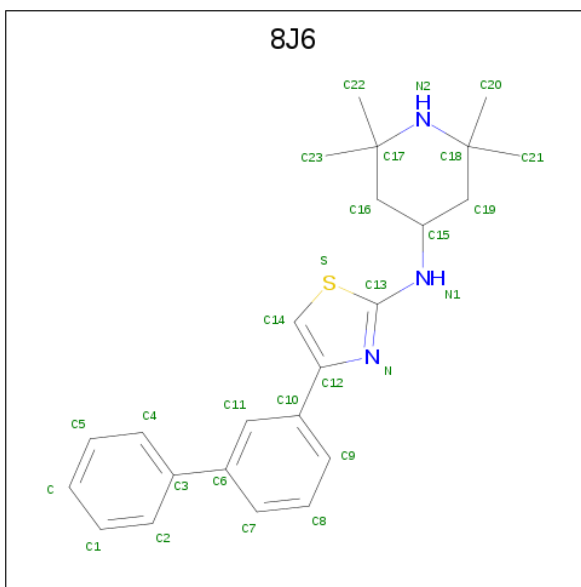
- Molecule 1 is a protein called Tyrosine-protein phosphatase non-receptor type 11.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	519	4094	2565	731	779	19	0	0	0

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-22	MET	-	expression tag	UNP Q06124
A	-21	GLY	-	expression tag	UNP Q06124
A	-20	SER	-	expression tag	UNP Q06124
A	-19	SER	-	expression tag	UNP Q06124
A	-18	HIS	-	expression tag	UNP Q06124
A	-17	HIS	-	expression tag	UNP Q06124
A	-16	HIS	-	expression tag	UNP Q06124
A	-15	HIS	-	expression tag	UNP Q06124
A	-14	HIS	-	expression tag	UNP Q06124
A	-13	HIS	-	expression tag	UNP Q06124
A	-12	SER	-	expression tag	UNP Q06124
A	-11	SER	-	expression tag	UNP Q06124
A	-10	GLY	-	expression tag	UNP Q06124
A	-9	LEU	-	expression tag	UNP Q06124
A	-8	VAL	-	expression tag	UNP Q06124
A	-7	PRO	-	expression tag	UNP Q06124
A	-6	ARG	-	expression tag	UNP Q06124
A	-5	GLY	-	expression tag	UNP Q06124
A	-4	SER	-	expression tag	UNP Q06124
A	-3	HIS	-	expression tag	UNP Q06124
A	-2	MET	-	expression tag	UNP Q06124
A	-1	ALA	-	expression tag	UNP Q06124
A	0	SER	-	expression tag	UNP Q06124
A	76	ALA	GLU	engineered mutation	UNP Q06124

- Molecule 2 is 4-(3-phenylphenyl)-N-(2,2,6,6-tetramethylpiperidin-4-yl)-1,3-thiazol-2-amine (three-letter code: 8J6) (formula: C₂₄H₂₉N₃S).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	S		
2	A	1	28	24	3	1	0	0

- Molecule 3 is water.

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

4 Data and refinement statistics

Property	Value	Source
Space group	P 2 21 21	Depositor
Cell constants a, b, c, α , β , γ	41.28Å 55.75Å 219.19Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.86 – 2.80 54.80 – 2.80	Depositor EDS
% Data completeness (in resolution range)	99.9 (19.86-2.80) 99.9 (54.80-2.80)	Depositor EDS
R_{merge}	0.17	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.29 (at 2.81Å)	Xtrriage
Refinement program	PHENIX dev_2650	Depositor
R, R_{free}	0.240 , 0.281 0.244 , 0.281	Depositor DCC
R_{free} test set	659 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	52.4	Xtrriage
Anisotropy	0.720	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 38.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.45$, $\langle L^2 \rangle = 0.28$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	4124	wwPDB-VP
Average B, all atoms (Å ²)	61.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.72% 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: 8J6

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	0.24	0/4179	0.45	2/5656 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	92	ASN	N-CA-C	7.30	130.72	111.00
1	A	91	LYS	C-N-CA	5.53	135.53	121.70

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	4094	0	3907	65	0
2	A	28	0	0	0	0
3	A	2	0	0	0	0
All	All	4124	0	3907	65	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 (65) 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:91:LYS:HA	1:A:93:GLY:H	1.32	0.92
1:A:37:ASN:HB3	1:A:40:ASP:HB2	1.67	0.77
1:A:91:LYS:HA	1:A:93:GLY:N	2.01	0.75
1:A:91:LYS:CA	1:A:93:GLY:H	2.00	0.75
1:A:330:THR:HG23	1:A:458:HIS:HB3	1.70	0.73
1:A:2:THR:HG23	1:A:259:CYS:HB2	1.75	0.68
1:A:41:PHE:O	1:A:56:ILE:N	2.20	0.67
1:A:390:GLU:HG2	1:A:399:ARG:HG2	1.79	0.63
1:A:92:ASN:OD1	1:A:92:ASN:N	2.31	0.63
1:A:306:ASN:OD1	1:A:501:ARG:NH2	2.32	0.62
1:A:479:ILE:HD13	1:A:522:ILE:HD11	1.82	0.61
1:A:497:VAL:HG12	1:A:504:MET:HG3	1.84	0.59
1:A:204:GLU:HG3	1:A:208:THR:HG23	1.85	0.58
1:A:250:GLU:OE1	1:A:511:TYR:OH	2.21	0.58
1:A:290:VAL:HG11	1:A:344:MET:HG3	1.86	0.58
1:A:4:ARG:NH1	1:A:258:GLU:OE2	2.38	0.56
1:A:340:ASP:OD1	1:A:343:ARG:NH1	2.39	0.55
1:A:117:LEU:HB3	1:A:138:ARG:HB2	1.89	0.55
1:A:330:THR:OG1	1:A:331:GLN:N	2.40	0.54
1:A:377:LEU:HD21	1:A:384:ARG:HD3	1.88	0.54
1:A:490:VAL:HB	1:A:491:PRO:HD3	1.89	0.53
1:A:47:ARG:HH21	1:A:94:ASP:HB3	1.74	0.53
1:A:69:GLU:OE1	1:A:81:TYR:OH	2.22	0.53
1:A:297:PRO:HB3	1:A:299:GLU:HA	1.92	0.52
1:A:2:THR:HB	1:A:3:SER:HA	1.91	0.52
1:A:309:ILE:HD13	1:A:328:ILE:HG12	1.92	0.52
1:A:154:GLY:HA2	1:A:165:SER:HA	1.93	0.51
1:A:222:ASN:HB3	1:A:225:GLU:HG2	1.92	0.51
1:A:233:LEU:HA	1:A:246:GLY:HA3	1.93	0.50
1:A:13:GLY:H	1:A:35:LYS:HD2	1.77	0.50
1:A:340:ASP:HA	1:A:343:ARG:HD3	1.94	0.50
1:A:4:ARG:HD3	1:A:256:GLN:HA	1.94	0.49
1:A:56:ILE:HG12	1:A:65:LEU:HD13	1.95	0.49
1:A:81:TYR:CD1	1:A:88:LEU:HB2	2.48	0.49
1:A:54:ILE:HD11	1:A:96:ILE:HD12	1.95	0.48
1:A:151:VAL:HG11	1:A:210:LEU:HB3	1.95	0.48
1:A:310:ILE:HB	1:A:327:TYR:HB2	1.97	0.47
1:A:16:ALA:HB1	1:A:30:LEU:HD11	1.98	0.46
1:A:26:ASP:HA	1:A:46:ARG:HG2	1.96	0.46
1:A:153:THR:O	1:A:166:LYS:N	2.48	0.46
1:A:34:SER:HB3	1:A:40:ASP:HB3	1.98	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:462:GLY:O	1:A:501:ARG:NH1	2.49	0.45
1:A:126:LEU:O	1:A:130:GLY:HA3	2.17	0.44
1:A:377:LEU:HD13	1:A:386:ARG:HB2	1.99	0.44
1:A:436:LEU:HD21	1:A:517:ALA:HB1	1.98	0.44
1:A:221:ILE:HD13	1:A:229:ARG:HG2	1.99	0.44
1:A:91:LYS:N	1:A:93:GLY:H	2.16	0.43
1:A:41:PHE:N	1:A:56:ILE:O	2.49	0.43
1:A:292:LEU:HB2	1:A:303:ASP:HA	2.00	0.43
1:A:508:GLU:OE1	1:A:512:ARG:NH2	2.52	0.43
1:A:69:GLU:OE1	1:A:69:GLU:N	2.51	0.43
1:A:10:ASN:ND2	1:A:106:ASP:OD2	2.49	0.43
1:A:178:LYS:HD2	1:A:186:ARG:CZ	2.49	0.42
1:A:42:THR:HA	1:A:55:LYS:HA	2.01	0.42
1:A:394:HIS:O	1:A:395:ASP:HB2	2.20	0.41
1:A:512:ARG:O	1:A:516:MET:HG2	2.20	0.41
1:A:57:GLN:HG2	1:A:64:ASP:HB3	2.02	0.41
1:A:299:GLU:HA	1:A:300:PRO:HA	1.78	0.41
1:A:280:LYS:H	1:A:280:LYS:HD2	1.84	0.41
1:A:28:SER:HA	1:A:100:TYR:O	2.20	0.41
1:A:219:THR:CG2	1:A:491:PRO:HD2	2.51	0.41
1:A:372:PRO:HG3	1:A:378:LYS:HB2	2.02	0.41
1:A:490:VAL:O	1:A:494:ILE:HG12	2.20	0.41
1:A:90:GLU:HB3	1:A:92:ASN:ND2	2.36	0.40
1:A:459:CYS:SG	1:A:460:SER:N	2.94	0.40

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	513/557 (92%)	473 (92%)	31 (6%)	9 (2%)	9 30

All (9) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	92	ASN
1	A	160	SER
1	A	213	LYS
1	A	156	ASP
1	A	302	SER
1	A	459	CYS
1	A	505	VAL
1	A	49	GLY
1	A	301	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.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	433/494 (88%)	411 (95%)	22 (5%)	26 59

All (22) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	10	ASN
1	A	19	LEU
1	A	35	LYS
1	A	74	LEU
1	A	79	GLN
1	A	92	ASN
1	A	118	SER
1	A	204	GLU
1	A	208	THR
1	A	209	VAL
1	A	214	GLN
1	A	259	CYS
1	A	265	ARG
1	A	280	LYS
1	A	292	LEU
1	A	330	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	335	GLN
1	A	366	LYS
1	A	477	ASP
1	A	525	LEU
1	A	531	GLU
1	A	534	LYS

Some 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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

1 ligand is modelled in this entry.

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	$\# Z > 2$
2	8J6	A	601	-	27,31,31	1.41	3 (11%)	38,46,46	1.82	4 (10%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the

Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	8J6	A	601	-	-	0/10/28/28	0/4/4/4

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	601	8J6	C16-C15	-2.60	1.49	1.52
2	A	601	8J6	C6-C3	2.84	1.56	1.49
2	A	601	8J6	C10-C12	4.98	1.56	1.48

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	601	8J6	C12-C14-S	-6.43	103.89	111.79
2	A	601	8J6	C17-C16-C15	-5.09	108.71	113.72
2	A	601	8J6	C14-C12-C10	-4.78	122.79	129.44
2	A	601	8J6	C21-C18-C19	-2.99	106.98	110.18

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 [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	519/557 (93%)	0.17	19 (3%) 41 31	26, 57, 116, 202	0

All (19) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	295	GLY	6.6
1	A	297	PRO	5.0
1	A	296	ASP	4.0
1	A	294	ASP	3.6
1	A	238	GLU	3.4
1	A	182	GLY	3.1
1	A	237	ALA	3.1
1	A	259	CYS	3.1
1	A	1	MET	2.8
1	A	300	PRO	2.8
1	A	298	ASN	2.7
1	A	40	ASP	2.6
1	A	160	SER	2.3
1	A	486	CYS	2.3
1	A	177	LEU	2.2
1	A	31	ALA	2.2
1	A	242	LYS	2.2
1	A	299	GLU	2.0
1	A	94	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](#)

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(Å ²)	Q<0.9
2	8J6	A	601	28/28	0.92	0.22	32,40,50,53	0

6.5 Other polymers [i](#)

There are no such residues in this entry.