Full wwPDB X-ray Structure Validation Report ⓘ

May 25, 2020 – 05:45 pm BST

PDB ID	:	5W0B
Title	:	Structure of human TUT7 catalytic module (CM)
Authors	:	Faehnle, C.R.; Walleshauser, J.; Joshua-Tor, L.
Deposited on	:	2017-05-30
Resolution	:	2.61 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:

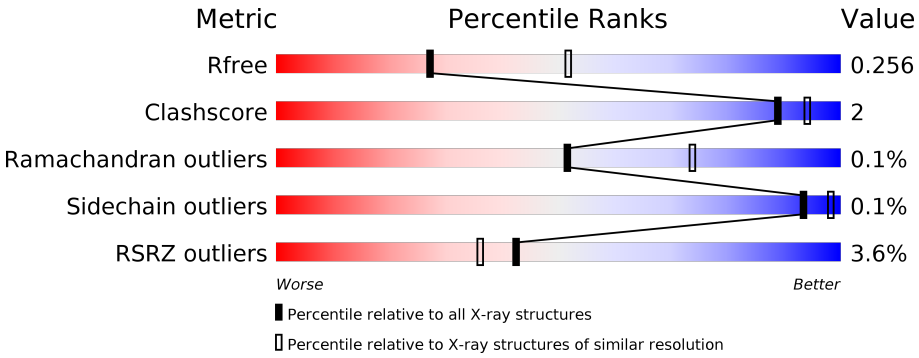
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, 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 2.61 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	3797 (2.64-2.60)
Clashscore	141614	4168 (2.64-2.60)
Ramachandran outliers	138981	4093 (2.64-2.60)
Sidechain outliers	138945	4093 (2.64-2.60)
RSRZ outliers	127900	3731 (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 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 <=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	389	3%, 82%, 14%
1	B	389	2%, 86%, 10%
1	C	389	5%, 85%, 6%, 9%

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 16803 atoms, of which 8381 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 Terminal uridylyltransferase 7.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	334	Total	C	H	N	O	S	0	0	0
			5373	1717	2702	454	481	19			
1	B	349	Total	C	H	N	O	S	0	0	0
			5659	1835	2832	466	512	14			
1	C	353	Total	C	H	N	O	S	0	0	0
			5710	1857	2847	473	519	14			

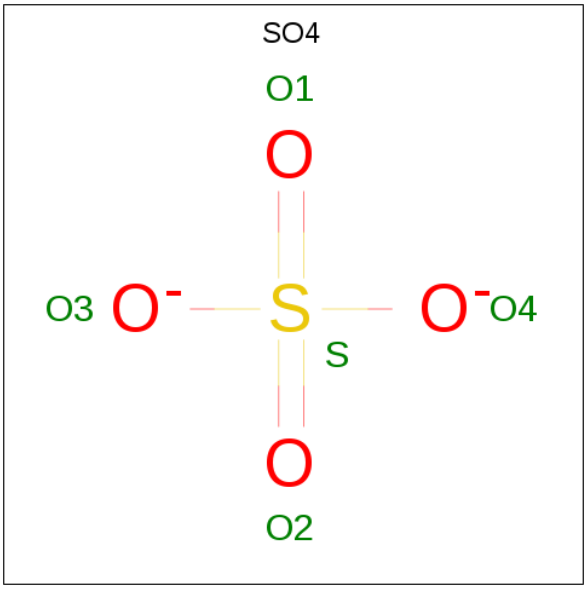
There are 21 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	977	GLY	-	expression tag	UNP Q5VYS8
A	978	ALA	-	expression tag	UNP Q5VYS8
A	979	GLY	-	expression tag	UNP Q5VYS8
A	980	ALA	-	expression tag	UNP Q5VYS8
A	981	GLY	-	expression tag	UNP Q5VYS8
A	982	SER	-	expression tag	UNP Q5VYS8
A	1060	ALA	ASP	engineered mutation	UNP Q5VYS8
B	977	GLY	-	expression tag	UNP Q5VYS8
B	978	ALA	-	expression tag	UNP Q5VYS8
B	979	GLY	-	expression tag	UNP Q5VYS8
B	980	ALA	-	expression tag	UNP Q5VYS8
B	981	GLY	-	expression tag	UNP Q5VYS8
B	982	SER	-	expression tag	UNP Q5VYS8
B	1060	ALA	ASP	engineered mutation	UNP Q5VYS8
C	977	GLY	-	expression tag	UNP Q5VYS8
C	978	ALA	-	expression tag	UNP Q5VYS8
C	979	GLY	-	expression tag	UNP Q5VYS8
C	980	ALA	-	expression tag	UNP Q5VYS8
C	981	GLY	-	expression tag	UNP Q5VYS8
C	982	SER	-	expression tag	UNP Q5VYS8
C	1060	ALA	ASP	engineered mutation	UNP Q5VYS8

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total Zn 1 1	0	0

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



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

- Molecule 4 is IODIDE ION (three-letter code: IOD) (formula: I).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total I 1 1	0	0
4	C	2	Total I 2 2	0	0

- Molecule 5 is water.

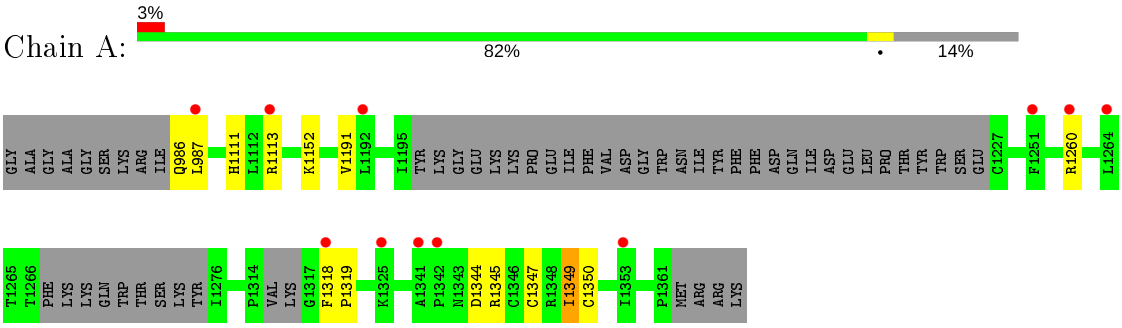
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	5	Total O 5 5	0	0
5	B	2	Total O 2 2	0	0
5	C	15	Total O 15 15	0	0

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.

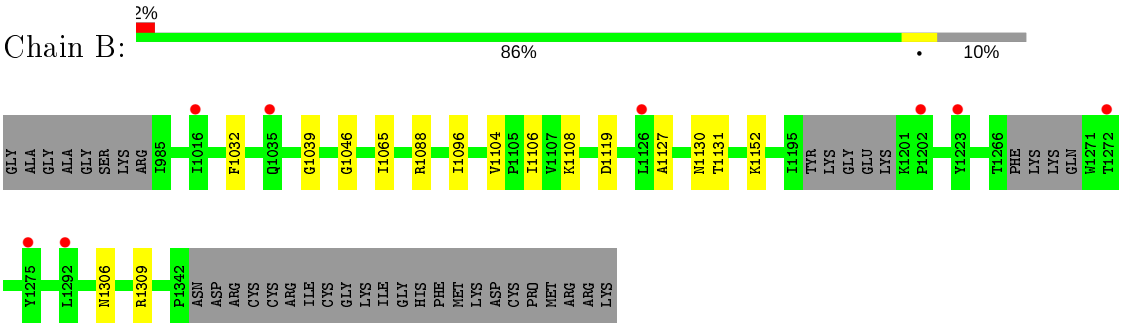
- Molecule 1: Terminal uridylyltransferase 7

Chain A:



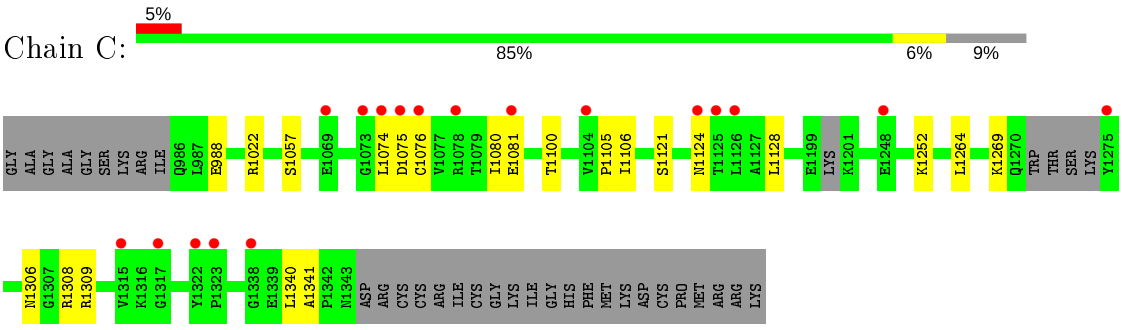
- Molecule 1: Terminal uridylyltransferase 7

Chain B:



- Molecule 1: Terminal uridylyltransferase 7

Chain C:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 61	Depositor
Cell constants a, b, c, α, β, γ	141.23Å 141.23Å 174.47Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	61.16 – 2.61 71.02 – 2.61	Depositor EDS
% Data completeness (in resolution range)	99.9 (61.16-2.61) 100.0 (71.02-2.61)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.28 (at 2.62Å)	Xtriage
Refinement program	PHENIX 1.10_2155	Depositor
R, R_{free}	0.231 , 0.255 0.231 , 0.256	Depositor DCC
R_{free} test set	3007 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å^2)	66.4	Xtriage
Anisotropy	0.348	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 49.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.037 for h,-h-k,-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	16803	wwPDB-VP
Average B, all atoms (Å^2)	95.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.34% 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

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, IOD, 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	0.25	0/2728	0.43	0/3682
1	B	0.25	0/2896	0.42	0/3921
1	C	0.26	0/2932	0.44	0/3965
All	All	0.25	0/8556	0.43	0/11568

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	2671	2702	2699	10	0
1	B	2827	2832	2838	11	0
1	C	2863	2847	2872	13	0
2	A	1	0	0	0	0
3	A	10	0	0	1	0
3	B	15	0	0	1	0
3	C	10	0	0	0	0
4	B	1	0	0	0	0
4	C	2	0	0	0	0
5	A	5	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	B	2	0	0	0	0
5	C	15	0	0	0	0
All	All	8422	8381	8409	32	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 (32) 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:1306:ASN:OD1	1:B:1309:ARG:NH2	2.13	0.81
1:C:1306:ASN:OD1	1:C:1309:ARG:NH2	2.35	0.60
1:A:986:GLN:OE1	1:A:987:LEU:N	2.36	0.58
1:A:1349:ILE:HD12	1:B:1032:PHE:HA	1.88	0.53
1:B:1046:GLY:H	1:B:1131:THR:HG21	1.75	0.52
1:A:1349:ILE:CD1	1:B:1032:PHE:HA	2.40	0.51
1:A:1191:VAL:HG11	1:A:1260:ARG:HA	1.92	0.51
1:C:1121:SER:OG	1:C:1124:ASN:OD1	2.13	0.50
1:A:1318:PHE:CD1	1:A:1319:PRO:HA	2.47	0.48
1:A:1152:LYS:NZ	3:A:1402:SO4:O2	2.39	0.47
1:B:1039:GLY:O	1:B:1065:ILE:HA	2.15	0.46
1:C:1100:THR:HA	1:C:1105:PRO:HB3	1.98	0.46
1:B:1130:ASN:OD1	1:B:1131:THR:N	2.48	0.46
1:C:1106:ILE:HG22	1:C:1121:SER:HB3	1.98	0.46
1:C:1074:LEU:O	1:C:1076:CYS:N	2.50	0.44
1:A:1344:ASP:OD1	1:A:1345:ARG:N	2.51	0.43
1:C:988:GLU:O	1:C:1308:ARG:NH2	2.39	0.43
1:C:1081:GLU:OE2	1:C:1105:PRO:HG2	2.18	0.43
1:C:1340:LEU:HD23	1:C:1341:ALA:O	2.18	0.43
1:B:1088:ARG:NH2	1:B:1096:ILE:O	2.52	0.43
1:A:1347:CYS:HB3	1:A:1350:CYS:SG	2.59	0.43
1:C:1264:LEU:HD21	1:C:1269:LYS:HD2	2.01	0.43
1:B:1152:LYS:NZ	3:B:1402:SO4:O2	2.49	0.42
1:C:1074:LEU:HD11	1:C:1080:ILE:HG12	2.01	0.42
1:C:1022:ARG:NH2	1:C:1057:SER:O	2.53	0.42
1:B:1108:LYS:NZ	1:B:1119:ASP:OD1	2.51	0.42
1:C:1252:LYS:HD3	1:C:1252:LYS:O	2.20	0.42
1:C:1124:ASN:O	1:C:1128:LEU:HD12	2.20	0.41
1:A:1347:CYS:SG	1:A:1349:ILE:HG23	2.60	0.41
1:A:1111:HIS:CE1	1:A:1113:ARG:HB2	2.56	0.41
1:B:1104:VAL:O	1:B:1106:ILE:HG23	2.21	0.40

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:1127:ALA:O	1:B:1131:THR:HG23	2.20	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	326/389 (84%)	318 (98%)	8 (2%)	0	100	100
1	B	343/389 (88%)	338 (98%)	5 (2%)	0	100	100
1	C	347/389 (89%)	333 (96%)	13 (4%)	1 (0%)	41	62
All	All	1016/1167 (87%)	989 (97%)	26 (3%)	1 (0%)	51	74

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	1075	ASP

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	300/348 (86%)	299 (100%)	1 (0%)	92	97
1	B	316/348 (91%)	316 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	C	319/348 (92%)	319 (100%)	0	100 100
All	All	935/1044 (90%)	934 (100%)	1 (0%)	93 98

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1349	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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

Of 11 ligands modelled in this entry, 4 are monoatomic - leaving 7 for Mogul analysis.

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	$\#\|Z\| > 2$
3	SO4	A	1403	-	4,4,4	0.14	0	6,6,6	0.05	0
3	SO4	B	1403	-	4,4,4	0.14	0	6,6,6	0.07	0
3	SO4	C	1402	-	4,4,4	0.13	0	6,6,6	0.05	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
3	SO4	A	1402	-	4,4,4	0.14	0	6,6,6	0.05	0
3	SO4	B	1402	-	4,4,4	0.14	0	6,6,6	0.05	0
3	SO4	C	1401	-	4,4,4	0.15	0	6,6,6	0.05	0
3	SO4	B	1401	-	4,4,4	0.13	0	6,6,6	0.06	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.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1402	SO4	1	0
3	B	1402	SO4	1	0

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

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, 95^{th} 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	334/389 (85%)	0.43	11 (3%) 46 40	56, 92, 161, 218	0
1	B	349/389 (89%)	0.39	8 (2%) 60 55	44, 76, 126, 155	0
1	C	353/389 (90%)	0.47	18 (5%) 28 22	48, 72, 139, 171	0
All	All	1036/1167 (88%)	0.43	37 (3%) 42 36	44, 79, 142, 218	0

All (37) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	1075	ASP	6.2
1	B	1223	TYR	6.0
1	C	1078	ARG	4.6
1	A	987	LEU	4.5
1	A	1192	LEU	4.4
1	A	1251	PHE	4.4
1	C	1315	VAL	4.3
1	C	1317	GLY	4.2
1	B	1275	TYR	4.1
1	C	1338	GLY	4.0
1	C	1104	VAL	3.8
1	C	1125	THR	3.2
1	C	1323	PRO	3.1
1	B	1292	LEU	3.0
1	C	1275	TYR	2.8
1	A	1341	ALA	2.7
1	C	1248	GLU	2.7
1	C	1076	CYS	2.7
1	A	1113	ARG	2.7
1	A	1260	ARG	2.6
1	C	1069	GLU	2.6
1	A	1342	PRO	2.5
1	C	1074	LEU	2.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	1325	LYS	2.4
1	C	1081	GLU	2.4
1	C	1073	GLY	2.4
1	C	1124	ASN	2.4
1	A	1318	PHE	2.3
1	A	1264	LEU	2.2
1	B	1202	PRO	2.2
1	A	1353	ILE	2.2
1	C	1126	LEU	2.1
1	B	1035	GLN	2.1
1	B	1272	THR	2.1
1	B	1016	ILE	2.1
1	B	1126	LEU	2.0
1	C	1322	TYR	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, 95^{th} 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
3	SO4	A	1403	5/5	0.86	0.14	139,139,139,140	0
3	SO4	C	1401	5/5	0.89	0.20	111,111,112,113	0
2	ZN	A	1401	1/1	0.92	0.11	107,107,107,107	0
3	SO4	B	1401	5/5	0.93	0.22	85,86,86,86	0
3	SO4	A	1402	5/5	0.94	0.19	94,96,96,97	0
3	SO4	C	1402	5/5	0.95	0.19	90,91,92,94	0
4	IOD	B	1404	1/1	0.97	0.19	152,152,152,152	0
4	IOD	C	1404	1/1	0.97	0.15	144,144,144,144	0
3	SO4	B	1402	5/5	0.97	0.28	80,83,86,87	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	SO4	B	1403	5/5	0.98	0.16	91,93,96,98	0
4	IOD	C	1403	1/1	1.00	0.22	103,103,103,103	0

6.5 Other polymers (i)

There are no such residues in this entry.