



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 14, 2017 – 10:26 pm GMT

PDB ID : 2PJR
Title : HELICASE PRODUCT COMPLEX
Authors : Velankar, S.S.; Soultanas, P.; Dillingham, M.S.; Subramanya, H.S.; Wigley, D.B.
Deposited on : 1999-03-12
Resolution : 2.90 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

<http://wwpdb.org/validation/2016/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:

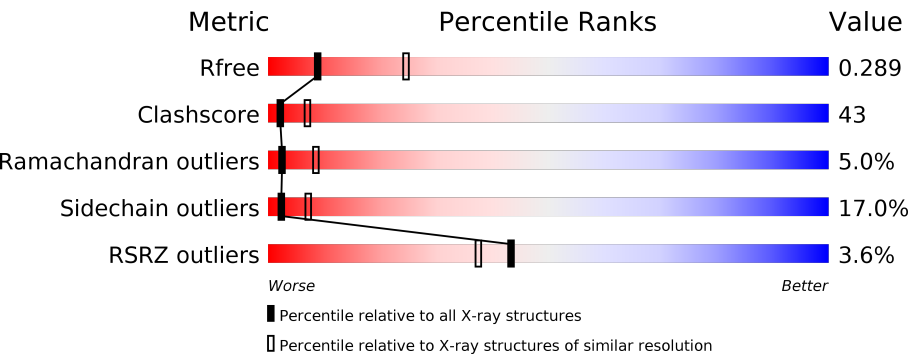
MolProbity	:	4.02b-467
Mogul	:	1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix)	:	1.9-1692
EDS	:	trunk28620
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	recalc28949

1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.90 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}	100719	1586 (2.90-2.90)
Clashscore	112137	1807 (2.90-2.90)
Ramachandran outliers	110173	1768 (2.90-2.90)
Sidechain outliers	110143	1770 (2.90-2.90)
RSRZ outliers	101464	1596 (2.90-2.90)

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	C	5	<div><div></div><div>100%</div></div>
1	D	5	<div><div></div><div>100%</div></div>
2	H	2	<div><div></div><div>100%</div><div><div></div><div>50%</div><div>50%</div></div></div>
3	I	5	<div><div></div><div>20%</div><div>60%</div><div>20%</div></div>
4	A	548	<div><div></div><div>2%</div><div><div></div><div>40%</div><div>50%</div><div>7%</div><div>..</div></div></div>
4	F	548	<div><div></div><div>4%</div><div><div></div><div>39%</div><div>45%</div><div>14%</div><div>..</div></div></div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	B	95	<div><div></div><div>7%</div><div>31%</div><div>47%</div><div>19%</div><div></div></div>
5	G	95	<div><div></div><div>5%</div><div>31%</div><div>52%</div><div>16%</div><div></div></div>

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 10671 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 DNA chain called DNA (5'-D(*TP*TP*TP*TP*T)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	C	5	Total	C	N	O	P	0	0	0
			97	50	10	33	4			
1	D	5	Total	C	N	O	P	0	0	0
			97	50	10	33	4			

- Molecule 2 is a DNA chain called DNA (5'-D(*GP*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	2	Total	C	N	O	P	0	0	0
			38	19	8	10	1			

- Molecule 3 is a DNA chain called DNA (5'-D(*AP*CP*TP*GP*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	I	5	Total	C	N	O	P	0	0	0
			98	48	18	28	4			

- Molecule 4 is a protein called PROTEIN (HELICASE PCRA).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	A	542	Total	C	N	O	S	0	0	0
			4409	2792	775	829	13			
4	F	544	Total	C	N	O	S	0	0	0
			4424	2802	777	832	13			

- Molecule 5 is a protein called PROTEIN (HELICASE PCRA).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	B	95	Total	C	N	O	S	0	0	0
			749	471	125	147	6			
5	G	95	Total	C	N	O	S	0	0	0
			749	471	125	147	6			

- Molecule 6 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	F	1	Total	O	S	0	0
			5	4	1		
6	A	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: DNA (5'-D(*TP*TP*TP*TP*T)-3')

Chain C:  100%

T13
T14
T15
T16
T17

- Molecule 1: DNA (5'-D(*TP*TP*TP*TP*T)-3')

Chain D:  100%

T13
T14
T15
T16
T17

- Molecule 2: DNA (5'-D(*GP*C)-3')

Chain H:  100%
50% 50%


G1
G2

- Molecule 3: DNA (5'-D(*AP*CP*TP*GP*C)-3')

Chain I:  20% 60% 20%

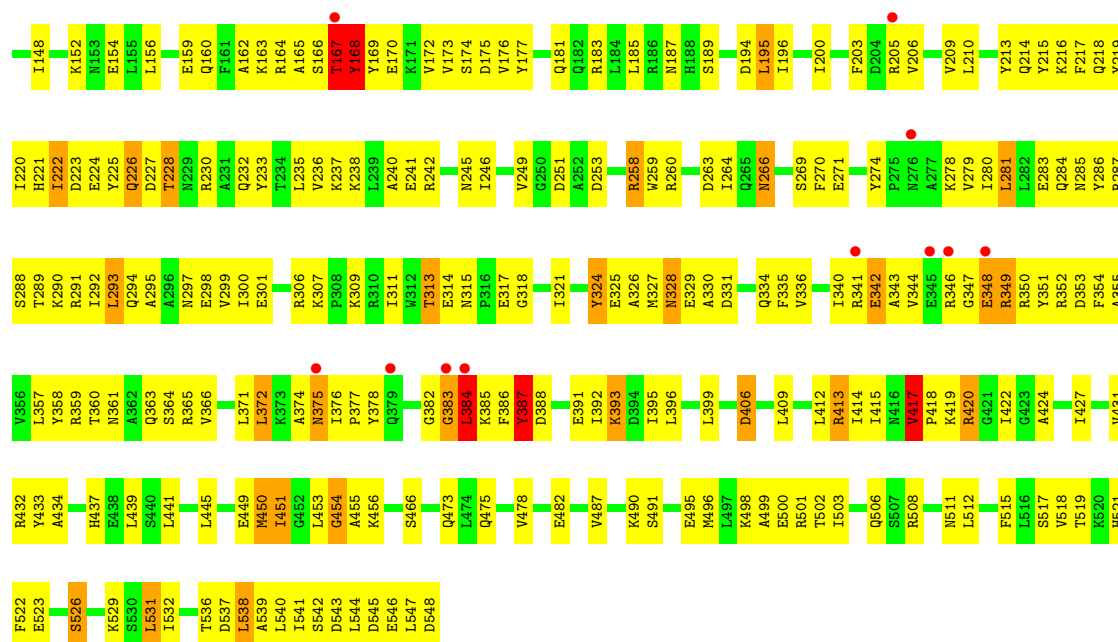
A30
C31
T32
G33
C34

- Molecule 4: PROTEIN (HELICASE PCRA)

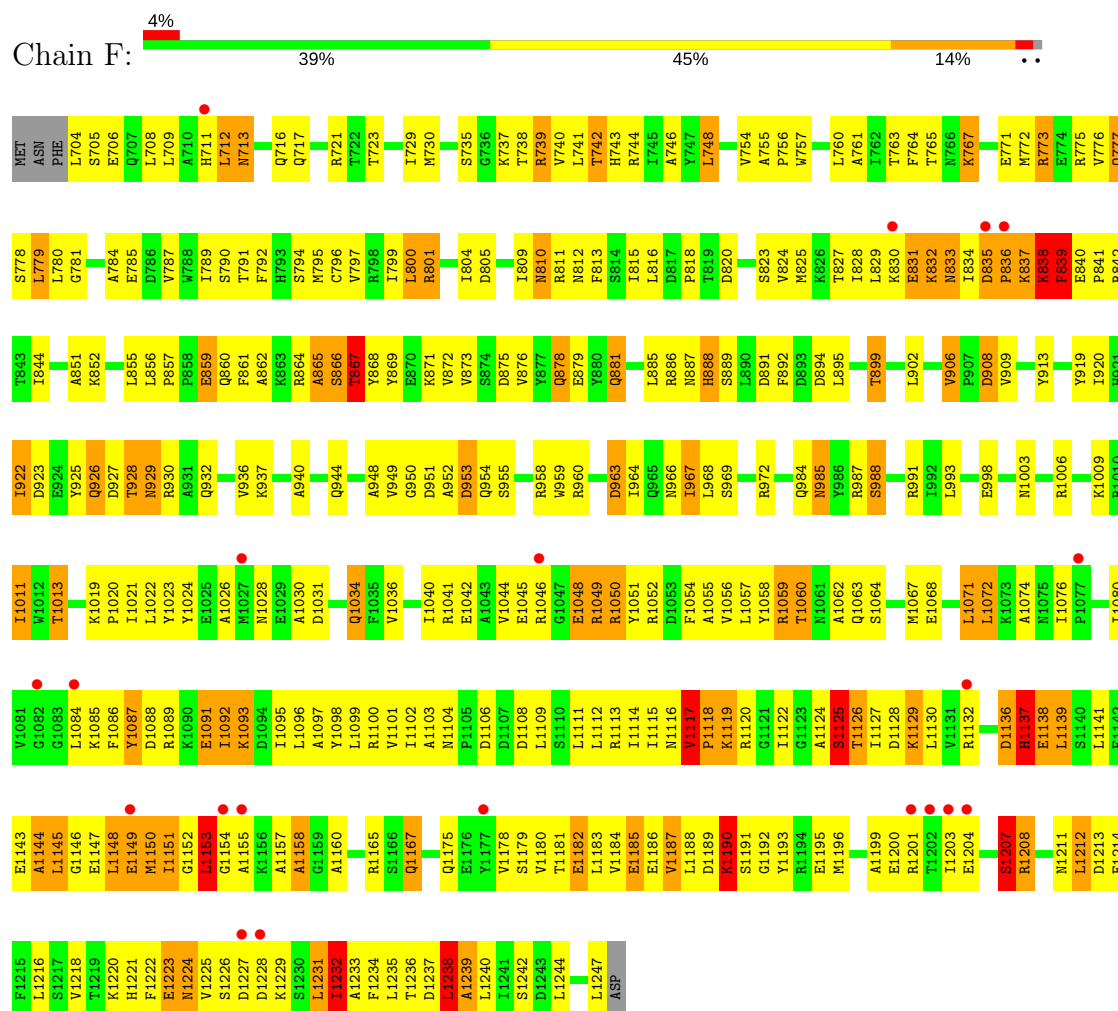
Chain A:  2% 40% 50% 7% ..

MET ASN PHE LEU SER GLU Q7 L8 L9 A10 H11 H12 L13 N14 K15 E16 Q16 Q17 E18 A19 V20 V21 R21 T22 T23 E24 L28 L29 M30 A31 G36 K37 T38 R39 V40 L41 T42 H43 R44 L48 L49 M49 A50 E51 A55 P56 W57 N58 I59 L60 T63 F64 K67 R70 E71 M72

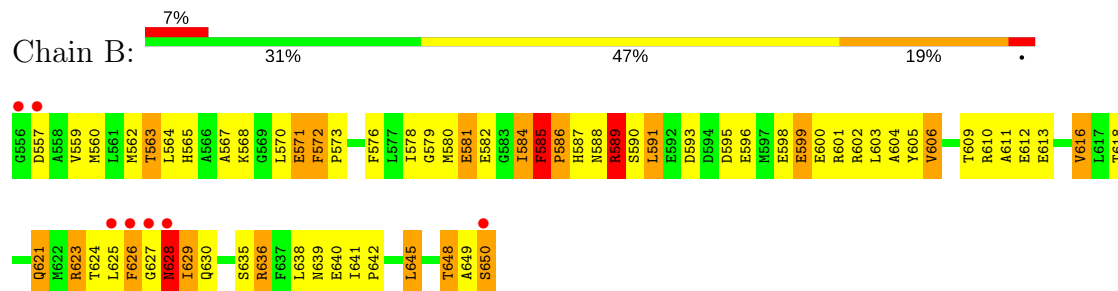
R73 V76 V77 L80 G81 D86 W87 W88 T89 S90 F91 F92 H93 S94 N95 C96 V97 L100 R101 R102 D103 I104 T107 N110 R111 N112 L116 D117 P118 T119 D120 S123 V124 M125 K126 T127 T128 L129 K130 E131 K132 N133 T134 D135 P136 K137 K138 F139 E140 T143 I144



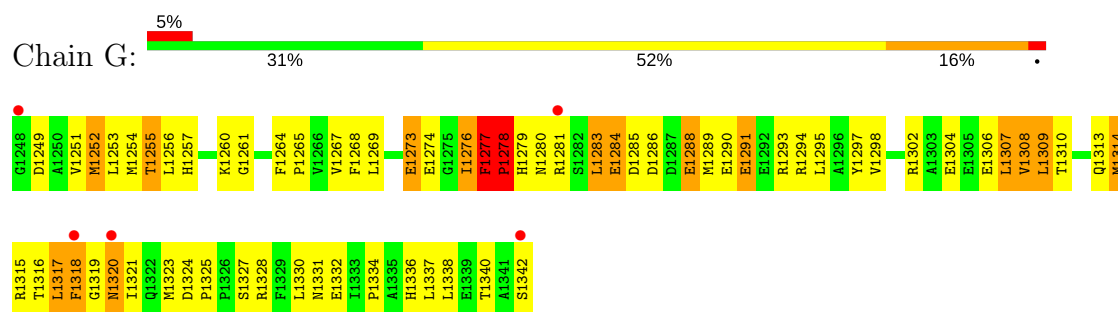
• Molecule 4: PROTEIN (HELICASE PCRA)



- Molecule 5: PROTEIN (HELICASE PCRA)



- Molecule 5: PROTEIN (HELICASE PCRA)



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	85.05Å 62.60Å 141.83Å 90.00° 95.84° 90.00°	Depositor
Resolution (Å)	15.00 – 2.90 14.98 – 2.90	Depositor EDS
% Data completeness (in resolution range)	97.9 (15.00-2.90) 97.9 (14.98-2.90)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.59 (at 2.91Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.240 , 0.296 0.240 , 0.289	Depositor DCC
R_{free} test set	1644 reflections (5.07%)	DCC
Wilson B-factor (Å ²)	45.9	Xtriage
Anisotropy	0.295	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 67.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.89	EDS
Total number of atoms	10671	wwPDB-VP
Average B, all atoms (Å ²)	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.57% 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	C	0.62	0/106	0.72	0/162
1	D	0.67	0/106	1.13	1/162 (0.6%)
2	H	1.79	2/42 (4.8%)	1.56	1/63 (1.6%)
3	I	2.09	4/109 (3.7%)	1.79	2/166 (1.2%)
4	A	0.52	0/4485	0.77	3/6059 (0.0%)
4	F	0.49	0/4500	0.78	8/6079 (0.1%)
5	B	0.63	1/762 (0.1%)	0.83	1/1028 (0.1%)
5	G	0.56	0/762	0.82	1/1028 (0.1%)
All	All	0.57	7/10872 (0.1%)	0.81	17/14747 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
3	I	0	1
4	A	0	1
4	F	0	1
All	All	0	3

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	I	32	DT	O3'-P	-10.64	1.48	1.61
5	B	650	SER	CA-CB	7.21	1.63	1.52
3	I	33	DG	O3'-P	5.79	1.68	1.61
3	I	34	DC	C4'-C3'	5.35	1.58	1.53
3	I	34	DC	C3'-O3'	5.32	1.50	1.44

The worst 5 of 17 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	17	DT	O4'-C4'-C3'	-7.29	101.58	104.50
3	I	30	DA	O5'-C5'-C4'	7.16	128.89	111.00
3	I	33	DG	OP2-P-O3'	6.65	119.84	105.20
4	F	1117	VAL	C-N-CD	6.57	142.21	128.40
4	A	417	VAL	C-N-CD	6.39	141.82	128.40

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
4	A	387	TYR	Sidechain
4	F	1193	TYR	Sidechain
3	I	34	DC	Sidechain

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	C	97	0	62	10	0
1	D	97	0	62	9	0
2	H	38	0	24	0	0
3	I	98	0	58	11	0
4	A	4409	0	4432	355	0
4	F	4424	0	4450	447	0
5	B	749	0	731	104	0
5	G	749	0	731	89	0
6	A	5	0	0	1	0
6	F	5	0	0	0	0
All	All	10671	0	10550	919	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 43.

The worst 5 of 919 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)
4:F:1096:LEU:HD21	4:F:1238:LEU:HD23	1.23	1.15

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
5:G:1316:THR:HG23	5:G:1320:ASN:HA	1.24	1.14
4:A:326:ALA:O	5:B:621:GLN:HG2	1.55	1.06
4:A:327:MET:HA	5:B:621:GLN:HG3	1.33	1.05
4:A:396:LEU:HD11	4:A:538:LEU:HD23	1.38	1.04

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	
4	A	540/548 (98%)	456 (84%)	60 (11%)	24 (4%)	3	11
4	F	542/548 (99%)	448 (83%)	63 (12%)	31 (6%)	2	6
5	B	93/95 (98%)	80 (86%)	8 (9%)	5 (5%)	2	7
5	G	93/95 (98%)	81 (87%)	8 (9%)	4 (4%)	3	12
All	All	1268/1286 (99%)	1065 (84%)	139 (11%)	64 (5%)	2	8

5 of 64 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	A	134	ILE
4	A	139	PHE
4	A	167	THR
4	A	168	TYR
4	A	348	GLU

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	
4	A	472/478 (99%)	406 (86%)	66 (14%)	4	12
4	F	474/478 (99%)	390 (82%)	84 (18%)	2	6
5	B	81/81 (100%)	62 (76%)	19 (24%)	1	2
5	G	81/81 (100%)	62 (76%)	19 (24%)	1	2
All	All	1108/1118 (99%)	920 (83%)	188 (17%)	2	7

5 of 188 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	F	705	SER
4	F	839	PHE
5	G	1277	PHE
4	F	713	ASN
4	F	778	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 54 such sidechains are listed below:

Mol	Chain	Res	Type
4	F	878	GLN
4	F	926	GLN
5	G	1257	HIS
4	F	882	GLN
4	F	888	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

2 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	$\# Z > 2$
6	SO4	A	901	-	4,4,4	0.55	0	6,6,6	0.48	0
6	SO4	F	1249	-	4,4,4	0.48	0	6,6,6	0.75	0

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
6	SO4	A	901	-	-	0/0/0/0	0/0/0/0
6	SO4	F	1249	-	-	0/0/0/0	0/0/0/0

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.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	A	901	SO4	1	0

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 [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	C	5/5 (100%)	-0.51	0 100 100	22, 25, 37, 37	0
1	D	5/5 (100%)	-0.33	0 100 100	29, 32, 41, 55	0
2	H	2/2 (100%)	2.96	2 (100%) 0 0	53, 53, 53, 64	0
3	I	5/5 (100%)	0.64	0 100 100	39, 39, 43, 47	0
4	A	542/548 (98%)	-0.22	12 (2%) 62 59	7, 31, 68, 100	0
4	F	544/548 (99%)	-0.00	20 (3%) 42 37	16, 42, 74, 101	0
5	B	95/95 (100%)	0.07	7 (7%) 15 11	11, 39, 79, 91	0
5	G	95/95 (100%)	0.14	5 (5%) 27 23	19, 41, 79, 91	0
All	All	1293/1303 (99%)	-0.07	46 (3%) 43 37	7, 37, 74, 101	0

The worst 5 of 46 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	A	345	GLU	9.6
5	G	1318	PHE	5.7
5	B	557	ASP	5.3
5	B	650	SER	5.1
4	F	836	PRO	4.6

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(\AA^2)	Q<0.9
6	SO4	A	901	5/5	0.96	0.13	-0.05	44,49,52,52	0
6	SO4	F	1249	5/5	0.97	0.10	-1.02	48,49,53,54	0

6.5 Other polymers [i](#)

There are no such residues in this entry.