



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 20, 2026 – 03:44 PM UTC

PDB ID : 5XYR / pdb_00005xyr
Title : Crystal structure of a serine protease from Streptococcus species
Authors : Jobichen, C.; Sivaraman, J.
Deposited on : 2017-07-10
Resolution : 2.80 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

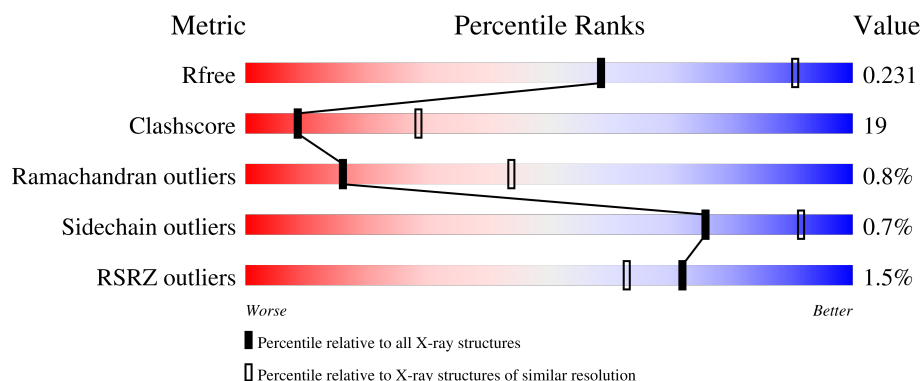
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.80 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	3866 (2.80-2.80)
Clashscore	190562	4276 (2.80-2.80)
Ramachandran outliers	187476	4196 (2.80-2.80)
Sidechain outliers	187428	4198 (2.80-2.80)
RSRZ outliers	180081	3869 (2.80-2.80)

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	1647	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	CL	A	1701	-	-	X	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	MLA	A	1706	-	X	-	-

2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 10518 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 Chemokine protease C.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1380	Total	C	N	O	Se	0	0	0
			10481	6584	1799	2074	24			

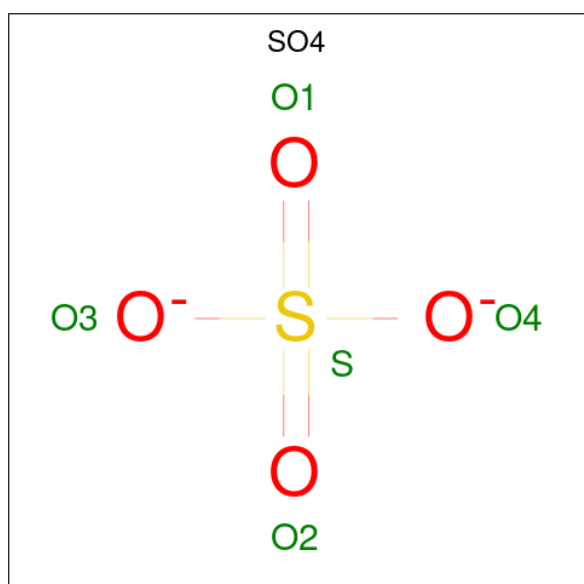
- Molecule 2 is CHLORIDE ION (CCD ID: CL) (formula: Cl).

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

- Molecule 3 is CALCIUM ION (CCD ID: CA) (formula: Ca).

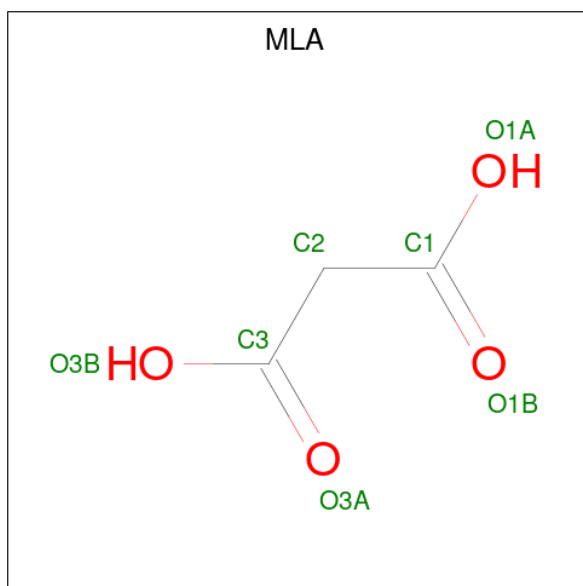
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	3	Total	Ca	0	0
			3	3		

- Molecule 4 is SULFATE ION (CCD ID: SO4) (formula: O₄S).



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

- Molecule 5 is MALONIC ACID (CCD ID: MLA) (formula: $C_3H_4O_4$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			7	3	4		

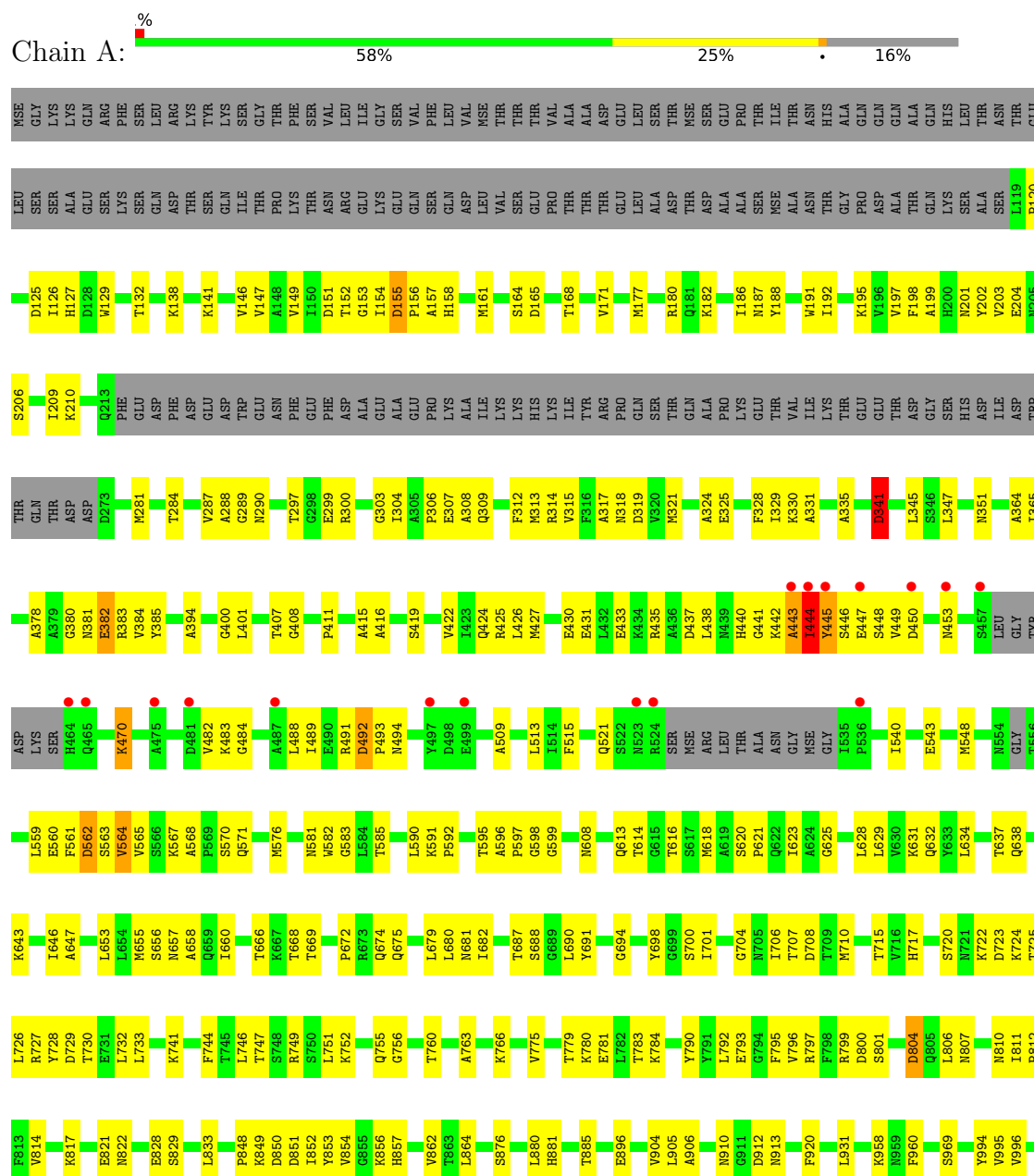
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	21	Total	O	0	0
			21	21		

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: Chemokine protease C



ALA	Y999	V1115	I1290	M1487
LEU	G1004	A1116	T1291	T1498
THR	I1117	A1118	Q1292	M1489
ALA	R1007	L1124	H1303	S1493
SER	M1010	F1125	D1307	Q1494
THR	T1011	I1133	I1317	D1500
PRO	F1012	R1134	V1318	L1503
ALA	D1013	L1135	R1319	L1503
LYS	M1014	K1136	E1320	A1514
THR	D1017	L1152	E1321	Q1515
THR	R1018	Q1156	Y1324	Y1526
PRO	Q1019	T1159	K1353	V1527
ALA	V1022	G1160	Y1359	K1533
THR	L1023	L1161	T1360	G1538
LYS	S1024	L1180	D1365	V1544
LEU	Q1025	M1183	S1370	S1545
PRO	F1028	D1186	D1371	L1546
SER	N1033	F1187	Y1376	Y1550
GLY	E1038	F1188	E1377	R1551
LYS	K1041	R1197	D1378	I1552
THR	D1042	K1197	R1379	E1553
ALA	R1043	V1200	V1383	L1570
LEU	A1046	Y1210	R1389	R1571
ARG	R1049	N1211	K1392	L1572
ILE	K1050	D1212	A1393	V1573
VAL	D1051	M1216	V1394	K1574
GLY	F1054	S1231	L1407	VAL
LEU	Y1074	S1232	P1408	GLY
THR	V1077	Q1233	V1409	ASP
CYS	E1078	G1237	P1410	ASP
VAL	D1079	A1238	E1411	ALA
VAL	N1080	M1257	D1412	SER
PHE	K1081	R1268	I1415	SER
SER	E1085	D1269	E1453	THR
ARG	R1086	K1273	L1454	LYS
LYS	L1094	E1274	L1455	ASP
SER	P1095	A1461	T1456	HIS
THR	L1096	I1281	Y1457	LYS
LYS	M1107	D1285	VAL	VAL
ASP	V1108	M1289	NSE	NSE
	E1109		SER	SER
	D1110		LYS	LYS
	F1111		ASN	ASN
	A1112		ASN	ASN
			GLN	GLN

4 Data and refinement statistics

Property	Value	Source
Space group	P 62 2 2	Depositor
Cell constants a, b, c, α , β , γ	191.62Å 191.62Å 250.96Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	19.99 – 2.80 19.99 – 2.80	Depositor EDS
% Data completeness (in resolution range)	99.3 (19.99-2.80) 94.8 (19.99-2.80)	Depositor EDS
R_{merge}	0.21	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.77 (at 2.79Å)	Xtriage
Refinement program	PHENIX (dev_2733: ???)	Depositor
R, R_{free}	0.209 , 0.257 (Not available) , 0.231	Depositor DCC
R_{free} test set	2000 reflections (2.93%)	wwPDB-VP
Wilson B-factor (Å ²)	65.5	Xtriage
Anisotropy	0.035	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.24 , 7.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.45$, $\langle L^2 \rangle = 0.27$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	10518	wwPDB-VP
Average B, all atoms (Å ²)	68.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.59% 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 [i](#)

5.1 Standard geometry [i](#)

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

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	0.56	0/10657	0.79	8/14404 (0.1%)

There are no bond length outliers.

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	1409	VAL	CA-C-N	6.67	126.63	119.76
1	A	1409	VAL	C-N-CA	6.67	126.63	119.76
1	A	445	TYR	CB-CA-C	-6.54	99.45	110.43
1	A	382	GLU	CA-C-N	6.43	133.82	121.54
1	A	382	GLU	C-N-CA	6.43	133.82	121.54

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts [i](#)

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	10481	0	10083	389	0
2	A	1	0	0	9	0
3	A	3	0	0	0	0
4	A	5	0	0	0	0
5	A	7	0	2	0	0
6	A	21	0	0	3	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	10518	0	10085	389	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 19.

The worst 5 of 389 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:383:ARG:NH2	1:A:746:LEU:CD1	1.77	1.45
1:A:383:ARG:NH2	1:A:746:LEU:HD12	1.06	1.36
1:A:447:GLU:OE2	1:A:453:ASN:CB	1.74	1.34
1:A:445:TYR:HE1	1:A:561:PHE:CE1	1.44	1.34
1:A:383:ARG:CZ	1:A:746:LEU:CD1	2.20	1.19

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	1370/1647 (83%)	1206 (88%)	153 (11%)	11 (1%)	16 44

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	482	VAL
1	A	540	ILE
1	A	564	VAL
1	A	1412	ASP
1	A	449	VAL

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	1104/1369 (81%)	1096 (99%)	8 (1%)	76 91

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1552	ILE
1	A	1411	GLU
1	A	1049	ARG
1	A	747	THR
1	A	1115	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	878	ASN
1	A	929	GLN
1	A	1494	GLN
1	A	1207	ASN
1	A	1402	ASN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

5.6 Ligand geometry

Of 6 ligands modelled in this entry, 4 are monoatomic - leaving 2 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$
5	MLA	A	1706	-	6,6,6	1.82	2 (33%)	7,7,7	1.98	4 (57%)
4	SO4	A	1705	-	4,4,4	0.54	0	6,6,6	0.95	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
5	MLA	A	1706	-	-	2/4/4/4	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	1706	MLA	C2-C3	2.86	1.55	1.51
5	A	1706	MLA	C2-C1	2.40	1.54	1.51

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	1706	MLA	O1A-C1-O1B	-2.75	116.25	123.33
5	A	1706	MLA	O3B-C3-C2	2.55	122.42	114.51
5	A	1706	MLA	O3B-C3-O3A	-2.52	116.85	123.33
5	A	1706	MLA	O1A-C1-C2	2.27	121.53	114.51

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	1706	MLA	O1B-C1-C2-C3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
5	A	1706	MLA	O1A-C1-C2-C3

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 [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	A	1355/1647 (82%)	-0.42	20 (1%) 72 63	35, 65, 118, 148	0

The worst 5 of 20 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	447	GLU	5.3
1	A	457	SER	4.8
1	A	464	HIS	4.4
1	A	445	TYR	4.3
1	A	1237	GLY	4.2

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 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
3	CA	A	1703	1/1	0.74	0.25	91,91,91,91	0
4	SO4	A	1705	5/5	0.85	0.17	66,71,88,105	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($\text{\AA}^2$)	Q<0.9
5	MLA	A	1706	7/7	0.91	0.20	57,64,69,71	0
3	CA	A	1704	1/1	0.99	0.02	51,51,51,51	0
3	CA	A	1702	1/1	0.99	0.04	44,44,44,44	0
2	CL	A	1701	1/1	0.99	0.06	45,45,45,45	0

6.5 Other polymers [i](#)

There are no such residues in this entry.