WORLDWIDE PDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Feb 22, 2024 – 06:11 PM EST

PDB ID	:	4RY8
Title	:	Crystal structure of 5-methylthioribose transporter solute binding protein TLET_1677 from Thermotoga lettingae TMO TARGET EFI-511109 in complex with 5-methylthioribose
Authors	:	Patskovsky, Y.; Toro, R.; Bhosle, R.; Al obaidi, N.; Morisco, L.L.; Wasserman, S.R.; Chamala, S.; Attonito, J.D.; Scott glenn, A.; Chowdhury, S.; Lafleur, J.; Hillerich, B.; Siede, R.D.; Love, J.; Whalen, K.L.; Gerlt, J.A.; Almo, S.C.; Enzyme Function Initiative (EFI)
Deposited on	:	2014-12-13
Resolution	:	1.75 Å(reported)

The following versions of software and data (see references (i)) 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	:	2.36
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)

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 1.75 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.
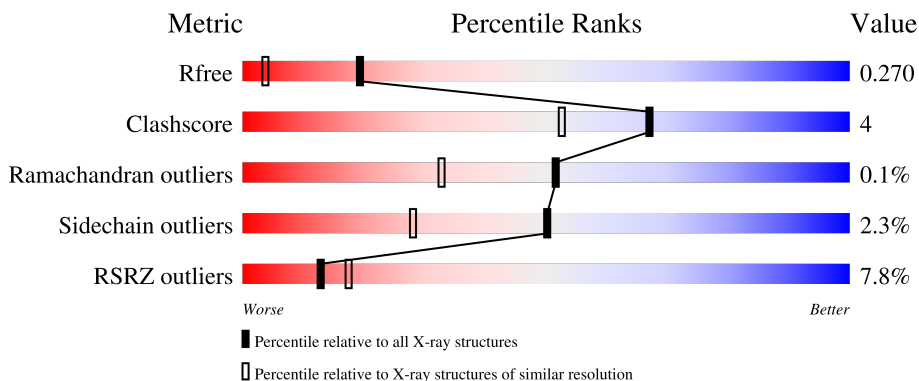
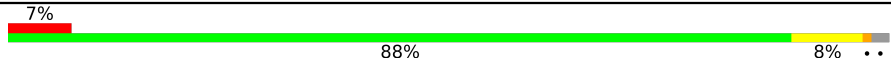| Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2340 (1.76-1.76)
Clashscore	141614	2466 (1.76-1.76)
Ramachandran outliers	138981	2437 (1.76-1.76)
Sidechain outliers	138945	2437 (1.76-1.76)
RSRZ outliers	127900	2298 (1.76-1.76)

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	329	9% / 88% / 7% · ·
1	B	329	7% / 86% / 11% ·
1	C	329	7% / 87% / 9% · ·

Continued on next page...

Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	D	329	

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 10412 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.
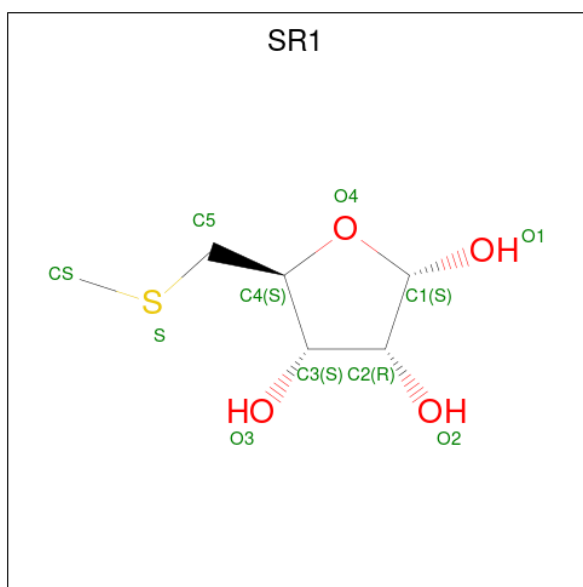
- Molecule 1 is a protein called Periplasmic binding protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	315	Total	C	N	O	S	0	1	0
			2435	1551	418	453	13			
1	B	321	Total	C	N	O	S	0	1	0
			2501	1592	431	465	13			
1	C	319	Total	C	N	O	S	0	1	0
			2473	1579	424	457	13			
1	D	321	Total	C	N	O	S	0	0	0
			2483	1582	425	463	13			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	20	SER	-	expression tag	UNP A8F7U7
A	21	MET	-	expression tag	UNP A8F7U7
B	20	SER	-	expression tag	UNP A8F7U7
B	21	MET	-	expression tag	UNP A8F7U7
C	20	SER	-	expression tag	UNP A8F7U7
C	21	MET	-	expression tag	UNP A8F7U7
D	20	SER	-	expression tag	UNP A8F7U7
D	21	MET	-	expression tag	UNP A8F7U7

- Molecule 2 is 5-S-methyl-5-thio-alpha-D-ribofuranose (three-letter code: SR1) (formula: $C_6H_{12}O_4S$).

| Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total 11	C 6	O 4	S 1	0	0
2	B	1	Total 11	C 6	O 4	S 1	0	0
2	C	1	Total 11	C 6	O 4	S 1	0	0
2	D	1	Total 11	C 6	O 4	S 1	0	0

- Molecule 3 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	113	Total 113	O 113	0	0
3	B	125	Total 125	O 125	0	0
3	C	121	Total 121	O 121	0	0
3	D	117	Total 117	O 117	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: Periplasmic binding protein
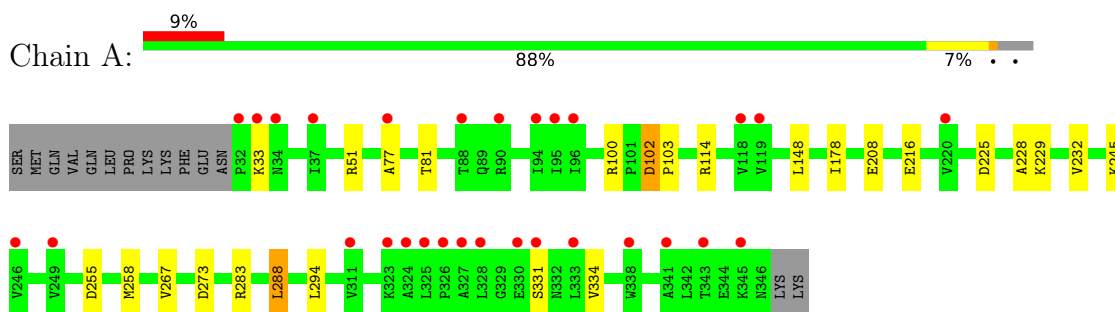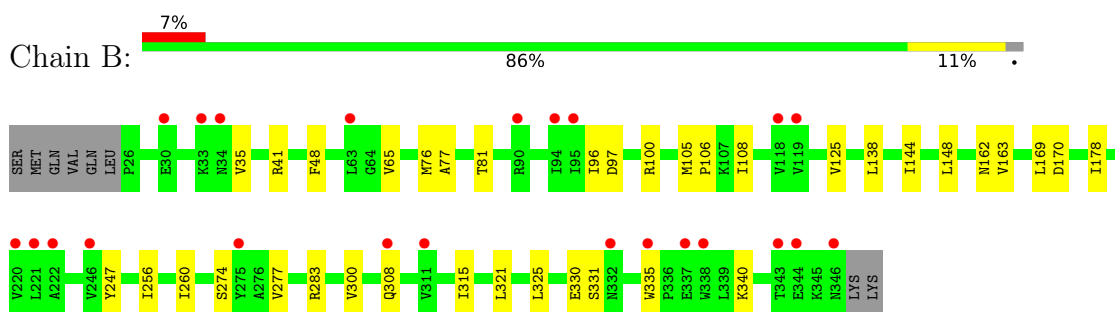
Chain A:

- Molecule 1: Periplasmic binding protein

Chain B:

- Molecule 1: Periplasmic binding protein

Chain C:

- Molecule 1: Periplasmic binding protein

Chain D:

SER MET GLN VAL GLN LEU P26 E30 N31 P32 K33 F48 A77 R78 T81 T88 I94 I95 I96 D97 H98 D102 K107 E110 R114 V118 V119 T120 V125 L148 I178 A194 R206 M207 E208 E216 A222 D225 K229

Q243 F244 K245 V249 D250 V251 D255 M258 Q261 V267 S274 R283 A284 A285 E301 R307 Q308 D319 K323 S331 N332 E337 W338 A341 N346 LYS LYS
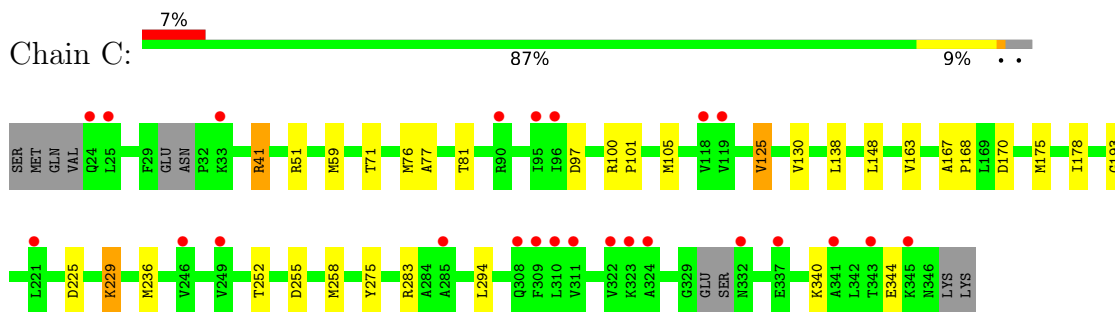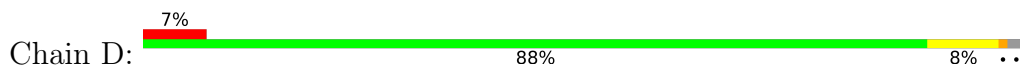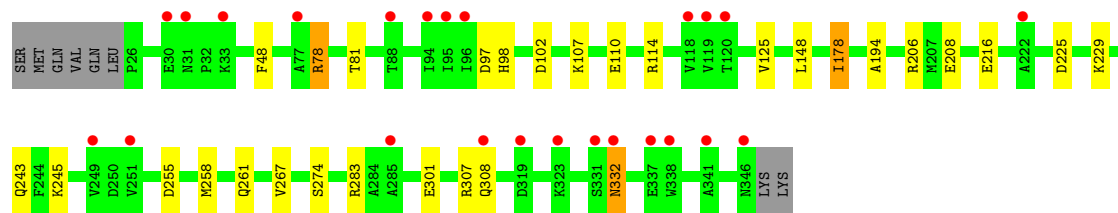
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	57.62Å 67.85Å 83.65Å 91.83° 92.39° 89.93°	Depositor
Resolution (Å)	50.00 – 1.75 24.67 – 1.75	Depositor EDS
% Data completeness (in resolution range)	97.0 (50.00-1.75) 96.7 (24.67-1.75)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.49 (at 1.75Å)	Xtriage
Refinement program	REFMAC 5.8.0073	Depositor
R, R_{free}	0.214 , 0.264 0.224 , 0.270	Depositor DCC
R_{free} test set	3775 reflections (3.05%)	wwPDB-VP
Wilson B-factor (Å^2)	18.2	Xtriage
Anisotropy	1.198	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 47.1	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.130 for -h,k,-l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	10412	wwPDB-VP
Average B, all atoms (Å^2)	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.61% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: SR1

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	0.68	0/2480	0.83	0/3364
1	B	0.67	0/2548	0.83	1/3453 (0.0%)
1	C	0.65	0/2518	0.82	3/3412 (0.1%)
1	D	0.65	0/2527	0.81	1/3427 (0.0%)
All	All	0.66	0/10073	0.82	5/13656 (0.0%)

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
1	A	0	1

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	41	ARG	NE-CZ-NH1	6.84	123.72	120.30
1	C	41	ARG	NE-CZ-NH2	-6.76	116.92	120.30
1	D	78	ARG	NE-CZ-NH2	5.57	123.08	120.30
1	B	170	ASP	CB-CG-OD1	5.25	123.03	118.30
1	C	170	ASP	CB-CG-OD1	5.20	122.98	118.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	331	SER	Peptide

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	2435	0	2465	20	0
1	B	2501	0	2538	22	0
1	C	2473	0	2506	20	0
1	D	2483	0	2510	24	0
2	A	11	0	0	0	0
2	B	11	0	0	0	0
2	C	11	0	0	0	0
2	D	11	0	0	0	0
3	A	113	0	0	1	0
3	B	125	0	0	0	0
3	C	121	0	0	1	0
3	D	117	0	0	3	0
All	All	10412	0	10019	78	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 4.

The worst 5 of 78 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:51[B]:ARG:HH21	1:A:51[B]:ARG:HG3	1.26	0.96
1:D:178:ILE:C	1:D:178:ILE:HD12	2.02	0.79
1:D:78:ARG:NH2	1:D:81:THR:OG1	2.23	0.71
1:A:51[B]:ARG:HH21	1:A:51[B]:ARG:CG	2.04	0.71
1:A:51[B]:ARG:HG3	1:A:51[B]:ARG:NH2	2.05	0.65

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	314/329 (95%)	307 (98%)	7 (2%)	0	100	100
1	B	320/329 (97%)	313 (98%)	6 (2%)	1 (0%)	41	22
1	C	314/329 (95%)	308 (98%)	6 (2%)	0	100	100
1	D	319/329 (97%)	310 (97%)	9 (3%)	0	100	100
All	All	1267/1316 (96%)	1238 (98%)	28 (2%)	1 (0%)	51	33

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	331	SER

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	254/272 (93%)	248 (98%)	6 (2%)	49	26
1	B	263/272 (97%)	258 (98%)	5 (2%)	57	37
1	C	258/272 (95%)	253 (98%)	5 (2%)	57	37
1	D	260/272 (96%)	252 (97%)	8 (3%)	40	17
All	All	1035/1088 (95%)	1011 (98%)	24 (2%)	50	28

5 of 24 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	229	LYS
1	D	107	LYS
1	D	102	ASP
1	D	125	VAL
1	B	125	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11

such sidechains are listed below:

Mol	Chain	Res	Type
1	C	85	ASN
1	C	157	ASN
1	D	157	ASN
1	D	85	ASN
1	B	85	ASN

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 monosaccharides in this entry.

5.6 Ligand geometry ⓘ

4 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	$\#|Z| > 2$
2	SR1	D	401	-	11,11,11	0.70	0	14,15,15	2.05	4 (28%)
2	SR1	A	401	-	11,11,11	0.73	0	14,15,15	2.07	4 (28%)
2	SR1	B	401	-	11,11,11	0.89	0	14,15,15	1.95	2 (14%)
2	SR1	C	401	-	11,11,11	0.92	1 (9%)	14,15,15	1.94	5 (35%)

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	SR1	D	401	-	-	0/3/19/19	0/1/1/1
2	SR1	A	401	-	-	0/3/19/19	0/1/1/1
2	SR1	B	401	-	-	0/3/19/19	0/1/1/1
2	SR1	C	401	-	-	0/3/19/19	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	401	SR1	O4-C4	-2.47	1.39	1.45

The worst 5 of 15 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	401	SR1	O1-C1-O4	-6.07	103.36	111.13
2	D	401	SR1	CS-S-C5	-5.18	91.78	101.30
2	A	401	SR1	O1-C1-O4	-4.61	105.23	111.13
2	A	401	SR1	CS-S-C5	-4.35	93.31	101.30
2	C	401	SR1	O4-C1-C2	4.05	109.44	104.46

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	315/329 (95%)	0.57	29 (9%)	9	11	17, 30, 55, 84	0
1	B	321/329 (97%)	0.53	23 (7%)	15	20	16, 30, 57, 85	0
1	C	319/329 (96%)	0.50	24 (7%)	14	19	17, 30, 57, 87	0
1	D	321/329 (97%)	0.55	24 (7%)	14	19	17, 30, 61, 82	0
All	All	1276/1316 (96%)	0.54	100 (7%)	13	17	16, 30, 59, 87	0

The worst 5 of 100 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	24	GLN	8.0
1	A	331	SER	7.7
1	A	33	LYS	4.6
1	D	341	ALA	4.1
1	B	118	VAL	3.9

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 ⓘ

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	SR1	A	401	11/11	0.94	0.14	17,22,35,37	0
2	SR1	B	401	11/11	0.94	0.13	18,21,30,31	0
2	SR1	C	401	11/11	0.95	0.13	20,22,30,43	0
2	SR1	D	401	11/11	0.96	0.13	19,21,30,40	0

6.5 Other polymers ⓘ

There are no such residues in this entry.