



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 22, 2024 – 06:38 PM EST

PDB ID : 4S37  
Title : Crystal structure of R2 pyocin membrane-piercing spike  
Authors : Browning, C.B.; Leiman, P.G.; Shneider, M.M.  
Deposited on : 2015-01-26  
Resolution : 2.20 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](mailto: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 : 4.02b-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)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.36

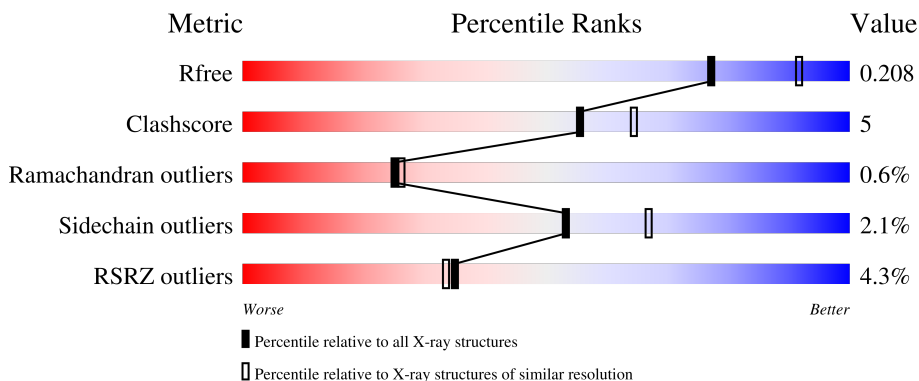
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.20 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	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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  $\geq 3$ , 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	185	 82% 9% 9%
1	B	185	 2% 84% 9% • 6%
1	C	185	 79% 14% • 7%
1	D	185	 79% 14% • 6%
1	E	185	 % 81% 12% 7%

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain	
1	F	185		82% 12% • 5%
1	G	185		84% 10% • 5%
1	H	185		85% 10% 5%
1	I	185		77% 16% • 5%
1	J	185		85% 9% 5%
1	K	185		88% 6% • 5%
1	L	185		84% 10% • 5%
1	M	185		78% 14% • 2%
1	N	185		80% 11% • 8%
1	O	185		83% 10% • 3%
1	P	185		76% 14% • 21%
1	Q	185		75% 15% 20%
1	R	185		75% 16% 19%

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	SO4	B	202	-	-	X	-
2	SO4	D	202	-	-	X	-
2	SO4	F	202	-	-	X	X
2	SO4	M	201	-	-	X	-
2	SO4	M	202	-	-	X	X
3	CL	G	205	-	-	X	-

## 2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 26421 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 Phage baseplate protein.

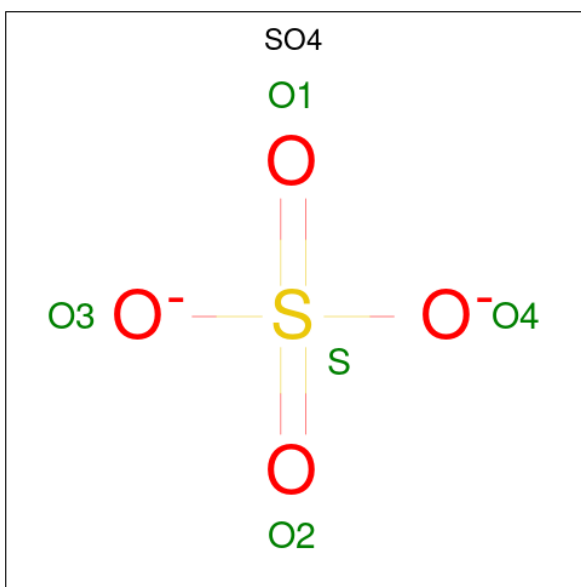
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	168	1287	806	233	243	5	0	6	0
1	B	174	1347	844	244	254	5	0	9	0
1	C	172	1346	838	248	255	5	0	9	0
1	D	173	1360	849	245	260	6	0	11	0
1	E	172	1343	837	243	258	5	0	10	0
1	F	175	1348	842	243	257	6	0	7	0
1	G	175	1342	837	242	258	5	0	6	0
1	H	176	1374	854	247	268	5	7	10	0
1	I	175	1340	839	241	255	5	0	6	0
1	J	175	1331	833	238	255	5	0	5	0
1	K	176	1368	853	246	263	6	0	9	0
1	L	175	1338	835	239	259	5	0	6	0
1	M	172	1293	809	233	246	5	8	3	0
1	N	171	1322	822	245	250	5	0	7	0
1	O	174	1294	809	231	249	5	0	2	0
1	P	166	1217	765	216	232	4	0	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	Q	166	Total	C	N	O	S	8	2	0
			1231	774	218	236	3			
1	R	169	Total	C	N	O	S	0	3	0
			1269	793	231	241	4			

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



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

Continued on next page...

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	E	1	Total	O	S	0	0
			5	4	1		
2	F	1	Total	O	S	0	0
			5	4	1		
2	F	1	Total	O	S	0	0
			5	4	1		
2	G	1	Total	O	S	0	0
			5	4	1		
2	G	1	Total	O	S	0	0
			5	4	1		
2	H	1	Total	O	S	0	0
			5	4	1		
2	H	1	Total	O	S	0	0
			5	4	1		
2	H	1	Total	O	S	0	0
			5	4	1		
2	I	1	Total	O	S	0	0
			5	4	1		
2	J	1	Total	O	S	0	0
			5	4	1		
2	J	1	Total	O	S	0	0
			5	4	1		
2	J	1	Total	O	S	0	0
			5	4	1		
2	K	1	Total	O	S	0	0
			5	4	1		
2	K	1	Total	O	S	0	0
			5	4	1		
2	L	1	Total	O	S	0	0
			5	4	1		
2	L	1	Total	O	S	0	0
			5	4	1		
2	M	1	Total	O	S	0	0
			5	4	1		
2	M	1	Total	O	S	0	0
			5	4	1		
2	M	1	Total	O	S	0	0
			5	4	1		
2	N	1	Total	O	S	0	0
			5	4	1		
2	O	1	Total	O	S	0	0
			5	4	1		

*Continued on next page...*

*Continued from previous page...*

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

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Cl	0	0
			1	1		
3	C	1	Total	Cl	0	0
			1	1		
3	D	2	Total	Cl	0	0
			2	2		
3	E	2	Total	Cl	0	0
			2	2		
3	F	1	Total	Cl	0	0
			1	1		
3	G	2	Total	Cl	0	0
			2	2		
3	H	2	Total	Cl	0	0
			2	2		
3	I	2	Total	Cl	0	0
			2	2		
3	J	2	Total	Cl	0	0
			2	2		
3	K	2	Total	Cl	0	0
			2	2		
3	L	1	Total	Cl	0	0
			1	1		
3	O	1	Total	Cl	0	0
			1	1		

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total C O 4 2 2	0	0
4	D	1	Total C O 4 2 2	0	0
4	F	1	Total C O 4 2 2	0	0
4	G	1	Total C O 4 2 2	0	0
4	K	1	Total C O 4 2 2	0	0
4	L	1	Total C O 4 2 2	0	0
4	N	1	Total C O 4 2 2	0	0

- Molecule 5 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	B	1	Total Ca 1 1	0	0
5	D	1	Total Ca 1 1	0	0
5	F	1	Total Ca 1 1	0	0
5	H	1	Total Ca 1 1	0	0
5	J	1	Total Ca 1 1	0	0

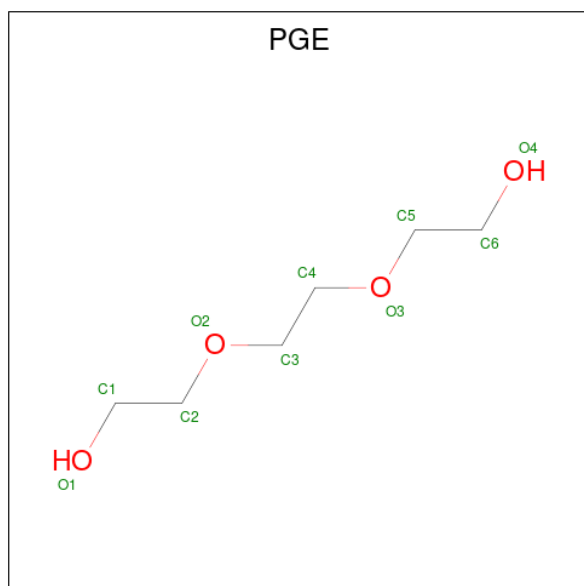
*Continued on next page...*



Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	O	1	Total	Ca	0	0
			1	1		

- Molecule 6 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula:  $C_6H_{14}O_4$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	H	1	Total	C	O	0	0
			10	6	4		

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	151	Total	O	0	0
			151	151		
7	B	166	Total	O	0	0
			166	166		
7	C	133	Total	O	0	0
			133	133		
7	D	173	Total	O	0	0
			173	173		
7	E	147	Total	O	0	0
			147	147		
7	F	149	Total	O	0	0
			149	149		
7	G	174	Total	O	0	0
			174	174		

Continued on next page...


*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
7	H	147	Total 147	O 147	0	0
7	I	138	Total 138	O 138	0	0
7	J	159	Total 159	O 159	0	0
7	K	140	Total 140	O 140	0	0
7	L	149	Total 149	O 149	0	0
7	M	105	Total 105	O 105	0	0
7	N	111	Total 111	O 111	0	0
7	O	106	Total 106	O 106	0	0
7	P	100	Total 100	O 100	0	0
7	Q	90	Total 90	O 90	0	0
7	R	95	Total 95	O 95	0	0

### 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: Phage baseplate protein

Chain A: 




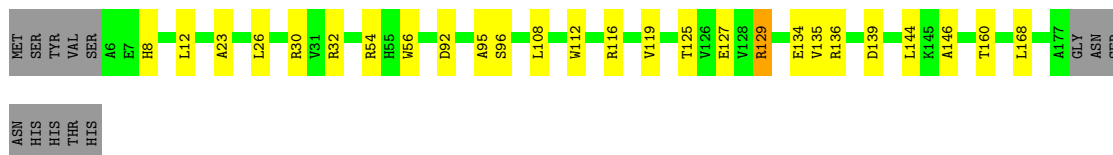
- Molecule 1: Phage baseplate protein

Chain B: 




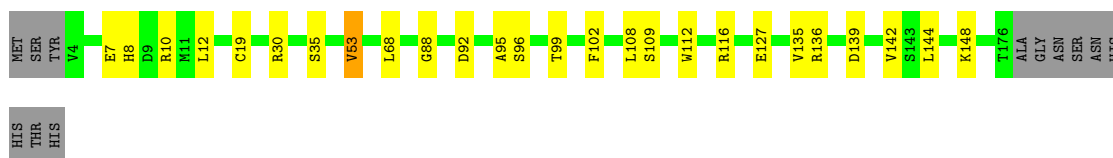
- Molecule 1: Phage baseplate protein

Chain C: 




- Molecule 1: Phage baseplate protein

Chain D: 

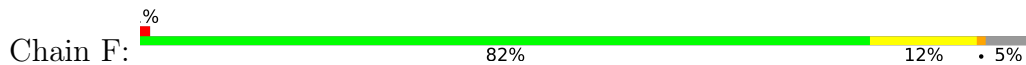


- Molecule 1: Phage baseplate protein

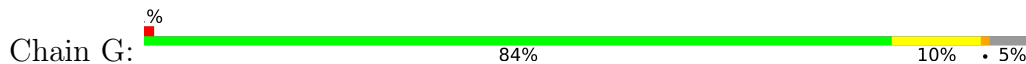
Chain E: 



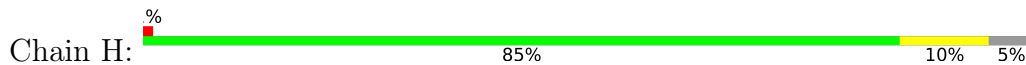
● Molecule 1: Phage baseplate protein



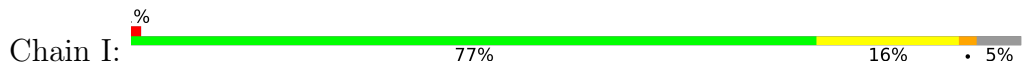
● Molecule 1: Phage baseplate protein



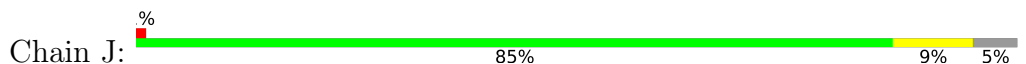
● Molecule 1: Phage baseplate protein



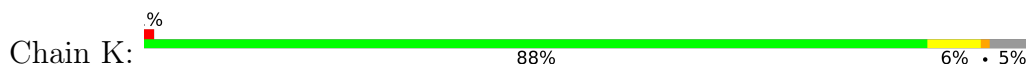
● Molecule 1: Phage baseplate protein




● Molecule 1: Phage baseplate protein



● Molecule 1: Phage baseplate protein




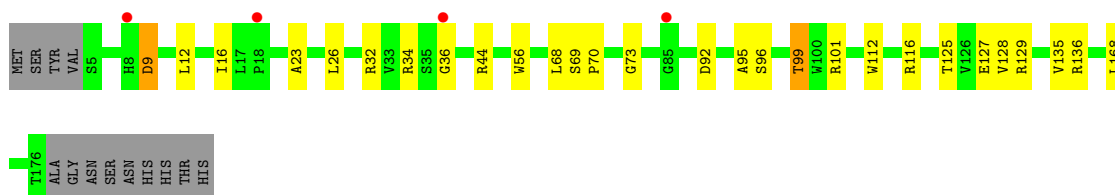
• Molecule 1: Phage baseplate protein

Chain L:  84% 10% • 5%




• Molecule 1: Phage baseplate protein

Chain M:  78% 14% • 7%




• Molecule 1: Phage baseplate protein

Chain N:  80% 11% • 8%




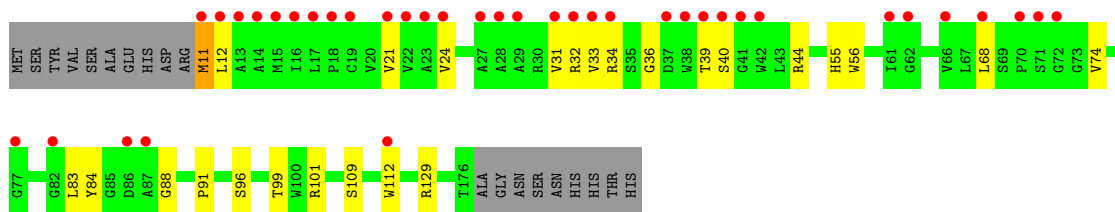
• Molecule 1: Phage baseplate protein

Chain O:  83% 10% • 6%




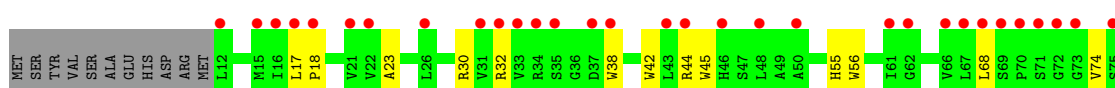
• Molecule 1: Phage baseplate protein

Chain P:  76% 14% • 10%



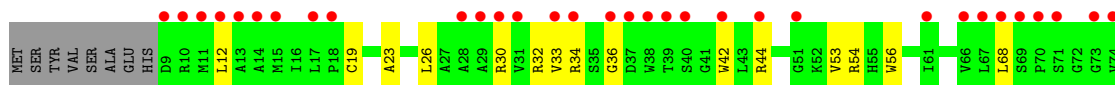
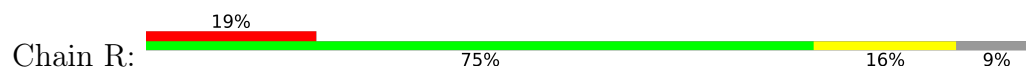
• Molecule 1: Phage baseplate protein

Chain Q:  75% 15% 10%





- Molecule 1: Phage baseplate protein



## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	94.94Å 137.47Å 164.76Å 90.00° 106.56° 90.00°	Depositor
Resolution (Å)	49.19 – 2.20 49.19 – 2.20	Depositor EDS
% Data completeness (in resolution range)	99.1 (49.19-2.20) 99.1 (49.19-2.20)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	0.06	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.66 (at 2.20Å)	Xtrriage
Refinement program	PHENIX (phenix.refine: 1.8.2_1309)	Depositor
R, $R_{free}$	0.163 , 0.208 0.163 , 0.208	Depositor DCC
$R_{free}$ test set	10160 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	38.2	Xtrriage
Anisotropy	0.103	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 43.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.28$	Xtrriage
Estimated twinning fraction	0.077 for h,-k,-h-l	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	26421	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	54.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.22% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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: CL, SO4, EDO, PGE, 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	0.27	0/1314	0.49	0/1788
1	B	0.26	0/1384	0.49	0/1883
1	C	0.26	0/1374	0.53	0/1868
1	D	0.27	0/1388	0.49	0/1889
1	E	0.27	0/1371	0.48	0/1866
1	F	0.28	0/1377	0.48	0/1874
1	G	0.28	0/1371	0.49	0/1866
1	H	0.33	0/1403	0.48	0/1910
1	I	0.26	0/1372	0.48	0/1867
1	J	0.26	0/1360	0.48	0/1852
1	K	0.28	0/1397	0.49	0/1902
1	L	0.28	0/1367	0.50	0/1861
1	M	0.24	0/1321	0.44	0/1798
1	N	0.23	0/1350	0.44	0/1835
1	O	0.24	0/1322	0.48	0/1801
1	P	0.22	0/1244	0.43	0/1696
1	Q	0.22	0/1258	0.42	0/1716
1	R	0.22	0/1296	0.42	0/1764
All	All	0.26	0/24269	0.47	0/33036

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	1287	0	1276	12	0
1	B	1347	0	1341	17	0
1	C	1346	0	1324	19	0
1	D	1360	0	1338	23	0
1	E	1343	0	1316	21	0
1	F	1348	0	1324	21	0
1	G	1342	0	1312	22	0
1	H	1374	0	1334	17	0
1	I	1340	0	1320	25	0
1	J	1331	0	1307	12	0
1	K	1368	0	1337	10	0
1	L	1338	0	1305	15	0
1	M	1293	0	1274	22	0
1	N	1322	0	1299	16	0
1	O	1294	0	1268	15	0
1	P	1217	0	1204	19	0
1	Q	1231	0	1215	19	0
1	R	1269	0	1254	25	0
2	A	5	0	0	0	0
2	B	15	0	0	5	0
2	C	15	0	0	0	0
2	D	10	0	0	3	0
2	E	5	0	0	0	0
2	F	10	0	0	3	0
2	G	10	0	0	1	0
2	H	15	0	0	0	0
2	I	5	0	0	0	0
2	J	15	0	0	0	0
2	K	10	0	0	0	0
2	L	10	0	0	0	0
2	M	15	0	0	4	0
2	N	5	0	0	0	0
2	O	15	0	0	1	0
2	P	5	0	0	0	0
2	R	10	0	0	0	0
3	A	1	0	0	0	0
3	C	1	0	0	1	0
3	D	2	0	0	1	0
3	E	2	0	0	0	0
3	F	1	0	0	1	0
3	G	2	0	0	3	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	H	2	0	0	1	0
3	I	2	0	0	0	0
3	J	2	0	0	1	0
3	K	2	0	0	1	0
3	L	1	0	0	0	0
3	O	1	0	0	1	0
4	B	4	0	6	0	0
4	D	4	0	6	0	0
4	F	4	0	6	1	0
4	G	4	0	6	0	0
4	K	4	0	6	0	0
4	L	4	0	6	0	0
4	N	4	0	6	0	0
5	B	1	0	0	0	0
5	D	1	0	0	0	0
5	F	1	0	0	0	0
5	H	1	0	0	0	0
5	J	1	0	0	0	0
5	O	1	0	0	0	0
6	H	10	0	14	0	0
7	A	151	0	0	2	1
7	B	166	0	0	3	0
7	C	133	0	0	1	1
7	D	173	0	0	3	0
7	E	147	0	0	1	0
7	F	149	0	0	2	0
7	G	174	0	0	2	0
7	H	147	0	0	1	1
7	I	138	0	0	3	0
7	J	159	0	0	3	0
7	K	140	0	0	1	0
7	L	149	0	0	3	0
7	M	105	0	0	0	0
7	N	111	0	0	0	0
7	O	106	0	0	1	0
7	P	100	0	0	2	0
7	Q	90	0	0	1	1
7	R	95	0	0	0	0
All	All	26421	0	23404	242	2

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 5.

The worst 5 of 242 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:C:30[B]:ARG:NH2	3:C:204:CL:CL	2.47	0.84
1:F:30:ARG:NE	3:F:205:CL:CL	2.46	0.81
1:G:30:ARG:NE	3:G:204:CL:CL	2.49	0.81
1:H:30:ARG:NE	3:H:207:CL:CL	2.52	0.80
1:D:136[B]:ARG:NH2	2:D:202:SO4:S	2.56	0.79

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
7:H:347:HOH:O	7:Q:283:HOH:O[1_645]	2.17	0.03
7:A:354:HOH:O	7:C:387:HOH:O[2_646]	2.18	0.02

## 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	172/185 (93%)	167 (97%)	4 (2%)	1 (1%)	25	26
1	B	181/185 (98%)	175 (97%)	5 (3%)	1 (1%)	25	26
1	C	179/185 (97%)	172 (96%)	7 (4%)	0	100	100
1	D	182/185 (98%)	177 (97%)	4 (2%)	1 (0%)	29	31
1	E	180/185 (97%)	174 (97%)	5 (3%)	1 (1%)	25	26
1	F	180/185 (97%)	172 (96%)	8 (4%)	0	100	100
1	G	179/185 (97%)	173 (97%)	6 (3%)	0	100	100
1	H	184/185 (100%)	176 (96%)	7 (4%)	1 (0%)	29	31
1	I	179/185 (97%)	171 (96%)	6 (3%)	2 (1%)	14	12
1	J	178/185 (96%)	172 (97%)	4 (2%)	2 (1%)	14	12

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	K	183/185 (99%)	175 (96%)	8 (4%)	0	100	100
1	L	179/185 (97%)	173 (97%)	5 (3%)	1 (1%)	25	26
1	M	173/185 (94%)	165 (95%)	7 (4%)	1 (1%)	25	26
1	N	176/185 (95%)	168 (96%)	6 (3%)	2 (1%)	14	12
1	O	174/185 (94%)	165 (95%)	7 (4%)	2 (1%)	14	12
1	P	164/185 (89%)	151 (92%)	12 (7%)	1 (1%)	25	26
1	Q	166/185 (90%)	157 (95%)	8 (5%)	1 (1%)	25	26
1	R	170/185 (92%)	163 (96%)	7 (4%)	0	100	100
All	All	3179/3330 (96%)	3046 (96%)	116 (4%)	17 (0%)	25	31

5 of 17 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	M	73	GLY
1	N	73	GLY
1	P	88	GLY
1	I	81	PRO
1	J	81	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	134/142 (94%)	130 (97%)	4 (3%)	41	53
1	B	141/142 (99%)	136 (96%)	5 (4%)	36	46
1	C	139/142 (98%)	135 (97%)	4 (3%)	42	54
1	D	143/142 (101%)	141 (99%)	2 (1%)	67	80
1	E	140/142 (99%)	139 (99%)	1 (1%)	84	91
1	F	141/142 (99%)	138 (98%)	3 (2%)	53	67
1	G	140/142 (99%)	138 (99%)	2 (1%)	67	80
1	H	144/142 (101%)	143 (99%)	1 (1%)	84	91

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	I	139/142 (98%)	135 (97%)	4 (3%)	42	54
1	J	139/142 (98%)	139 (100%)	0	100	100
1	K	143/142 (101%)	140 (98%)	3 (2%)	53	67
1	L	140/142 (99%)	138 (99%)	2 (1%)	67	80
1	M	134/142 (94%)	130 (97%)	4 (3%)	41	53
1	N	137/142 (96%)	133 (97%)	4 (3%)	42	54
1	O	134/142 (94%)	131 (98%)	3 (2%)	52	65
1	P	126/142 (89%)	123 (98%)	3 (2%)	49	62
1	Q	127/142 (89%)	124 (98%)	3 (2%)	49	62
1	R	131/142 (92%)	129 (98%)	2 (2%)	65	78
All	All	2472/2556 (97%)	2422 (98%)	50 (2%)	53	69

5 of 50 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	K	129	ARG
1	N	10	ARG
1	R	148	LYS
1	L	12	LEU
1	M	99	THR

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

### 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

Of 68 ligands modelled in this entry, 25 are monoatomic - leaving 43 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
2	SO4	H	203	-	4,4,4	0.14	0	6,6,6	0.06	0
2	SO4	L	203	-	4,4,4	0.14	0	6,6,6	0.06	0
2	SO4	H	204	-	4,4,4	0.14	0	6,6,6	0.06	0
2	SO4	J	203	-	4,4,4	0.14	0	6,6,6	0.05	0
2	SO4	F	202	-	4,4,4	0.14	0	6,6,6	0.05	0
4	EDO	D	201	-	3,3,3	0.47	0	2,2,2	0.29	0
2	SO4	F	203	-	4,4,4	0.14	0	6,6,6	0.06	0
4	EDO	L	201	-	3,3,3	0.46	0	2,2,2	0.30	0
2	SO4	R	201	-	4,4,4	0.14	0	6,6,6	0.04	0
2	SO4	E	201	-	4,4,4	0.15	0	6,6,6	0.05	0
2	SO4	C	202	-	4,4,4	0.14	0	6,6,6	0.05	0
2	SO4	J	201	-	4,4,4	0.14	0	6,6,6	0.06	0
2	SO4	M	202	-	4,4,4	0.14	0	6,6,6	0.05	0
2	SO4	D	202	-	4,4,4	0.13	0	6,6,6	0.10	0
2	SO4	L	202	-	4,4,4	0.14	0	6,6,6	0.08	0
2	SO4	C	203	-	4,4,4	0.14	0	6,6,6	0.05	0
2	SO4	D	203	-	4,4,4	0.15	0	6,6,6	0.04	0
4	EDO	N	201	-	3,3,3	0.46	0	2,2,2	0.33	0
2	SO4	A	201	-	4,4,4	0.14	0	6,6,6	0.06	0
4	EDO	B	201	-	3,3,3	0.46	0	2,2,2	0.34	0
2	SO4	O	201	-	4,4,4	0.16	0	6,6,6	0.07	0
2	SO4	G	202	-	4,4,4	0.14	0	6,6,6	0.04	0
2	SO4	J	202	-	4,4,4	0.14	0	6,6,6	0.04	0
2	SO4	G	203	-	4,4,4	0.14	0	6,6,6	0.06	0
2	SO4	N	202	-	4,4,4	0.14	0	6,6,6	0.04	0
2	SO4	B	202	-	4,4,4	0.13	0	6,6,6	0.12	0
2	SO4	C	201	-	4,4,4	0.14	0	6,6,6	0.06	0
2	SO4	M	201	-	4,4,4	0.14	0	6,6,6	0.06	0
4	EDO	K	201	-	3,3,3	0.44	0	2,2,2	0.34	0
2	SO4	B	203	-	4,4,4	0.14	0	6,6,6	0.08	0
2	SO4	P	201	-	4,4,4	0.14	0	6,6,6	0.05	0
2	SO4	I	201	-	4,4,4	0.14	0	6,6,6	0.06	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	SO4	K	202	-	4,4,4	0.14	0	6,6,6	0.05	0
4	EDO	F	201	-	3,3,3	0.48	0	2,2,2	0.29	0
2	SO4	R	202	-	4,4,4	0.14	0	6,6,6	0.06	0
2	SO4	B	204	-	4,4,4	0.14	0	6,6,6	0.08	0
2	SO4	H	202	-	4,4,4	0.14	0	6,6,6	0.06	0
2	SO4	K	203	-	4,4,4	0.14	0	6,6,6	0.07	0
4	EDO	G	201	-	3,3,3	0.46	0	2,2,2	0.33	0
2	SO4	M	203	-	4,4,4	0.14	0	6,6,6	0.05	0
2	SO4	O	202	-	4,4,4	0.16	0	6,6,6	0.12	0
6	PGE	H	201	-	9,9,9	0.31	0	8,8,8	0.28	0
2	SO4	O	203	-	4,4,4	0.14	0	6,6,6	0.05	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
4	EDO	B	201	-	-	0/1/1/1	-
4	EDO	G	201	-	-	0/1/1/1	-
4	EDO	K	201	-	-	0/1/1/1	-
6	PGE	H	201	-	-	2/7/7/7	-
4	EDO	D	201	-	-	0/1/1/1	-
4	EDO	L	201	-	-	0/1/1/1	-
4	EDO	N	201	-	-	0/1/1/1	-
4	EDO	F	201	-	-	0/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	H	201	PGE	C1-C2-O2-C3
6	H	201	PGE	C3-C4-O3-C5

There are no ring outliers.

10 monomers are involved in 18 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	F	202	SO4	2	0
2	F	203	SO4	1	0
2	M	202	SO4	2	0
2	D	202	SO4	3	0
2	G	202	SO4	1	0
2	B	202	SO4	4	0
2	M	201	SO4	2	0
4	F	201	EDO	1	0
2	B	204	SO4	1	0
2	O	202	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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	168/185 (90%)	-0.58	0 100 100	25, 38, 59, 99	0
1	B	174/185 (94%)	-0.51	3 (1%) 70 68	26, 37, 69, 135	0
1	C	172/185 (92%)	-0.55	0 100 100	25, 37, 59, 130	0
1	D	173/185 (93%)	-0.58	0 100 100	26, 35, 70, 139	0
1	E	172/185 (92%)	-0.50	1 (0%) 89 88	22, 37, 72, 135	0
1	F	175/185 (94%)	-0.44	2 (1%) 80 79	23, 37, 80, 182	0
1	G	175/185 (94%)	-0.54	2 (1%) 80 79	21, 35, 64, 102	0
1	H	176/185 (95%)	-0.43	1 (0%) 89 88	23, 43, 67, 92	0
1	I	175/185 (94%)	-0.52	2 (1%) 80 79	20, 39, 86, 144	0
1	J	175/185 (94%)	-0.55	1 (0%) 89 88	25, 37, 76, 106	0
1	K	176/185 (95%)	-0.53	2 (1%) 80 79	24, 38, 81, 119	0
1	L	175/185 (94%)	-0.58	0 100 100	26, 36, 68, 134	0
1	M	172/185 (92%)	-0.20	4 (2%) 60 58	32, 65, 101, 173	0
1	N	171/185 (92%)	-0.24	1 (0%) 89 88	32, 58, 103, 138	0
1	O	174/185 (94%)	-0.32	5 (2%) 51 49	31, 53, 109, 175	0
1	P	166/185 (89%)	0.82	38 (22%) 0 0	24, 97, 162, 196	0
1	Q	166/185 (89%)	0.86	37 (22%) 0 0	25, 87, 160, 178	0
1	R	169/185 (91%)	0.94	35 (20%) 1 1	26, 89, 155, 190	0
All	All	3104/3330 (93%)	-0.25	134 (4%) 35 33	20, 42, 131, 196	0

The worst 5 of 134 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	R	9	ASP	12.1
1	R	11	MET	11.7
1	Q	12	LEU	10.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	Q	71	SER	8.7
1	Q	72	GLY	8.2

## 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 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
2	SO4	R	201	5/5	0.52	0.40	198,200,201,203	0
5	CA	B	205	1/1	0.59	0.08	121,121,121,121	0
4	EDO	N	201	4/4	0.61	0.23	92,96,96,98	0
2	SO4	F	202	5/5	0.68	0.42	171,172,172,173	0
3	CL	O	205	1/1	0.70	0.14	108,108,108,108	0
2	SO4	M	202	5/5	0.73	0.50	180,180,181,182	0
5	CA	J	204	1/1	0.73	0.20	94,94,94,94	0
2	SO4	K	202	5/5	0.75	0.23	167,169,171,171	0
2	SO4	G	202	5/5	0.75	0.32	138,140,144,145	0
2	SO4	H	202	5/5	0.75	0.18	148,150,152,152	0
3	CL	A	202	1/1	0.75	0.10	106,106,106,106	0
2	SO4	M	203	5/5	0.76	0.23	144,144,145,146	0
2	SO4	J	202	5/5	0.77	0.15	174,174,175,176	0
2	SO4	N	202	5/5	0.77	0.24	176,178,179,181	0
2	SO4	O	201	5/5	0.79	0.32	133,140,141,145	0
3	CL	D	206	1/1	0.79	0.10	91,91,91,91	0
3	CL	H	207	1/1	0.79	0.09	92,92,92,92	0
3	CL	I	203	1/1	0.79	0.10	85,85,85,85	0
3	CL	E	203	1/1	0.81	0.16	95,95,95,95	0
3	CL	F	205	1/1	0.81	0.15	88,88,88,88	0
2	SO4	M	201	5/5	0.82	0.20	167,168,170,171	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	EDO	F	201	4/4	0.83	0.21	65,68,70,72	0
2	SO4	G	203	5/5	0.83	0.17	117,124,127,128	0
4	EDO	L	201	4/4	0.84	0.13	55,55,59,62	0
2	SO4	R	202	5/5	0.84	0.24	142,144,146,147	0
2	SO4	B	203	5/5	0.84	0.22	130,132,133,135	0
2	SO4	E	201	5/5	0.84	0.28	148,151,153,154	0
2	SO4	J	201	5/5	0.85	0.24	149,150,151,154	0
3	CL	I	202	1/1	0.85	0.30	97,97,97,97	0
2	SO4	P	201	5/5	0.85	0.26	146,148,151,152	0
5	CA	O	204	1/1	0.85	0.05	113,113,113,113	0
6	PGE	H	201	10/10	0.86	0.26	70,73,78,78	0
2	SO4	C	202	5/5	0.87	0.16	139,140,142,143	0
2	SO4	A	201	5/5	0.87	0.11	110,112,116,121	0
2	SO4	C	201	5/5	0.87	0.25	148,149,149,150	0
3	CL	K	205	1/1	0.87	0.08	83,83,83,83	0
3	CL	G	204	1/1	0.88	0.12	89,89,89,89	0
3	CL	E	202	1/1	0.89	0.08	93,93,93,93	0
5	CA	D	204	1/1	0.89	0.06	99,99,99,99	0
4	EDO	G	201	4/4	0.89	0.10	55,62,63,63	0
3	CL	J	206	1/1	0.89	0.11	96,96,96,96	0
4	EDO	D	201	4/4	0.89	0.12	46,51,55,58	0
2	SO4	O	203	5/5	0.90	0.09	129,129,131,132	0
3	CL	K	204	1/1	0.90	0.11	85,85,85,85	0
2	SO4	J	203	5/5	0.91	0.12	95,96,100,107	0
2	SO4	K	203	5/5	0.92	0.10	126,127,129,129	0
2	SO4	I	201	5/5	0.93	0.17	132,134,136,138	0
2	SO4	F	203	5/5	0.93	0.16	107,108,111,111	0
2	SO4	H	203	5/5	0.93	0.23	137,137,138,140	0
2	SO4	H	204	5/5	0.93	0.14	97,99,103,106	0
4	EDO	B	201	4/4	0.93	0.13	60,61,62,68	0
4	EDO	K	201	4/4	0.94	0.13	59,64,73,79	0
2	SO4	L	203	5/5	0.94	0.12	112,117,119,119	0
2	SO4	D	203	5/5	0.95	0.09	99,100,101,103	0
3	CL	L	204	1/1	0.95	0.07	74,74,74,74	0
2	SO4	C	203	5/5	0.95	0.12	109,110,113,114	0
3	CL	H	206	1/1	0.96	0.19	99,99,99,99	0
3	CL	D	205	1/1	0.96	0.11	79,79,79,79	0
2	SO4	B	202	5/5	0.96	0.17	67,89,97,106	0
2	SO4	B	204	5/5	0.97	0.15	81,81,85,89	0
3	CL	J	205	1/1	0.97	0.10	76,76,76,76	0
5	CA	H	205	1/1	0.97	0.19	69,69,69,69	0
2	SO4	D	202	5/5	0.98	0.11	54,55,58,73	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	SO4	L	202	5/5	0.98	0.10	58,61,62,62	0
3	CL	C	204	1/1	0.98	0.09	92,92,92,92	0
2	SO4	O	202	5/5	0.98	0.17	57,63,77,79	0
3	CL	G	205	1/1	0.99	0.19	65,65,65,65	0
5	CA	F	204	1/1	0.99	0.19	72,72,72,72	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.