



Full wwPDB X-ray Structure Validation Report ⓘ

May 16, 2020 – 11:14 am BST

PDB ID : 4BJ5
Title : Crystal structure of Rif2 in complex with the C-terminal domain of Rap1 (Rap1-RCT)
Authors : Shi, T.; Bunker, R.D.; Gut, H.; Scrima, A.; Thoma, N.H.
Deposited on : 2013-04-16
Resolution : 3.29 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 : 1.8.5 (274361), CSD as541be (2020)
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

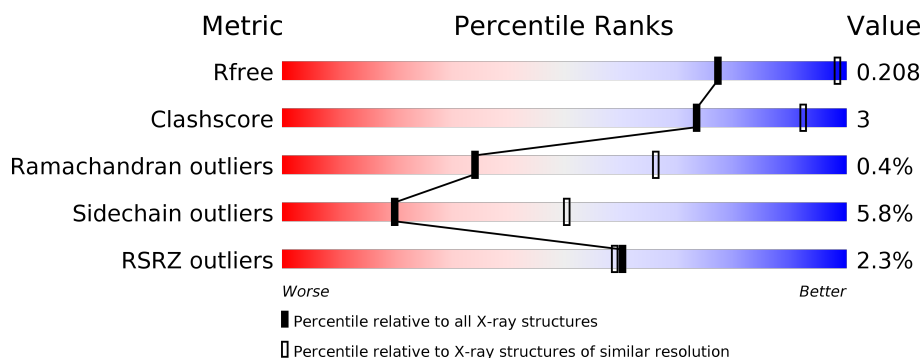
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.29 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	1149 (3.34-3.26)
Clashscore	141614	1205 (3.34-3.26)
Ramachandran outliers	138981	1183 (3.34-3.26)
Sidechain outliers	138945	1182 (3.34-3.26)
RSRZ outliers	127900	1115 (3.34-3.26)

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 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	399	<div> <div></div> <div> <div></div> <div>62%</div> <div>11%</div> <div>•</div> <div>27%</div> </div> </div>
1	B	399	<div> <div></div> <div> <div></div> <div>62%</div> <div>11%</div> <div>•</div> <div>27%</div> </div> </div>
2	C	202	<div> <div></div> <div> <div></div> <div>66%</div> <div>8%</div> <div>25%</div> </div> </div>
2	E	202	<div> <div></div> <div> <div></div> <div>66%</div> <div>8%</div> <div>25%</div> </div> </div>
3	D	13	<div> <div>15%</div> <div> <div></div> <div>85%</div> <div>15%</div> </div> </div>
3	F	13	<div> <div>15%</div> <div> <div></div> <div>85%</div> <div>15%</div> </div> </div>

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 7305 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 PROTEIN RIF2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	292	Total	C	N	O	S	0	0	0
			2347	1520	374	440	13			
1	B	291	Total	C	N	O	S	0	0	0
			2308	1495	364	436	13			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-3	GLY	-	expression tag	UNP Q06208
A	-2	GLY	-	expression tag	UNP Q06208
A	-1	GLY	-	expression tag	UNP Q06208
A	0	ARG	-	expression tag	UNP Q06208
B	-3	GLY	-	expression tag	UNP Q06208
B	-2	GLY	-	expression tag	UNP Q06208
B	-1	GLY	-	expression tag	UNP Q06208
B	0	ARG	-	expression tag	UNP Q06208

- Molecule 2 is a protein called DNA-BINDING PROTEIN RAP1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	151	Total	C	N	O	S	0	0	0
			1226	777	203	239	7			
2	E	151	Total	C	N	O	S	0	0	0
			1227	778	204	239	6			

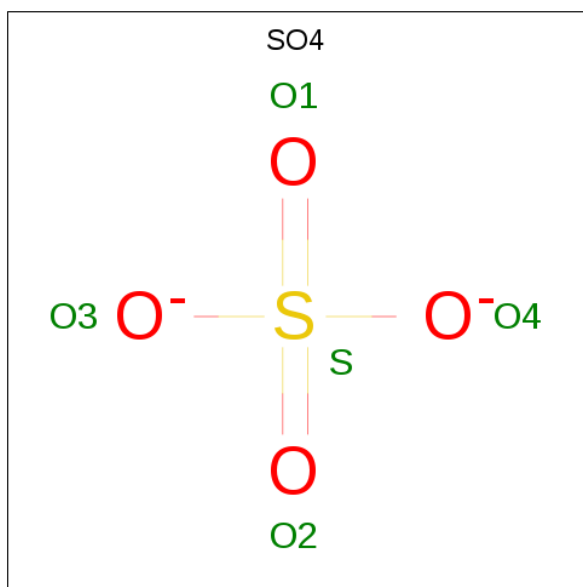
There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	626	GLY	-	expression tag	UNP P11938
E	626	GLY	-	expression tag	UNP P11938

- Molecule 3 is a protein called DNA-BINDING PROTEIN RAP1.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	D	13	Total	C	N	O	0	0	0
			84	56	14	14			
3	F	13	Total	C	N	O	0	0	0
			88	59	15	14			

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O₄S).

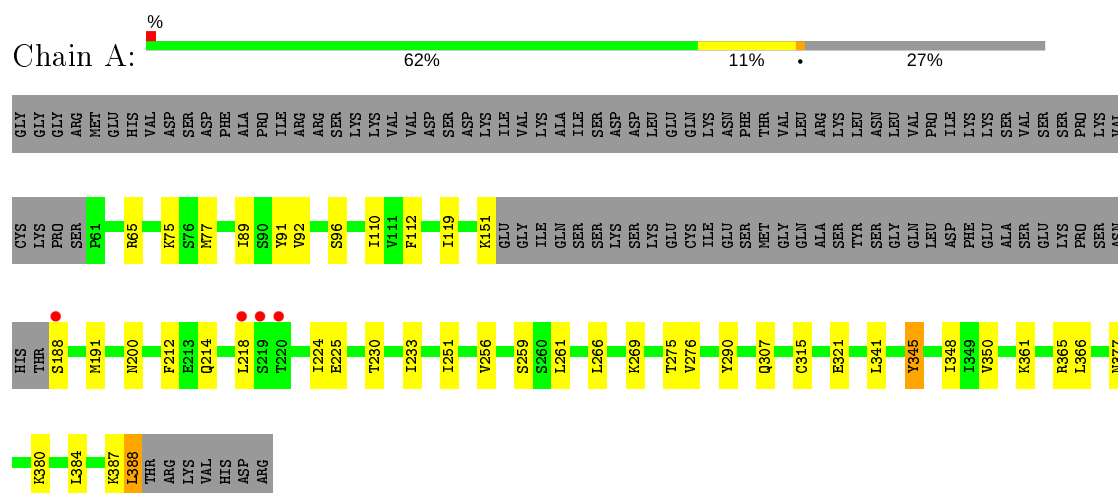


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	O	S	0	0
			5	4	1		
4	A	1	Total	O	S	0	0
			5	4	1		
4	B	1	Total	O	S	0	0
			5	4	1		
4	C	1	Total	O	S	0	0
			5	4	1		
4	E	1	Total	O	S	0	0
			5	4	1		

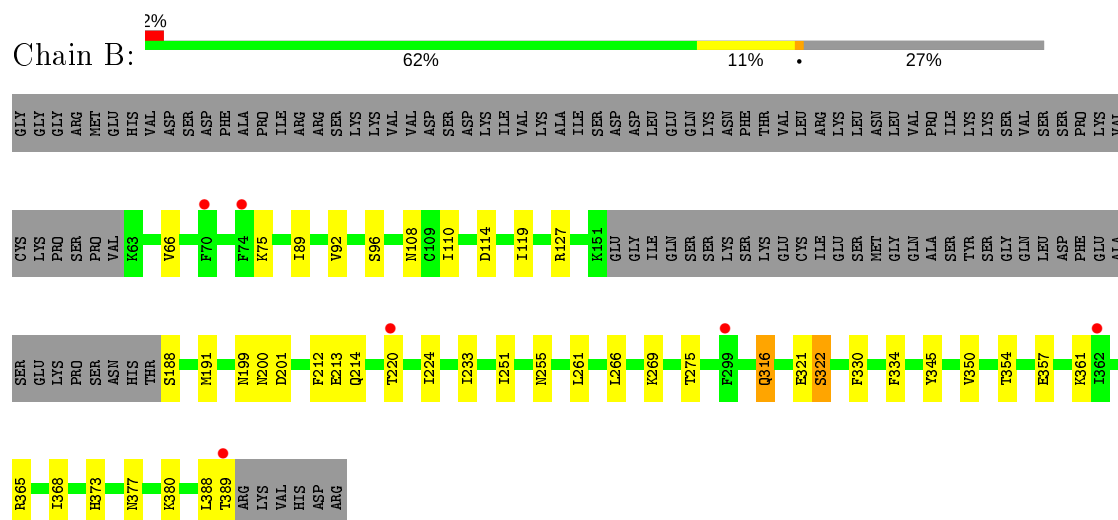
3 Residue-property plots [i](#)

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: PROTEIN RIF2

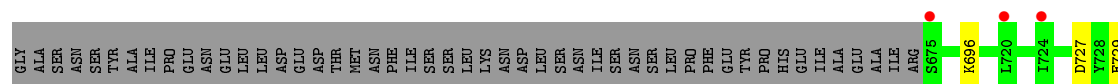


• Molecule 1: PROTEIN RIF2

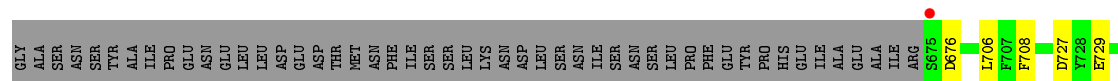


• Molecule 2: DNA-BINDING PROTEIN RAP1

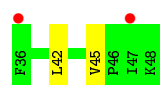
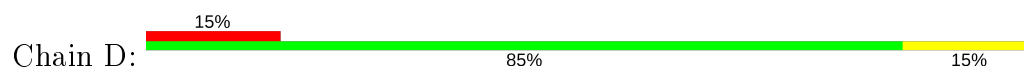




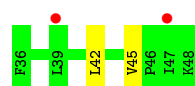
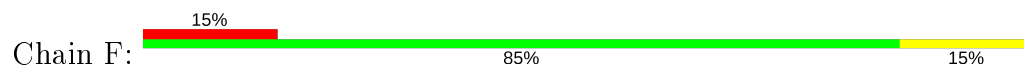
● Molecule 2: DNA-BINDING PROTEIN RAP1



● Molecule 3: DNA-BINDING PROTEIN RAP1



● Molecule 3: DNA-BINDING PROTEIN RAP1



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	108.44Å 113.86Å 140.95Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	70.47 – 3.29 68.60 – 3.29	Depositor EDS
% Data completeness (in resolution range)	99.5 (70.47-3.29) 99.5 (68.60-3.29)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.84 (at 3.26Å)	Xtriage
Refinement program	BUSTER 2.11.4	Depositor
R, R_{free}	0.176 , 0.193 0.191 , 0.208	Depositor DCC
R_{free} test set	1349 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	104.1	Xtriage
Anisotropy	0.380	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 90.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	0.013 for k,h,-l	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	7305	wwPDB-VP
Average B, all atoms (Å ²)	115.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.81% 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

Bond lengths and bond angles in the following residue types are not validated in this section: SO4

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	0.54	0/2393	0.72	0/3240
1	B	0.52	0/2353	0.70	0/3193
2	C	0.46	0/1257	0.59	0/1700
2	E	0.45	0/1258	0.61	0/1701
3	D	0.60	0/84	0.67	0/116
3	F	0.58	0/88	0.71	0/120
All	All	0.50	0/7433	0.67	0/10070

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	2347	0	2325	21	0
1	B	2308	0	2252	17	0
2	C	1226	0	1162	6	0
2	E	1227	0	1166	5	0
3	D	84	0	77	1	0
3	F	88	0	88	1	0
4	A	10	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	5	0	0	0	0
4	C	5	0	0	0	0
4	E	5	0	0	0	0
All	All	7305	0	7070	47	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 3.

All (47) 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:256:VAL:HA	1:A:259:SER:HB3	1.78	0.65
2:C:761:ASP:HA	3:D:45:VAL:HB	1.78	0.64
1:A:224:ILE:HD11	1:A:261:LEU:HD21	1.83	0.60
2:E:761:ASP:HA	3:F:45:VAL:HB	1.82	0.60
1:A:233:ILE:HG22	1:A:269:LYS:HG2	1.84	0.58
2:E:736:LEU:HD23	2:E:755:LEU:HD11	1.86	0.58
1:B:233:ILE:HG22	1:B:269:LYS:HG2	1.85	0.58
1:B:119:ILE:HG12	1:B:380:LYS:HA	1.87	0.56
1:B:75:LYS:HE3	1:B:350:VAL:HA	1.88	0.56
2:C:696:LYS:HB2	2:C:776:ASN:HD21	1.69	0.56
2:C:764:VAL:HG23	2:C:767:ARG:HD2	1.89	0.55
2:C:822:PHE:HA	2:C:825:ASP:HB2	1.88	0.54
1:B:354:THR:HG23	1:B:357:GLU:H	1.72	0.54
1:A:151:LYS:HG3	1:A:225:GLU:HG2	1.92	0.51
1:A:112:PHE:HB2	1:A:276:VAL:HG22	1.92	0.51
1:B:224:ILE:HD11	1:B:261:LEU:HD21	1.91	0.51
1:A:119:ILE:HG12	1:A:380:LYS:HA	1.93	0.50
1:A:75:LYS:HE3	1:A:350:VAL:HA	1.95	0.49
1:B:361:LYS:HA	1:B:365:ARG:HB2	1.94	0.49
2:C:758:LEU:HD13	2:C:764:VAL:HG13	1.95	0.49
1:B:334:PHE:HE2	1:B:368:ILE:HD12	1.78	0.49
1:A:387:LYS:O	1:A:388:LEU:HG	2.15	0.47
1:B:212:PHE:HB3	1:B:251:ILE:HD13	1.97	0.46
1:A:224:ILE:HD11	1:A:261:LEU:CD2	2.46	0.46
1:A:212:PHE:HB3	1:A:251:ILE:HD13	1.98	0.46
1:A:361:LYS:HA	1:A:365:ARG:HB2	1.98	0.46
1:B:224:ILE:HD11	1:B:261:LEU:CD2	2.46	0.45
1:B:330:PHE:HB3	1:B:334:PHE:CE2	2.51	0.45
1:A:151:LYS:CG	1:A:225:GLU:HG2	2.47	0.45
1:A:348:ILE:HD12	1:A:366:LEU:HD13	1.98	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:350:VAL:HG11	2:E:708:PHE:CG	2.52	0.44
1:B:220:THR:HA	1:B:255:ASN:HD21	1.83	0.43
1:A:269:LYS:HG3	1:A:269:LYS:O	2.17	0.43
1:A:91:TYR:HB3	1:A:290:TYR:CE2	2.53	0.43
2:E:737:VAL:HG22	2:E:751:SER:HB2	2.01	0.42
1:B:269:LYS:HG3	1:B:269:LYS:O	2.19	0.42
1:B:110:ILE:HG21	1:B:266:LEU:HD13	2.00	0.42
1:B:89:ILE:HG22	1:B:92:VAL:HG23	2.02	0.42
1:B:188:SER:HB2	1:B:191:MET:HB3	2.02	0.42
1:A:110:ILE:HG21	1:A:266:LEU:HD13	2.02	0.41
1:B:199:ASN:HD21	1:B:201:ASP:HB2	1.85	0.41
1:B:316:GLN:HE21	1:B:322:SER:HB2	1.85	0.41
1:A:384:LEU:HD23	2:E:706:LEU:HD12	2.01	0.41
1:A:188:SER:HB2	1:A:191:MET:HB3	2.02	0.41
1:A:89:ILE:HG22	1:A:92:VAL:HG23	2.02	0.41
2:C:737:VAL:HG22	2:C:751:SER:HB2	2.03	0.41
1:A:341:LEU:O	1:A:345:TYR:HB2	2.20	0.41

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	288/399 (72%)	277 (96%)	11 (4%)	0	100	100
1	B	287/399 (72%)	275 (96%)	12 (4%)	0	100	100
2	C	149/202 (74%)	143 (96%)	4 (3%)	2 (1%)	12	40
2	E	149/202 (74%)	142 (95%)	5 (3%)	2 (1%)	12	40
3	D	11/13 (85%)	11 (100%)	0	0	100	100
3	F	11/13 (85%)	11 (100%)	0	0	100	100
All	All	895/1228 (73%)	859 (96%)	32 (4%)	4 (0%)	34	66

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	E	732	GLN
2	C	732	GLN
2	C	781	PRO
2	E	781	PRO

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	268/378 (71%)	254 (95%)	14 (5%)	23	54
1	B	260/378 (69%)	243 (94%)	17 (6%)	17	46
2	C	140/188 (74%)	133 (95%)	7 (5%)	24	55
2	E	140/188 (74%)	132 (94%)	8 (6%)	20	51
3	D	7/13 (54%)	6 (86%)	1 (14%)	3	15
3	F	8/13 (62%)	7 (88%)	1 (12%)	4	19
All	All	823/1158 (71%)	775 (94%)	48 (6%)	20	50

All (48) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	65	ARG
1	A	77	MET
1	A	96	SER
1	A	200	ASN
1	A	214	GLN
1	A	218	LEU
1	A	230	THR
1	A	275	THR
1	A	307	GLN
1	A	315	CYS
1	A	321	GLU
1	A	345	TYR
1	A	377	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	388	LEU
1	B	66	VAL
1	B	96	SER
1	B	108	ASN
1	B	114	ASP
1	B	127	ARG
1	B	200	ASN
1	B	213	GLU
1	B	214	GLN
1	B	275	THR
1	B	316	GLN
1	B	321	GLU
1	B	322	SER
1	B	345	TYR
1	B	373	HIS
1	B	377	ASN
1	B	388	LEU
1	B	389	THR
2	C	727	ASP
2	C	729	GLU
2	C	757	CYS
2	C	765	PHE
2	C	782	ASN
2	C	795	LEU
2	C	825	ASP
3	D	42	LEU
2	E	676	ASP
2	E	727	ASP
2	E	729	GLU
2	E	741	CYS
2	E	757	CYS
2	E	765	PHE
2	E	782	ASN
2	E	795	LEU
3	F	42	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (9) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	199	ASN
1	A	244	ASN
1	B	199	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	244	ASN
1	B	316	GLN
1	B	370	ASN
1	B	377	ASN
2	C	776	ASN
2	E	709	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 carbohydrates in this entry.

5.6 Ligand geometry ⓘ

5 ligands are 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	$\# Z > 2$
4	SO4	C	1826	-	4,4,4	0.13	0	6,6,6	0.25	0
4	SO4	B	1390	-	4,4,4	0.20	0	6,6,6	0.29	0
4	SO4	E	1826	-	4,4,4	0.17	0	6,6,6	0.48	0
4	SO4	A	1389	-	4,4,4	0.42	0	6,6,6	0.22	0
4	SO4	A	1390	-	4,4,4	0.13	0	6,6,6	0.25	0

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

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	292/399 (73%)	0.26	4 (1%) 75 75	68, 99, 150, 193	0
1	B	291/399 (72%)	0.29	6 (2%) 63 62	77, 110, 186, 201	0
2	C	151/202 (74%)	0.33	5 (3%) 46 44	92, 123, 183, 202	0
2	E	151/202 (74%)	0.30	2 (1%) 77 77	76, 110, 170, 186	0
3	D	13/13 (100%)	0.68	2 (15%) 2 2	143, 160, 176, 177	0
3	F	13/13 (100%)	0.80	2 (15%) 2 2	126, 146, 171, 172	0
All	All	911/1228 (74%)	0.30	21 (2%) 60 59	68, 110, 174, 202	0

All (21) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	E	675	SER	4.4
1	B	299	PHE	3.7
1	B	389	THR	3.7
1	B	220	THR	3.5
3	F	39	LEU	3.4
1	B	74	PHE	2.9
3	D	36	PHE	2.8
1	A	188	SER	2.7
1	B	362	ILE	2.7
1	A	218	LEU	2.5
1	A	219	SER	2.4
2	C	675	SER	2.3
1	B	70	PHE	2.3
3	F	47	ILE	2.3
2	E	821	PHE	2.2
3	D	47	ILE	2.2
2	C	724	ILE	2.2
2	C	720	LEU	2.1
2	C	780	PRO	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	C	765	PHE	2.0
1	A	220	THR	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(\AA^2)	Q<0.9
4	SO4	E	1826	5/5	0.88	0.25	131,132,134,134	0
4	SO4	C	1826	5/5	0.89	0.15	172,173,173,173	0
4	SO4	A	1390	5/5	0.96	0.23	128,128,131,132	0
4	SO4	B	1390	5/5	0.99	0.22	78,81,83,86	0
4	SO4	A	1389	5/5	1.00	0.26	79,81,81,82	0

6.5 Other polymers [i](#)

There are no such residues in this entry.