WORLDWIDE
wwPDB
PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report ⓘ

Sep 19, 2023 – 09:42 PM EDT

PDB ID	:	5IBS
Title	:	Structure of E76Q, a Cancer-Associated Mutation of the Oncogenic Phosphatase SHP2
Authors	:	Blacklow, S.C.; Stams, T.; Fodor, M.; LaRochelle, J.R.
Deposited on	:	2016-02-22
Resolution	:	2.32 Å(reported)

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	:	2.35.1
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)	:	2.35.1

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 2.32 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.
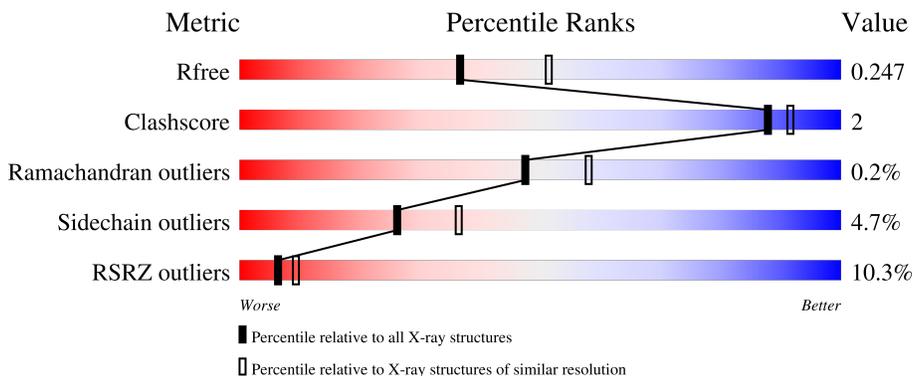
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	5974 (2.34-2.30)
Clashscore	141614	6604 (2.34-2.30)
Ramachandran outliers	138981	6523 (2.34-2.30)
Sidechain outliers	138945	6523 (2.34-2.30)
RSRZ outliers	127900	5855 (2.34-2.30)

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	526	10% / 82% / 10% / 8%
1	B	526	9% / 82% / 8% · 9%

2   Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 8215 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
1	A	485	Total 3937	C 2483	N 701	O 735	S 18	0	1	0
1	B	479	Total 3882	C 2451	N 692	O 721	S 18	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	SER	-	expression tag	UNP Q06124
A	76	GLN	GLU	engineered mutation	UNP Q06124
B	0	SER	-	expression tag	UNP Q06124
B	76	GLN	GLU	engineered mutation	UNP Q06124

- Molecule 2 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	183	Total 183	O 183	0	0
2	B	213	Total 213	O 213	0	0

3 Residue-property plots ⓘ

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 11
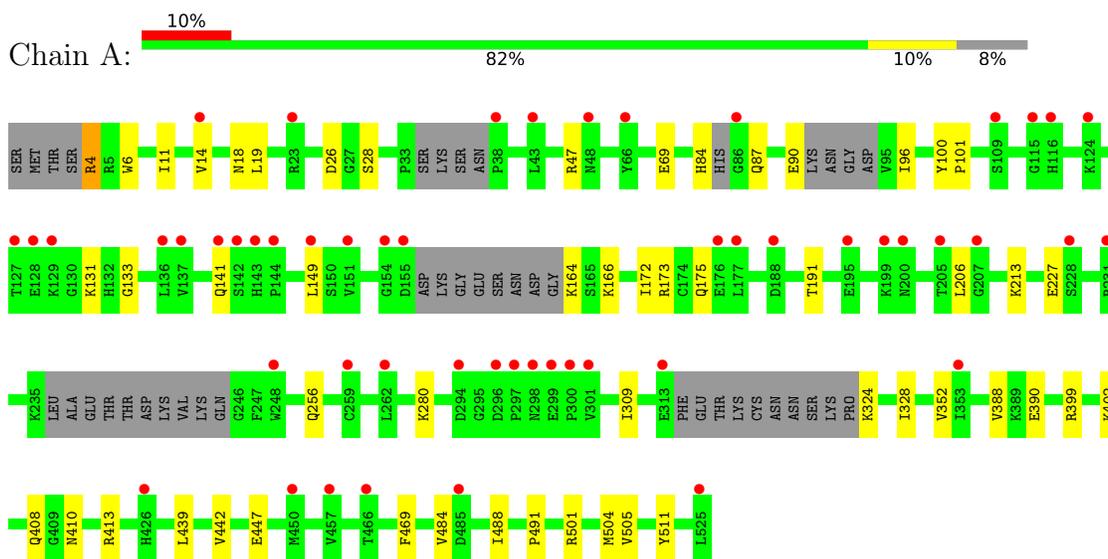
• Molecule 1: Tyrosine-protein phosphatase non-receptor type 11
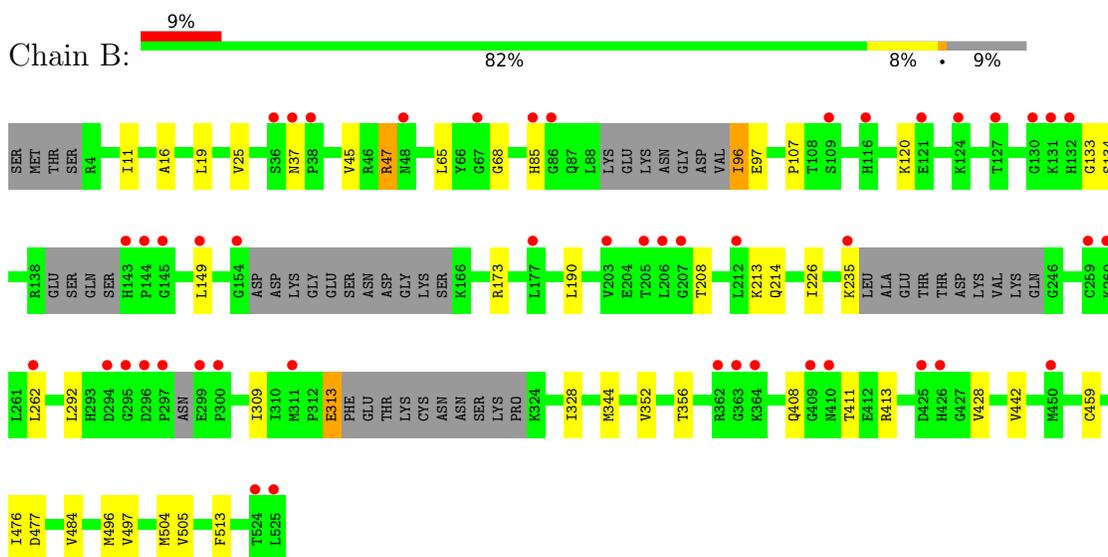
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	45.46Å 214.77Å 55.38Å 90.00° 95.59° 90.00°	Depositor
Resolution (Å)	23.39 – 2.32 23.19 – 2.32	Depositor EDS
% Data completeness (in resolution range)	98.8 (23.39-2.32) 98.8 (23.19-2.32)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.28 (at 2.33Å)	Xtriage
Refinement program	BUSTER 2.11.5	Depositor
R, R_{free}	0.186 , 0.238 0.193 , 0.247	Depositor DCC
R_{free} test set	2217 reflections (4.94%)	wwPDB-VP
Wilson B-factor (Å^2)	34.0	Xtriage
Anisotropy	0.384	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 62.1	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	0.94	EDS
Total number of atoms	8215	wwPDB-VP
Average B, all atoms (Å^2)	43.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.56% 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 ⓘ

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	0.44	0/4020	0.62	0/5419
1	B	0.45	0/3963	0.66	0/5345
All	All	0.44	0/7983	0.64	0/10764

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	3937	0	3883	16	0
1	B	3882	0	3825	16	0
2	A	183	0	0	0	0
2	B	213	0	0	0	0
All	All	8215	0	7708	31	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.

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:45:VAL:HB	1:B:96:ILE:HD11	1.80	0.61

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:84:HIS:HB3	1:A:87:GLN:HG3	1.83	0.60
1:A:388:VAL:HG21	1:A:402:LYS:HD3	1.86	0.58
1:B:133:GLY:HA3	1:B:213:LYS:HB2	1.85	0.57
1:A:175:GLN:HB3	1:B:25:VAL:HG11	1.87	0.56
1:B:313:GLU:HB2	1:B:496:MET:HG3	1.89	0.53
1:B:352:VAL:HG11	1:B:442:VAL:HG13	1.90	0.53
1:A:133:GLY:HA3	1:A:213:LYS:HB2	1.91	0.52
1:A:501:ARG:O	1:A:504:MET:HG2	2.11	0.51
1:A:309:ILE:HD13	1:A:328:ILE:HG12	1.94	0.50
1:A:390:GLU:HG2	1:A:399:ARG:HG2	1.94	0.49
1:A:69:GLU:OE1	1:A:280:LYS:HE3	2.12	0.49
1:B:65:LEU:HD12	1:B:68:GLY:HA3	1.94	0.49
1:A:90:GLU:HB3	1:A:96:ILE:HD11	1.96	0.48
1:A:28:SER:HA	1:A:100:TYR:O	2.13	0.47
1:A:491:PRO:HG3	1:A:511:TYR:OH	2.15	0.47
1:A:6:TRP:HB3	1:A:101:PRO:HB3	1.98	0.46
1:A:352:VAL:HG11	1:A:442:VAL:HG13	1.97	0.45
1:B:47:ARG:HG2	1:B:96:ILE:HD12	1.99	0.45
1:B:107:PRO:HG3	1:B:190:LEU:HD12	2.00	0.44
1:B:309:ILE:HD13	1:B:328:ILE:HG12	2.01	0.43
1:A:4:ARG:HB3	1:A:256:GLN:HE22	1.83	0.43
1:B:428:VAL:HG21	1:B:513:PHE:HB2	2.00	0.43
1:B:134:SER:HA	1:B:214:GLN:O	2.19	0.42
1:A:149:LEU:HB2	1:A:172:ILE:HD11	2.02	0.42
1:B:356:THR:OG1	1:B:459:CYS:HB3	2.19	0.42
1:A:439:LEU:HD21	1:A:469:PHE:HD2	1.85	0.41
1:B:292:LEU:HD21	1:B:344:MET:HB2	2.01	0.41
1:B:408:GLN:HB3	1:B:411:THR:HG23	2.02	0.41
1:B:497:VAL:HG12	1:B:504:MET:HG3	2.03	0.41
1:B:11:ILE:HD12	1:B:16:ALA:HB2	2.04	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	472/526 (90%)	462 (98%)	9 (2%)	1 (0%)	47	58
1	B	465/526 (88%)	446 (96%)	18 (4%)	1 (0%)	47	58
All	All	937/1052 (89%)	908 (97%)	27 (3%)	2 (0%)	47	58

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.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	432/468 (92%)	410 (95%)	22 (5%)	24	33
1	B	425/468 (91%)	407 (96%)	18 (4%)	30	41
All	All	857/936 (92%)	817 (95%)	40 (5%)	26	36

All (40) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	4	ARG
1	A	11	ILE
1	A	14	VAL
1	A	18	ASN
1	A	19	LEU
1	A	26	ASP
1	A	47	ARG
1	A	131	LYS
1	A	141	GLN
1	A	164	LYS
1	A	166	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	173	ARG
1	A	191	THR
1	A	206	LEU
1	A	227	GLU
1	A	324	LYS
1	A	408	GLN
1	A	410	ASN
1	A	413	ARG
1	A	447	GLU
1	A	484	VAL
1	A	488	ILE
1	B	19	LEU
1	B	37	ASN
1	B	47	ARG
1	B	85	HIS
1	B	96	ILE
1	B	97	GLU
1	B	120	LYS
1	B	149	LEU
1	B	173	ARG
1	B	208	THR
1	B	226	ILE
1	B	235	LYS
1	B	262	LEU
1	B	313	GLU
1	B	413	ARG
1	B	476	ILE
1	B	477	ASP
1	B	484	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	256	GLN
1	B	37	ASN
1	B	85	HIS

5.3.3   RNA (i)

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 monosaccharides 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, 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	485/526 (92%)	0.58	52 (10%)	6	8	20, 44, 71, 108	0
1	B	479/526 (91%)	0.53	47 (9%)	7	10	20, 40, 69, 88	0
All	All	964/1052 (91%)	0.55	99 (10%)	6	9	20, 42, 71, 108	0

All (99) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	363	GLY	5.4
1	A	525	LEU	5.2
1	B	85	HIS	5.0
1	A	262	LEU	4.9
1	B	426	HIS	4.8
1	A	144	PRO	4.8
1	A	48	ASN	4.7
1	B	364	LYS	4.5
1	A	177	LEU	4.3
1	B	36	SER	4.3
1	A	297	PRO	4.2
1	B	294	ASP	4.2
1	A	188	ASP	4.1
1	B	144	PRO	4.1
1	B	177	LEU	4.1
1	B	262	LEU	4.0
1	A	294	ASP	3.7
1	A	149	LEU	3.7
1	B	362	ARG	3.6
1	B	425	ASP	3.5
1	A	115	GLY	3.4
1	A	142	SER	3.4
1	A	154	GLY	3.3
1	A	14	VAL	3.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	300	PRO	3.3
1	B	154	GLY	3.3
1	A	136	LEU	3.3
1	B	297	PRO	3.3
1	B	410	ASN	3.3
1	B	203	VAL	3.3
1	B	205	THR	3.3
1	B	207	GLY	3.3
1	B	450	MET	3.3
1	A	298	ASN	3.2
1	B	86	GLY	3.2
1	B	132	HIS	3.2
1	B	116	HIS	3.1
1	A	109	SER	3.1
1	A	127	THR	3.1
1	B	206	LEU	3.0
1	A	313	GLU	3.0
1	A	23	ARG	3.0
1	A	38	PRO	2.9
1	A	116	HIS	2.9
1	B	299	GLU	2.9
1	B	145	GLY	2.9
1	A	141	GLN	2.9
1	A	151	VAL	2.8
1	A	299	GLU	2.8
1	B	109	SER	2.8
1	B	130	GLY	2.8
1	B	67	GLY	2.7
1	B	48	ASN	2.7
1	A	176	GLU	2.7
1	A	199	LYS	2.7
1	B	143	HIS	2.7
1	A	300	PRO	2.6
1	A	143	HIS	2.6
1	A	86	GLY	2.6
1	A	296	ASP	2.6
1	A	301	VAL	2.6
1	B	121	GLU	2.6
1	B	260	LYS	2.5
1	A	155	ASP	2.5
1	B	409	GLY	2.5
1	B	127	THR	2.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	124	LYS	2.5
1	A	259	CYS	2.5
1	A	353	ILE	2.5
1	A	248	TRP	2.5
1	A	426	HIS	2.5
1	B	38	PRO	2.4
1	B	37	ASN	2.4
1	B	296	ASP	2.4
1	A	231	ARG	2.4
1	A	205	THR	2.4
1	B	259	CYS	2.4
1	A	195	GLU	2.4
1	A	128	GLU	2.3
1	A	450	MET	2.3
1	A	200	ASN	2.3
1	B	311	MET	2.3
1	B	131	LYS	2.3
1	A	66	TYR	2.2
1	A	228	SER	2.2
1	A	124	LYS	2.2
1	B	295	GLY	2.2
1	A	466	THR	2.2
1	A	207	GLY	2.2
1	B	212	LEU	2.1
1	B	524	THR	2.1
1	B	525	LEU	2.1
1	B	235	LYS	2.1
1	A	137	VAL	2.1
1	A	457	VAL	2.1
1	A	485	ASP	2.0
1	A	43	LEU	2.0
1	A	129	LYS	2.0
1	B	149	LEU	2.0

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.