



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 5, 2026 – 12:32 PM UTC

PDB ID : 3ZC4 / pdb_00003zc4
Title : The structure of Csa5 from Sulfolobus solfataricus.
Authors : Reeks, J.; Anderson, L.; White, M.F.; Naismith, J.H.
Deposited on : 2012-11-15
Resolution : 2.72 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

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

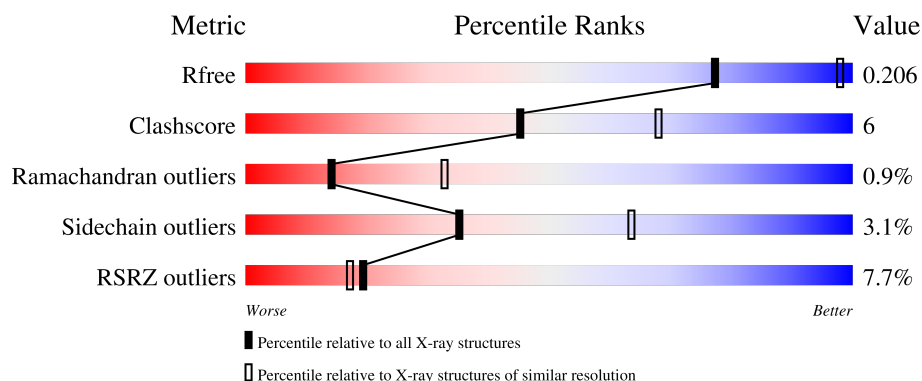
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.72 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}	180053	4348 (2.74-2.70)
Clashscore	190562	4665 (2.74-2.70)
Ramachandran outliers	187476	4584 (2.74-2.70)
Sidechain outliers	187428	4585 (2.74-2.70)
RSRZ outliers	180081	4348 (2.74-2.70)

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 $\leq 5\%$. 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	164	
1	B	164	
1	C	164	
1	D	164	
1	E	164	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	164	10% 80% 13% • 5%
1	G	164	7% 82% 12% • 5%
1	H	164	7% 82% 9% • 7%
1	I	164	6% 82% 12% • •

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 11552 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 SSO1398.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	157	Total	C	N	O	S	Se	0	0	0
			1270	824	205	238	2	1			
1	B	153	Total	C	N	O	S	Se	0	0	0
			1238	805	199	231	2	1			
1	C	154	Total	C	N	O	S	Se	0	0	0
			1246	809	200	234	2	1			
1	D	157	Total	C	N	O	S	Se	0	0	0
			1270	824	205	238	2	1			
1	E	158	Total	C	N	O	S	Se	0	1	0
			1288	837	206	242	2	1			
1	F	155	Total	C	N	O	S	Se	0	1	0
			1262	822	201	236	2	1			
1	G	155	Total	C	N	O	S	Se	0	0	0
			1254	813	202	236	2	1			
1	H	152	Total	C	N	O	S	Se	0	0	0
			1231	800	198	230	2	1			
1	I	157	Total	C	N	O	S	Se	0	0	0
			1270	824	205	238	2	1			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	GLY	-	cloning artifact	UNP Q97YC8
A	0	ALA	-	cloning artifact	UNP Q97YC8
B	-1	GLY	-	cloning artifact	UNP Q97YC8
B	0	ALA	-	cloning artifact	UNP Q97YC8
C	-1	GLY	-	cloning artifact	UNP Q97YC8
C	0	ALA	-	cloning artifact	UNP Q97YC8
D	-1	GLY	-	cloning artifact	UNP Q97YC8
D	0	ALA	-	cloning artifact	UNP Q97YC8
E	-1	GLY	-	cloning artifact	UNP Q97YC8
E	0	ALA	-	cloning artifact	UNP Q97YC8
F	-1	GLY	-	cloning artifact	UNP Q97YC8

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	0	ALA	-	cloning artifact	UNP Q97YC8
G	-1	GLY	-	cloning artifact	UNP Q97YC8
G	0	ALA	-	cloning artifact	UNP Q97YC8
H	-1	GLY	-	cloning artifact	UNP Q97YC8
H	0	ALA	-	cloning artifact	UNP Q97YC8
I	-1	GLY	-	cloning artifact	UNP Q97YC8
I	0	ALA	-	cloning artifact	UNP Q97YC8

- Molecule 2 is DI(HYDROXYETHYL)ETHER (CCD ID: PEG) (formula: $C_4H_{10}O_3$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	C	1	Total	C	O	0	0
			7	4	3		
2	D	1	Total	C	O	0	0
			7	4	3		
2	E	1	Total	C	O	0	0
			7	4	3		
2	F	1	Total	C	O	0	0
			7	4	3		
2	G	1	Total	C	O	0	0
			7	4	3		
2	H	1	Total	C	O	0	0
			7	4	3		
2	H	1	Total	C	O	0	0
			7	4	3		
2	I	1	Total	C	O	0	0
			7	4	3		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	I	1	Total	C	O	0	0
			7	4	3		

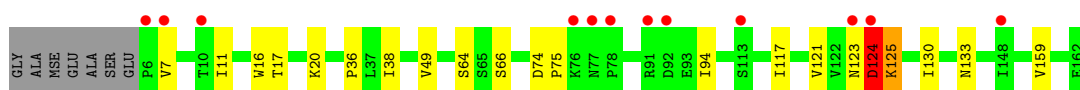
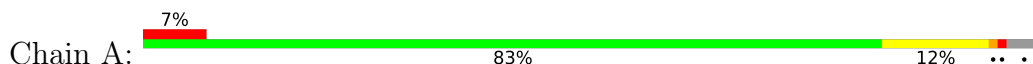
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	17	Total	O	0	0
			17	17		
3	B	13	Total	O	0	0
			13	13		
3	C	13	Total	O	0	0
			13	13		
3	D	22	Total	O	0	0
			22	22		
3	E	36	Total	O	0	0
			36	36		
3	F	26	Total	O	0	0
			26	26		
3	G	8	Total	O	0	0
			8	8		
3	H	11	Total	O	0	0
			11	11		
3	I	14	Total	O	0	0
			14	14		

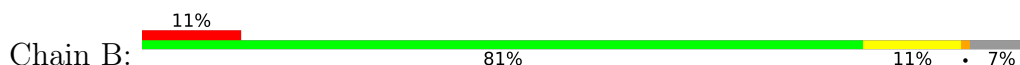
3 Residue-property plots [i](#)

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.

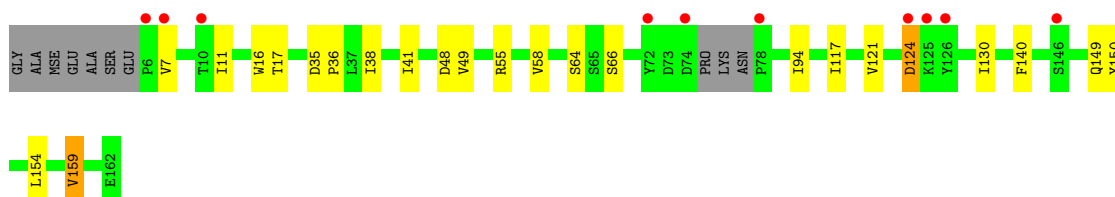
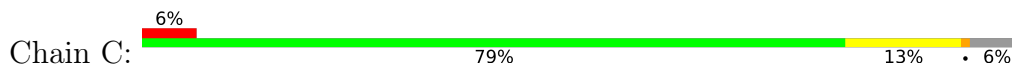
- Molecule 1: SSO1398



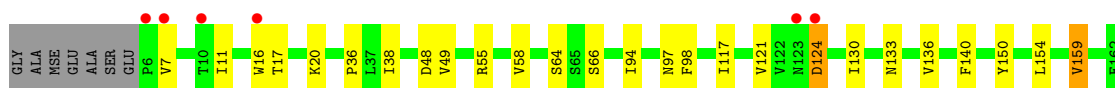
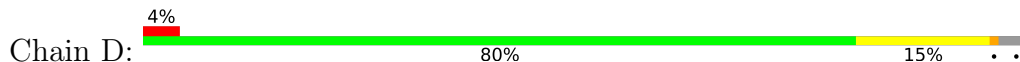
- Molecule 1: SSO1398



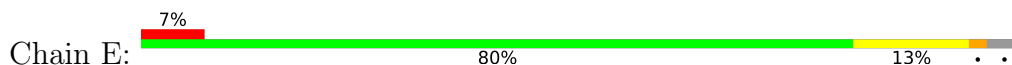
- Molecule 1: SSO1398



- Molecule 1: SSO1398



- Molecule 1: SSO1398





4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	285.00Å 53.72Å 216.68Å 90.00° 124.65° 90.00°	Depositor
Resolution (Å)	58.61 – 2.72 58.61 – 2.73	Depositor EDS
% Data completeness (in resolution range)	99.9 (58.61-2.72) 99.9 (58.61-2.73)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.58 (at 2.73Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, R_{free}	0.205 , 0.217 0.208 , 0.206	Depositor DCC
R_{free} test set	3694 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å ²)	55.0	Xtriage
Anisotropy	0.495	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 35.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	11552	wwPDB-VP
Average B, all atoms (Å ²)	42.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.79% 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: PEG

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	0.52	0/1298	0.77	0/1760
1	B	0.49	0/1264	0.70	1/1711 (0.1%)
1	C	0.49	0/1272	0.72	0/1722
1	D	0.53	0/1298	0.75	0/1760
1	E	0.57	0/1320	0.79	2/1791 (0.1%)
1	F	0.54	0/1293	0.76	0/1752
1	G	0.49	0/1280	0.71	0/1734
1	H	0.47	0/1256	0.70	0/1700
1	I	0.55	0/1298	0.76	0/1760
All	All	0.52	0/11579	0.74	3/15690 (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 (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	5	GLU	CA-C-N	5.20	126.34	119.84
1	E	5	GLU	C-N-CA	5.20	126.34	119.84
1	B	8	ALA	N-CA-C	5.04	117.55	111.40

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	123	ASN	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	1270	0	1291	15	0
1	B	1238	0	1261	17	0
1	C	1246	0	1265	21	0
1	D	1270	0	1291	25	0
1	E	1288	0	1305	20	0
1	F	1262	0	1281	18	0
1	G	1254	0	1270	17	0
1	H	1231	0	1253	15	0
1	I	1270	0	1291	19	0
2	C	7	0	10	0	0
2	D	7	0	10	0	0
2	E	7	0	10	0	0
2	F	7	0	10	0	0
2	G	7	0	10	0	0
2	H	14	0	20	0	0
2	I	14	0	20	0	0
3	A	17	0	0	2	0
3	B	13	0	0	0	0
3	C	13	0	0	0	0
3	D	22	0	0	1	0
3	E	36	0	0	1	0
3	F	26	0	0	0	0
3	G	8	0	0	0	0
3	H	11	0	0	0	0
3	I	14	0	0	1	0
All	All	11552	0	11598	135	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 6.

The worst 5 of 135 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:D:7:VAL:O	1:D:11:ILE:HD13	1.35	1.20
1:C:7:VAL:O	1:C:11:ILE:HD13	1.40	1.17
1:B:7:VAL:O	1:B:11:ILE:HD13	1.44	1.15
1:I:124:ASP:O	1:I:125:LYS:HB2	1.59	1.00
1:B:11:ILE:HD11	1:B:150:TYR:CE1	2.08	0.88

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	155/164 (94%)	148 (96%)	5 (3%)	2 (1%)	9	23
1	B	149/164 (91%)	144 (97%)	4 (3%)	1 (1%)	18	39
1	C	150/164 (92%)	144 (96%)	5 (3%)	1 (1%)	18	39
1	D	155/164 (94%)	150 (97%)	4 (3%)	1 (1%)	21	42
1	E	157/164 (96%)	149 (95%)	7 (4%)	1 (1%)	21	42
1	F	152/164 (93%)	145 (95%)	5 (3%)	2 (1%)	9	23
1	G	151/164 (92%)	145 (96%)	5 (3%)	1 (1%)	18	39
1	H	148/164 (90%)	142 (96%)	5 (3%)	1 (1%)	18	39
1	I	155/164 (94%)	147 (95%)	6 (4%)	2 (1%)	9	23
All	All	1372/1476 (93%)	1314 (96%)	46 (3%)	12 (1%)	14	33

5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	125	LYS
1	F	74	ASP
1	I	124	ASP
1	I	125	LYS
1	A	124	ASP

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	145/147 (99%)	142 (98%)	3 (2%)	47	73
1	B	141/147 (96%)	137 (97%)	4 (3%)	38	67
1	C	142/147 (97%)	136 (96%)	6 (4%)	26	53
1	D	145/147 (99%)	141 (97%)	4 (3%)	38	67
1	E	147/147 (100%)	142 (97%)	5 (3%)	32	60
1	F	144/147 (98%)	140 (97%)	4 (3%)	38	67
1	G	143/147 (97%)	138 (96%)	5 (4%)	32	59
1	H	140/147 (95%)	136 (97%)	4 (3%)	37	65
1	I	145/147 (99%)	140 (97%)	5 (3%)	32	60
All	All	1292/1323 (98%)	1252 (97%)	40 (3%)	35	63

5 of 40 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	96	GLN
1	I	14	ARG
1	G	124	ASP
1	H	66	SER
1	I	66	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	D	77	ASN
1	E	77	ASN
1	E	82	GLN
1	E	96	GLN
1	F	149	GLN

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 oligosaccharides in this entry.

5.6 Ligand geometry ⓘ

9 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	PEG	F	1163	-	6,6,6	0.47	0	5,5,5	0.25	0
2	PEG	H	1163	-	6,6,6	0.55	0	5,5,5	0.26	0
2	PEG	H	1164	-	6,6,6	0.50	0	5,5,5	0.22	0
2	PEG	G	1163	-	6,6,6	0.49	0	5,5,5	0.29	0
2	PEG	E	1163	-	6,6,6	0.70	0	5,5,5	0.42	0
2	PEG	C	1163	-	6,6,6	0.50	0	5,5,5	0.30	0
2	PEG	I	1163	-	6,6,6	0.52	0	5,5,5	0.19	0
2	PEG	D	1163	-	6,6,6	0.55	0	5,5,5	0.38	0
2	PEG	I	1164	-	6,6,6	0.63	0	5,5,5	0.40	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
2	PEG	F	1163	-	-	2/4/4/4	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	PEG	H	1163	-	-	2/4/4/4	-
2	PEG	H	1164	-	-	3/4/4/4	-
2	PEG	G	1163	-	-	2/4/4/4	-
2	PEG	E	1163	-	-	2/4/4/4	-
2	PEG	C	1163	-	-	1/4/4/4	-
2	PEG	I	1163	-	-	3/4/4/4	-
2	PEG	D	1163	-	-	0/4/4/4	-
2	PEG	I	1164	-	-	1/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 16 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	G	1163	PEG	O1-C1-C2-O2
2	I	1164	PEG	O1-C1-C2-O2
2	H	1163	PEG	O1-C1-C2-O2
2	H	1164	PEG	O1-C1-C2-O2
2	E	1163	PEG	O1-C1-C2-O2

There are no ring outliers.

No monomer is involved in short contacts.

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, 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	156/164 (95%)	0.27	12 (7%) 19 17	25, 40, 76, 108	0
1	B	152/164 (92%)	0.35	18 (11%) 9 8	28, 43, 90, 112	0
1	C	153/164 (93%)	0.26	10 (6%) 25 22	27, 42, 77, 99	0
1	D	156/164 (95%)	0.02	6 (3%) 44 40	17, 35, 56, 74	0
1	E	157/164 (95%)	0.18	11 (7%) 22 20	13, 30, 53, 71	1 (0%)
1	F	154/164 (93%)	0.24	16 (10%) 11 10	17, 32, 68, 101	1 (0%)
1	G	154/164 (93%)	0.25	12 (7%) 19 17	26, 39, 90, 116	0
1	H	151/164 (92%)	0.30	12 (7%) 18 16	31, 44, 86, 100	0
1	I	156/164 (95%)	0.35	10 (6%) 25 23	28, 41, 73, 109	0
All	All	1389/1476 (94%)	0.25	107 (7%) 19 17	13, 39, 78, 116	2 (0%)

The worst 5 of 107 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	H	79	VAL	8.2
1	B	16	TRP	7.1
1	F	75	PRO	6.9
1	I	6	PRO	6.6
1	C	125	LYS	5.8

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 oligosaccharides 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(Å ²)	Q<0.9
2	PEG	H	1164	7/7	0.76	0.24	66,68,71,72	0
2	PEG	D	1163	7/7	0.78	0.21	43,50,55,57	0
2	PEG	I	1164	7/7	0.79	0.20	42,46,51,53	0
2	PEG	F	1163	7/7	0.83	0.23	53,54,64,66	0
2	PEG	H	1163	7/7	0.86	0.15	45,49,50,51	0
2	PEG	C	1163	7/7	0.87	0.19	63,64,66,67	0
2	PEG	E	1163	7/7	0.88	0.16	32,40,46,47	0
2	PEG	I	1163	7/7	0.89	0.15	57,57,62,63	0
2	PEG	G	1163	7/7	0.93	0.13	49,52,54,55	0

6.5 Other polymers [i](#)

There are no such residues in this entry.