



wwPDB X-ray Structure Validation Summary Report ⓘ

May 6, 2021 – 02:11 PM JST

PDB ID : 7EDD
Title : Crystal structure of a serine protease from Streptococcus pyogenes
Authors : Jobichen, C.; Sivaraman, J.
Deposited on : 2021-03-15
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

<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	:	2.18
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)	:	2.18

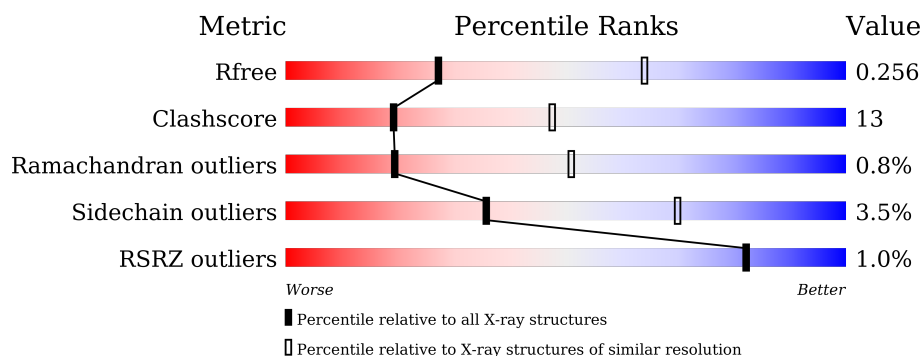
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.90 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}	130704	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (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 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	1519	<div> <div style="width: 100%; height: 10px; position: relative;"> <div style="position: absolute; top: -10px; left: 0; width: 100%;">%</div> <div style="position: absolute; top: 10px; left: 0; width: 100%;"> </div> </div> </div>

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
3	CA	A	1704	-	-	-	X
3	CA	A	1705	-	-	-	X

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	CA	A	1708	-	-	-	X
3	CA	A	1711	-	-	-	X
4	GOL	A	1714	-	X	-	X
7	CL	A	1720	-	-	X	-
7	CL	A	1721	-	-	X	-

2 Entry composition [i](#)

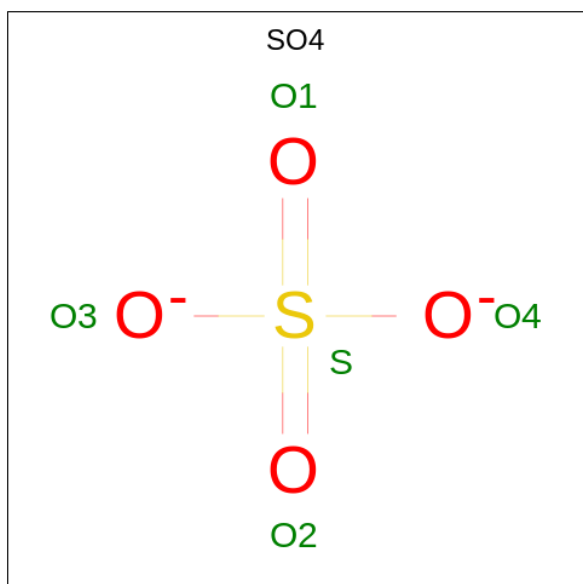
There are 8 unique types of molecules in this entry. The entry contains 10868 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 C5a peptidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1394	Total	C	N	O	Se	0	0	0
			10731	6749	1835	2121	26			

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O₄S).

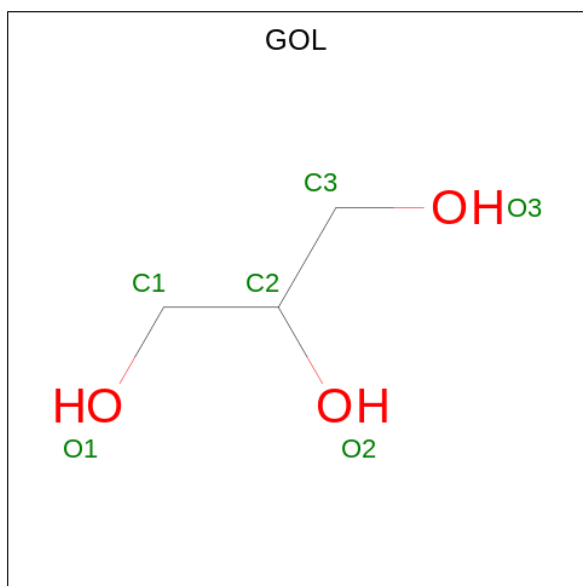


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	O	S	0	0
			5	4	1		
2	A	1	Total	O	S	0	0
			5	4	1		
2	A	1	Total	O	S	0	0
			5	4	1		

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	10	Total	Ca	0	0
			10	10		

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			6	3	3		
4	A	1	Total	C	O	0	0
			6	3	3		
4	A	1	Total	C	O	0	0
			6	3	3		
4	A	1	Total	C	O	0	0
			6	3	3		
4	A	1	Total	C	O	0	0
			6	3	3		

- Molecule 5 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	1	Total	Mg	0	0
			1	1		

- Molecule 6 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).



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

- Molecule 7 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

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

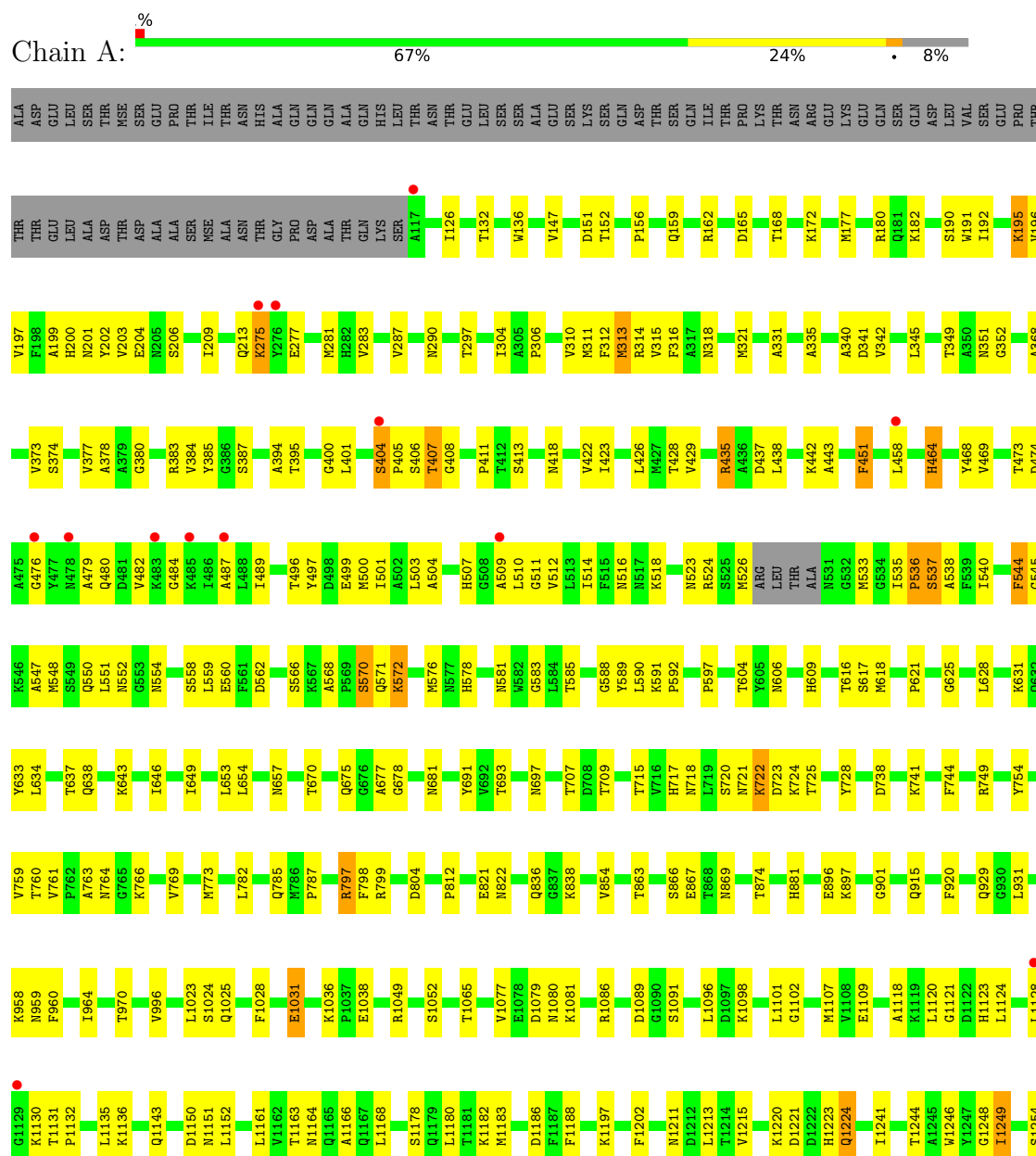
- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	72	Total	O	0	0
			72	72		

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: C5a peptidase



S1545	M1257
L1546	V1265
R1551	T1266
E1552	Y1267
E1553	R1268
G1554	D1269
N1555	E1270
E1564	H1271
V1565	Q1276
L1568	S1282
V1575	K1286
GLY	K1287
ASP	P1288
ALA	Y1288
SER	T1291
ASP	F1295
SER	D1296
THR	T1297
GLY	L1298
ASP	K1308
HIS	T1309
LYS	K1310
VAL	G1313
MSE	S1314
SER	S1315
LYS	G1316
ASN	I1317
ASN	F1323
GLN	K1327
ALA	R1331
LEU	V1335
THR	D1340
ALA	G1341
SER	I1342
ALA	K1348
THR	S1362
THR	K1363
THR	R1364
PRO	D1365
ALA	Y1373
THR	V1376
ALA	E1377
LYS	D1378
ALA	
	R1379
	V1383
	S1384
	T1387
	K1396
	V1400
	L1405
	V1409
	P1410
	E1411
	Q1414
	I1415
	V1416
	N1417
	L1421
	I1430
	Y1435
	Y1436
	G1440
	N1441
	S1442
	L1445
	K1449
	V1452
	L1455
	S1471
	F1480
	V1483
	M1487
	D1500
	H1501
	L1502
	S1507
	Y1526
	K1529
	V1544

4 Data and refinement statistics

Property	Value	Source
Space group	P 62 2 2	Depositor
Cell constants a, b, c, α , β , γ	191.97Å 191.97Å 252.79Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	48.37 – 2.90 48.37 – 2.90	Depositor EDS
% Data completeness (in resolution range)	99.0 (48.37-2.90) 99.0 (48.37-2.90)	Depositor EDS
R_{merge}	0.19	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.71 (at 2.91Å)	Xtriage
Refinement program	PHENIX (1.14_3260: ???)	Depositor
R, R_{free}	0.195 , 0.255 0.199 , 0.256	Depositor DCC
R_{free} test set	3004 reflections (4.93%)	wwPDB-VP
Wilson B-factor (Å ²)	55.8	Xtriage
Anisotropy	0.144	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 42.8	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	0.93	EDS
Total number of atoms	10868	wwPDB-VP
Average B, all atoms (Å ²)	52.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.57% 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 [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: SO4, PEG, CL, MG, GOL, 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	0.73	0/10917	0.84	0/14747

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 [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	10731	0	10467	269	0
2	A	15	0	0	1	0
3	A	10	0	0	0	0
4	A	30	0	39	4	0
5	A	1	0	0	1	0
6	A	7	0	10	0	0
7	A	2	0	0	6	0
8	A	72	0	0	5	0
All	All	10868	0	10516	275	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 13.

The worst 5 of 275 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)
7:A:1720:CL:CL	8:A:1869:HOH:O	2.04	1.12
7:A:1720:CL:CL	8:A:1867:HOH:O	2.15	0.99
1:A:378:ALA:HB1	1:A:616:THR:HG23	1.40	0.98
1:A:1363:LYS:HG3	1:A:1364:ARG:H	1.28	0.98
1:A:349:THR:HB	1:A:404:SER:HB2	1.48	0.96

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	1390/1519 (92%)	1261 (91%)	118 (8%)	11 (1%)	19 51

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	476	GLY
1	A	536	PRO
1	A	1128	LEU
1	A	1271	HIS
1	A	405	PRO

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	1158/1257 (92%)	1118 (96%)	40 (4%)	36 70

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

Mol	Chain	Res	Type
1	A	1086	ARG
1	A	1286	LYS
1	A	1220	LYS
1	A	1249	ILE
1	A	1310	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	318	ASN
1	A	480	GLN
1	A	494	ASN
1	A	507	HIS

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 monosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 22 ligands modelled in this entry, 13 are monoatomic - leaving 9 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	$\# Z > 2$
6	PEG	A	1719	-	6,6,6	0.74	0	5,5,5	0.99	0
4	GOL	A	1713	-	5,5,5	1.35	1 (20%)	5,5,5	1.06	0
4	GOL	A	1715	-	5,5,5	1.96	3 (60%)	5,5,5	0.65	0
2	SO4	A	1702	-	4,4,4	0.15	0	6,6,6	0.13	0
4	GOL	A	1714	-	5,5,5	2.80	3 (60%)	5,5,5	0.92	0
2	SO4	A	1701	-	4,4,4	0.54	0	6,6,6	0.75	0
4	GOL	A	1717	-	5,5,5	2.38	2 (40%)	5,5,5	1.31	0
4	GOL	A	1716	-	5,5,5	1.25	1 (20%)	5,5,5	1.21	0
2	SO4	A	1703	-	4,4,4	0.46	0	6,6,6	0.65	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	PEG	A	1719	-	-	3/4/4/4	-
4	GOL	A	1713	-	-	0/4/4/4	-
4	GOL	A	1715	-	-	0/4/4/4	-
4	GOL	A	1714	-	-	4/4/4/4	-
4	GOL	A	1717	-	-	2/4/4/4	-
4	GOL	A	1716	-	-	2/4/4/4	-

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	1717	GOL	C1-C2	4.37	1.69	1.51
4	A	1714	GOL	C3-C2	4.24	1.69	1.51
4	A	1714	GOL	C1-C2	3.71	1.67	1.51
4	A	1716	GOL	O2-C2	-2.63	1.35	1.43
4	A	1715	GOL	C1-C2	2.35	1.61	1.51

There are no bond angle outliers.

There are no chirality outliers.

5 of 11 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	1714	GOL	O1-C1-C2-O2
4	A	1714	GOL	O1-C1-C2-C3
4	A	1714	GOL	C1-C2-C3-O3
4	A	1714	GOL	O2-C2-C3-O3
6	A	1719	PEG	C4-C3-O2-C2

There are no ring outliers.

4 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	1714	GOL	2	0
2	A	1701	SO4	1	0
4	A	1717	GOL	1	0
4	A	1716	GOL	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 [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	1367/1519 (89%)	-0.27	14 (1%) 82 82	25, 50, 87, 120	0

The worst 5 of 14 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	487	ALA	3.3
1	A	404	SER	3.3
1	A	276	TYR	3.2
1	A	1414	GLN	2.8
1	A	1129	GLY	2.7

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 monosaccharides 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	1710	1/1	0.41	0.25	98,98,98,98	0
3	CA	A	1711	1/1	0.42	0.88	94,94,94,94	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
3	CA	A	1704	1/1	0.65	0.43	114,114,114,114	0
3	CA	A	1705	1/1	0.75	0.55	111,111,111,111	0
4	GOL	A	1714	6/6	0.77	0.69	56,61,62,65	0
3	CA	A	1708	1/1	0.79	0.41	99,99,99,99	0
6	PEG	A	1719	7/7	0.81	0.37	41,46,61,61	0
3	CA	A	1709	1/1	0.83	0.38	89,89,89,89	0
3	CA	A	1722	1/1	0.84	0.30	91,91,91,91	0
2	SO4	A	1703	5/5	0.88	0.27	60,69,103,114	0
2	SO4	A	1701	5/5	0.89	0.24	51,56,99,101	0
4	GOL	A	1717	6/6	0.91	0.25	38,51,53,58	0
4	GOL	A	1713	6/6	0.91	0.21	59,67,72,72	0
2	SO4	A	1702	5/5	0.94	0.16	93,94,99,111	0
3	CA	A	1707	1/1	0.95	0.15	41,41,41,41	0
4	GOL	A	1715	6/6	0.96	0.25	38,40,43,45	0
4	GOL	A	1716	6/6	0.96	0.16	39,44,45,48	0
3	CA	A	1712	1/1	0.98	0.12	45,45,45,45	0
3	CA	A	1706	1/1	0.99	0.09	39,39,39,39	0
7	CL	A	1721	1/1	0.99	0.26	60,60,60,60	0
7	CL	A	1720	1/1	1.00	0.21	10,10,10,10	0
5	MG	A	1718	1/1	1.00	0.22	11,11,11,11	0

6.5 Other polymers [i](#)

There are no such residues in this entry.