



# Full wwPDB X-ray Structure Validation Report ⓘ

May 21, 2020 – 10:14 pm BST

PDB ID : 4K38  
Title : Native anSMEcpe with bound AdoMet and Kp18Cys peptide  
Authors : Goldman, P.J.; Drennan, C.L.  
Deposited on : 2013-04-10  
Resolution : 1.83 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 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 : 2.11  
buster-report : 1.1.7 (2018)  
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) : 2.11

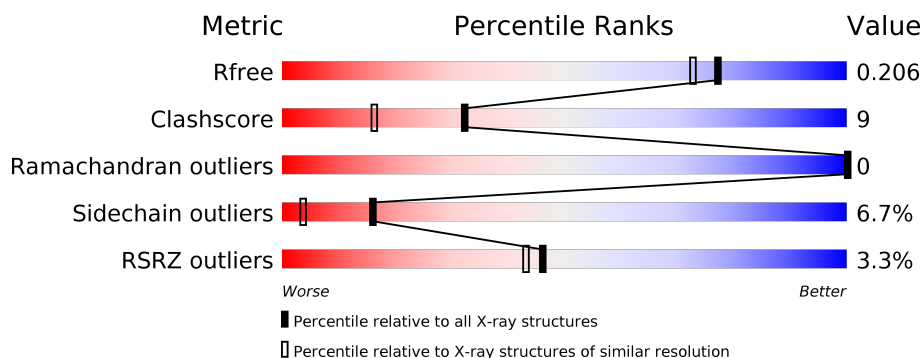
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.83 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	4003 (1.86-1.82)
Clashscore	141614	4233 (1.86-1.82)
Ramachandran outliers	138981	4185 (1.86-1.82)
Sidechain outliers	138945	4186 (1.86-1.82)
RSRZ outliers	127900	3957 (1.86-1.82)

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 on 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	370	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 0%, red 1%, green 1%, green 86%, yellow 86%, yellow 98%, red 98%, red 100%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>%</span> <span>86%</span> <span>12%</span> <span>.</span> </div> </div>
1	B	370	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 0%, red 4%, green 4%, green 77%, yellow 77%, yellow 94%, red 94%, red 100%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>4%</span> <span>77%</span> <span>17%</span> <span>...</span> </div> </div>
2	C	18	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 0%, red 17%, green 17%, green 39%, yellow 39%, yellow 56%, orange 56%, orange 62%, grey 62%, grey 100%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>17%</span> <span>39%</span> <span>17%</span> <span>6%</span> <span>39%</span> </div> </div>
2	D	18	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 0%, red 6%, green 6%, green 28%, yellow 28%, yellow 50%, orange 50%, orange 56%, grey 56%, grey 100%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>6%</span> <span>28%</span> <span>22%</span> <span>6%</span> <span>44%</span> </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 crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	GOL	B	510	-	-	X	-

## 2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 6764 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.

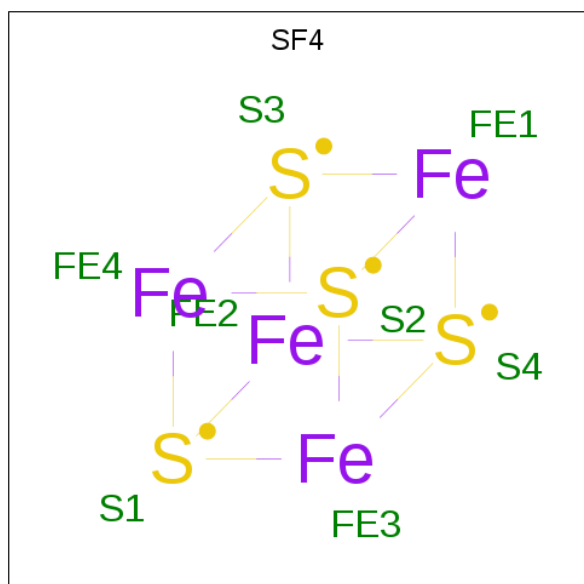
- Molecule 1 is a protein called Anaerobic sulfatase-maturing enzyme.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	369	Total	C	N	O	S	0	3	0
			3037	1948	507	557	25			
1	B	362	Total	C	N	O	S	0	2	0
			2971	1910	492	544	25			

- Molecule 2 is a protein called Kp18Cys peptide.

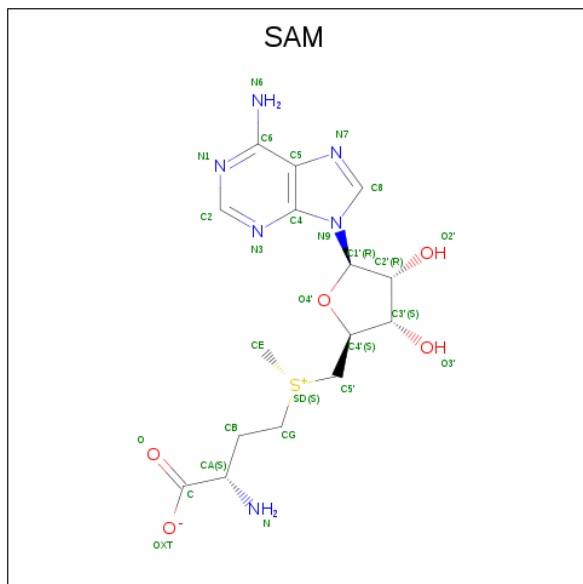
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	11	Total	C	N	O	S	0	0	0
			71	42	14	13	2			
2	D	10	Total	C	N	O	S	0	0	0
			68	40	13	13	2			

- Molecule 3 is IRON/SULFUR CLUSTER (three-letter code: SF4) (formula: Fe<sub>4</sub>S<sub>4</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	Fe	S	0	0
			8	4	4		
3	A	1	Total	Fe	S	0	0
			8	4	4		
3	A	1	Total	Fe	S	0	0
			8	4	4		
3	B	1	Total	Fe	S	0	0
			8	4	4		
3	B	1	Total	Fe	S	0	0
			8	4	4		
3	B	1	Total	Fe	S	0	0
			8	4	4		

- Molecule 4 is S-ADENOSYLMETHIONINE (three-letter code: SAM) (formula:  $C_{15}H_{22}N_6O_5S$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	N	O	S	0	0
			27	15	6	5	1		
4	B	1	Total	C	N	O	S	0	0
			27	15	6	5	1		

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



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

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	B	1	Total	C	O	0	0
			6	3	3		

- Molecule 6 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

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

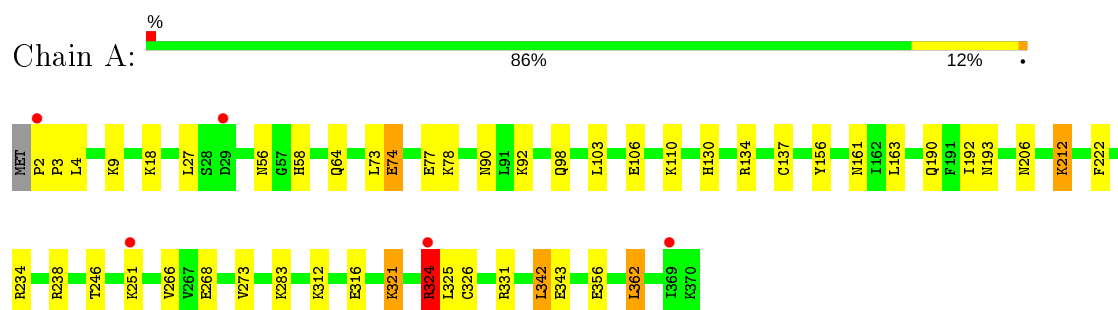
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	246	Total	O	0	0
			246	246		
7	B	165	Total	O	0	0
			165	165		
7	C	5	Total	O	0	0
			5	5		
7	D	6	Total	O	0	0
			6	6		

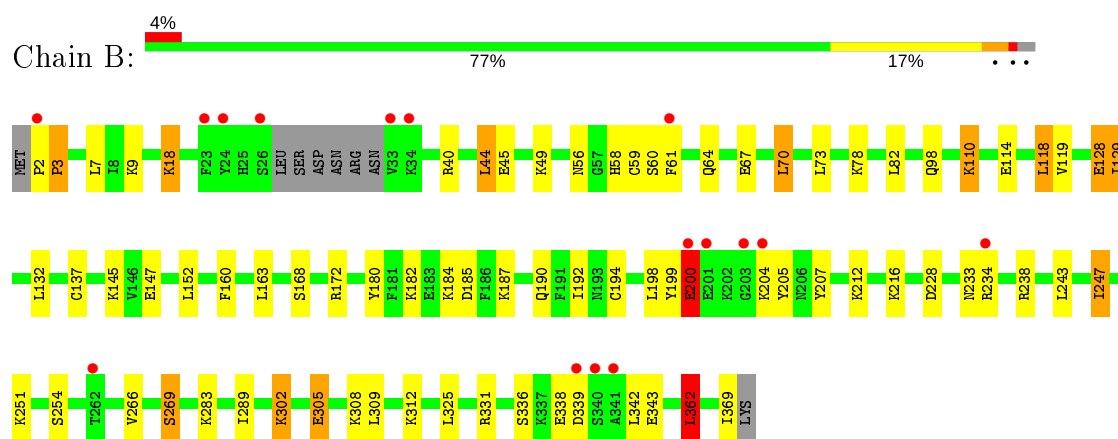
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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.

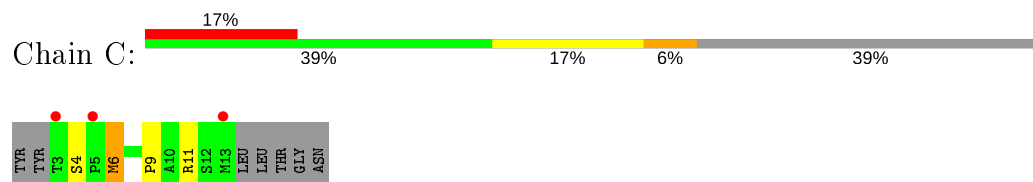
- Molecule 1: Anaerobic sulfatase-maturing enzyme



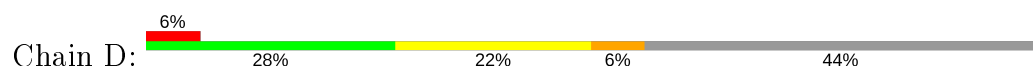
- Molecule 1: Anaerobic sulfatase-maturing enzyme



- Molecule 2: Kp18Cys peptide



- Molecule 2: Kp18Cys peptide







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	44.38Å 92.07Å 91.10Å 90.00° 91.21° 90.00°	Depositor
Resolution (Å)	41.09 – 1.83 41.08 – 1.83	Depositor EDS
% Data completeness (in resolution range)	99.2 (41.09-1.83) 99.1 (41.08-1.83)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.25 (at 1.83Å)	Xtriage
Refinement program	PHENIX 1.8.2_1309	Depositor
R, $R_{free}$	0.172 , 0.205 0.172 , 0.206	Depositor DCC
$R_{free}$ test set	3165 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	18.6	Xtriage
Anisotropy	0.725	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 53.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.012 for -h,-l,-k 0.000 for -h,l,k 0.032 for h,-k,-l	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	6764	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	25.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 23.54 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 4.5713e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, SF4, SAM, CL

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	0.60	1/3113 (0.0%)	0.66	1/4183 (0.0%)
1	B	0.69	1/3043 (0.0%)	0.74	8/4088 (0.2%)
2	C	0.74	0/72	0.70	0/97
2	D	0.50	0/69	0.57	0/93
All	All	0.65	2/6297 (0.0%)	0.70	9/8461 (0.1%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	3	PRO	N-CD	5.26	1.55	1.47
1	A	316	GLU	CD-OE1	-5.07	1.20	1.25

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	269[A]	SER	CA-C-O	5.76	132.19	120.10
1	B	269[B]	SER	CA-C-O	5.76	132.19	120.10
1	B	60[A]	SER	C-N-CA	5.56	135.60	121.70
1	B	60[B]	SER	C-N-CA	5.56	135.60	121.70
1	B	200	GLU	N-CA-C	-5.36	96.54	111.00
1	B	269[A]	SER	CA-C-N	-5.30	105.54	117.20
1	B	269[B]	SER	CA-C-N	-5.30	105.54	117.20
1	A	324	ARG	NE-CZ-NH2	5.21	122.91	120.30
1	B	362	LEU	CA-CB-CG	5.18	127.22	115.30

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	3037	0	2957	35	0
1	B	2971	0	2889	67	0
2	C	71	0	64	7	0
2	D	68	0	67	6	0
3	A	24	0	0	0	0
3	B	24	0	0	0	0
4	A	27	0	22	1	0
4	B	27	0	22	1	0
5	A	48	0	64	7	0
5	B	42	0	56	12	0
6	A	2	0	0	0	0
6	B	1	0	0	0	0
7	A	246	0	0	3	0
7	B	165	0	0	9	0
7	C	5	0	0	0	0
7	D	6	0	0	0	0
All	All	6764	0	6141	110	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 9.

All (110) 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:B:2:PRO:CB	1:B:3:PRO:HD3	1.83	1.07
1:B:2:PRO:CB	1:B:3:PRO:CD	2.39	0.99
1:B:145:LYS:HZ2	5:B:508:GOL:H2	1.31	0.96
1:B:145:LYS:NZ	5:B:508:GOL:H2	1.82	0.94
1:A:312:LYS:HE3	7:A:812:HOH:O	1.67	0.93
7:B:717:HOH:O	2:C:6:MET:HE1	1.75	0.86
1:A:106:GLU:OE2	1:A:156:TYR:HE2	1.58	0.85
1:B:180:TYR:CE1	1:B:184:LYS:HE2	2.13	0.83
1:B:132:LEU:O	1:B:132:LEU:HD12	1.80	0.81
1:A:163:LEU:HD12	1:A:192:ILE:HD11	1.63	0.80
1:A:2:PRO:N	1:A:4:LEU:HG	1.98	0.77

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:2:PRO:N	1:A:3:PRO:C	2.38	0.77
1:A:2:PRO:HA	1:A:56:ASN:H	1.49	0.76
1:B:305:GLU:HA	1:B:308:LYS:HG2	1.68	0.76
1:B:2:PRO:HA	1:B:56:ASN:HD22	1.51	0.76
1:A:161:ASN:HD22	1:A:190[B]:GLN:NE2	1.87	0.73
1:B:118:LEU:HD23	1:B:119:VAL:N	2.03	0.73
1:B:163:LEU:HD23	1:B:192:ILE:HD11	1.71	0.73
1:B:216:LYS:NZ	5:B:509:GOL:O1	2.22	0.72
1:A:161:ASN:HD22	1:A:190[B]:GLN:HE21	1.37	0.72
1:A:74:GLU:CD	1:A:74:GLU:H	1.92	0.72
1:A:106:GLU:OE2	1:A:156:TYR:CE2	2.42	0.71
1:A:238:ARG:NH2	2:D:4:SER:OG	2.25	0.69
1:B:247:ILE:HD12	1:B:247:ILE:C	2.14	0.68
1:A:74:GLU:OE1	1:A:74:GLU:N	2.26	0.67
7:B:715:HOH:O	2:C:6:MET:CE	2.43	0.66
1:B:302:LYS:HE3	7:B:736:HOH:O	1.95	0.65
7:B:715:HOH:O	2:C:6:MET:HE1	1.97	0.64
1:B:110:LYS:O	1:B:114:GLU:HG3	1.99	0.63
1:B:7:LEU:C	1:B:7:LEU:HD23	2.20	0.61
5:A:506:GOL:HO3	2:D:11:ARG:HH22	1.46	0.61
1:B:168:SER:O	1:B:172:ARG:HG3	2.00	0.61
1:B:234:ARG:HH11	1:B:234:ARG:HG3	1.66	0.61
5:A:506:GOL:O3	2:D:11:ARG:NH2	2.29	0.60
1:B:200:GLU:HA	1:B:200:GLU:OE1	2.01	0.60
1:A:234:ARG:HD2	5:A:510:GOL:O3	2.02	0.60
2:D:12:SER:O	2:D:12:SER:OG	2.14	0.60
1:A:190[A]:GLN:NE2	2:D:6:MET:O	2.36	0.59
1:B:234:ARG:HD2	5:B:510:GOL:H2	1.83	0.59
1:A:18:LYS:NZ	1:A:27:LEU:O	2.33	0.59
1:B:147:GLU:HG2	7:B:690:HOH:O	2.02	0.58
1:B:160:PHE:H	5:B:506:GOL:H31	1.68	0.58
1:B:3:PRO:HG2	1:B:58:HIS:CE1	2.39	0.58
1:B:234:ARG:NH1	1:B:234:ARG:HG3	2.19	0.57
1:A:90:ASN:ND2	5:A:508:GOL:O1	2.32	0.55
1:B:118:LEU:CD2	1:B:119:VAL:N	2.69	0.55
1:A:18:LYS:HD3	1:A:137:CYS:HA	1.90	0.54
1:A:234:ARG:HD2	5:A:510:GOL:C3	2.39	0.53
1:B:243:LEU:O	1:B:247:ILE:HG23	2.09	0.53
1:B:59:CYS:HG	1:B:61:PHE:HE1	1.56	0.53
1:B:118:LEU:HD23	1:B:119:VAL:H	1.71	0.52
7:A:699:HOH:O	5:B:508:GOL:H12	2.09	0.52

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:336:SER:OG	1:B:343:GLU:HG2	2.10	0.52
1:B:129:ILE:HD12	7:B:705:HOH:O	2.07	0.52
1:B:7:LEU:HD22	1:B:266:VAL:HG13	1.91	0.52
1:A:312:LYS:HE2	1:A:342:LEU:O	2.09	0.51
5:B:506:GOL:O3	2:C:11:ARG:NH1	2.32	0.51
1:B:238:ARG:NH1	2:C:4:SER:OG	2.43	0.51
1:B:132:LEU:HD12	1:B:132:LEU:C	2.29	0.51
1:B:190:GLN:HG2	1:B:238:ARG:HB2	1.93	0.51
1:B:234:ARG:HD2	5:B:510:GOL:C2	2.41	0.50
1:B:234:ARG:HD2	5:B:510:GOL:C3	2.41	0.50
1:B:343:GLU:HB2	7:B:735:HOH:O	2.11	0.50
1:A:98:GLN:OE1	1:A:163:LEU:HG	2.12	0.50
1:A:3:PRO:HB3	1:A:58:HIS:O	2.11	0.49
1:B:118:LEU:HD22	2:C:9:PRO:HB2	1.95	0.48
1:A:222:PHE:HE1	1:A:362:LEU:HD13	1.78	0.48
1:A:246:THR:HG23	1:A:324:ARG:NH1	2.29	0.48
1:A:9:LYS:HE2	1:A:268:GLU:OE2	2.14	0.48
1:B:45:GLU:HG2	1:B:49:LYS:HD2	1.96	0.47
1:B:180:TYR:CZ	1:B:184:LYS:HE2	2.50	0.47
1:B:228:ASP:HB3	1:B:233:ASN:O	2.14	0.47
1:B:9:LYS:NZ	1:B:64:GLN:OE1	2.48	0.46
1:B:204:LYS:HD3	1:B:205:TYR:CE1	2.50	0.46
1:B:18:LYS:HD3	1:B:137:CYS:HA	1.98	0.46
1:B:3:PRO:HB3	1:B:58:HIS:O	2.16	0.46
1:B:118:LEU:CD2	1:B:118:LEU:C	2.85	0.46
1:B:18:LYS:CG	1:B:137:CYS:HA	2.47	0.45
1:B:44:LEU:HD13	1:B:82:LEU:HD12	1.98	0.45
1:A:326:CYS:HA	5:A:505:GOL:H12	1.99	0.45
1:B:194:CYS:HG	1:B:207:TYR:HD1	1.61	0.45
1:B:187:LYS:NZ	1:B:233:ASN:OD1	2.50	0.44
1:B:243:LEU:HD13	1:B:362:LEU:HD21	1.99	0.44
1:B:212:LYS:HG2	7:B:620:HOH:O	2.17	0.44
1:B:67:GLU:O	1:B:70:LEU:HB2	2.18	0.44
1:A:266:VAL:O	1:A:273:VAL:HA	2.18	0.44
1:B:198:LEU:HG	1:B:199:TYR:CE2	2.53	0.44
1:A:234:ARG:HD2	5:A:510:GOL:H32	1.99	0.43
1:B:212:LYS:HD3	1:B:212:LYS:HA	1.77	0.43
1:B:180:TYR:HE1	1:B:184:LYS:HE2	1.79	0.43
1:A:212:LYS:HB3	1:A:212:LYS:HE2	1.74	0.43
1:A:283:LYS:HE3	1:A:283:LYS:HB2	1.69	0.43
1:B:2:PRO:HA	1:B:56:ASN:ND2	2.28	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:2:PRO:N	1:A:3:PRO:CA	2.81	0.42
1:B:40:ARG:HG3	7:B:609:HOH:O	2.19	0.42
1:A:206:ASN:HB2	7:A:810:HOH:O	2.19	0.42
1:A:64:GLN:HG2	4:A:504:SAM:HG2	2.00	0.42
1:B:160:PHE:H	5:B:506:GOL:C3	2.31	0.42
1:A:130:HIS:CE1	1:A:134:ARG:HD3	2.55	0.41
1:B:247:ILE:O	1:B:247:ILE:HD12	2.21	0.41
1:B:234:ARG:HD3	5:B:510:GOL:O3	2.20	0.41
1:B:238:ARG:NE	1:B:254:SER:OG	2.53	0.41
1:B:234:ARG:CD	5:B:510:GOL:O3	2.69	0.41
1:B:182:LYS:NZ	1:B:233:ASN:ND2	2.69	0.41
1:B:338:GLU:HG2	1:B:339:ASP:OD2	2.21	0.41
1:B:118:LEU:CD2	2:C:9:PRO:HB2	2.51	0.41
1:B:98:GLN:HG3	4:B:504:SAM:HB2	2.04	0.40
1:A:321:LYS:NZ	1:A:356:GLU:OE2	2.53	0.40
1:A:190[B]:GLN:HE22	2:D:9:PRO:HA	1.87	0.40
1:B:128:GLU:H	1:B:128:GLU:CD	2.25	0.40

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	370/370 (100%)	363 (98%)	7 (2%)	0	100	100
1	B	360/370 (97%)	353 (98%)	7 (2%)	0	100	100
2	C	9/18 (50%)	8 (89%)	1 (11%)	0	100	100
2	D	8/18 (44%)	7 (88%)	1 (12%)	0	100	100
All	All	747/776 (96%)	731 (98%)	16 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	336/339 (99%)	319 (95%)	17 (5%)	24	8
1	B	328/339 (97%)	301 (92%)	27 (8%)	11	2
2	C	7/15 (47%)	6 (86%)	1 (14%)	3	0
2	D	8/15 (53%)	7 (88%)	1 (12%)	4	0
All	All	679/708 (96%)	633 (93%)	46 (7%)	16	3

All (46) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	73	LEU
1	A	74	GLU
1	A	77	GLU
1	A	78	LYS
1	A	92	LYS
1	A	103	LEU
1	A	110	LYS
1	A	193	ASN
1	A	212	LYS
1	A	251	LYS
1	A	321	LYS
1	A	324	ARG
1	A	325	LEU
1	A	331	ARG
1	A	342	LEU
1	A	343	GLU
1	A	362	LEU
1	B	18	LYS
1	B	44	LEU
1	B	70	LEU
1	B	73	LEU
1	B	78	LYS
1	B	110	LYS
1	B	118	LEU

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	B	128	GLU
1	B	129	ILE
1	B	152	LEU
1	B	185	ASP
1	B	200	GLU
1	B	247	ILE
1	B	251	LYS
1	B	269[A]	SER
1	B	269[B]	SER
1	B	283	LYS
1	B	289	ILE
1	B	302	LYS
1	B	305	GLU
1	B	309	LEU
1	B	312	LYS
1	B	325	LEU
1	B	331	ARG
1	B	342	LEU
1	B	362	LEU
1	B	369	ILE
2	C	6	MET
2	D	12	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	56	ASN
1	B	133	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 carbohydrates in this entry.

## 5.6 Ligand geometry

Of 26 ligands modelled in this entry, 3 are monoatomic - leaving 23 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	GOL	A	508	-	5,5,5	0.34	0	5,5,5	0.71	0
5	GOL	A	511	-	5,5,5	0.43	0	5,5,5	0.60	0
3	SF4	A	501	1	0,12,12	0.00	-	-		
5	GOL	B	505	-	5,5,5	0.32	0	5,5,5	0.39	0
4	SAM	A	504	3	21,29,29	1.28	2 (9%)	18,42,42	1.66	2 (11%)
3	SF4	A	503	1,4	0,12,12	0.00	-	-		
5	GOL	A	507	-	5,5,5	0.34	0	5,5,5	0.43	0
5	GOL	B	507	-	5,5,5	0.33	0	5,5,5	0.46	0
5	GOL	A	512	-	5,5,5	0.28	0	5,5,5	0.42	0
5	GOL	A	509	-	5,5,5	0.22	0	5,5,5	0.24	0
5	GOL	B	506	-	5,5,5	0.38	0	5,5,5	0.37	0
3	SF4	A	502	1	0,12,12	0.00	-	-		
3	SF4	B	503	1,4	0,12,12	0.00	-	-		
3	SF4	B	502	1	0,12,12	0.00	-	-		
5	GOL	B	509	-	5,5,5	0.31	0	5,5,5	0.33	0
5	GOL	A	506	-	5,5,5	0.26	0	5,5,5	0.28	0
3	SF4	B	501	1	0,12,12	0.00	-	-		
5	GOL	B	508	-	5,5,5	0.98	0	5,5,5	1.38	0
5	GOL	A	510	-	5,5,5	0.39	0	5,5,5	0.37	0
5	GOL	B	510	-	5,5,5	0.24	0	5,5,5	0.42	0
5	GOL	A	505	-	5,5,5	0.75	0	5,5,5	1.07	0
5	GOL	B	511	-	5,5,5	0.30	0	5,5,5	0.25	0
4	SAM	B	504	3	21,29,29	1.20	2 (9%)	18,42,42	1.46	1 (5%)

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	GOL	A	508	-	-	2/4/4/4	-

*Continued on next page...*

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	GOL	A	511	-	-	4/4/4/4	-
3	SF4	A	501	1	-	-	0/6/5/5
5	GOL	B	505	-	-	0/4/4/4	-
4	SAM	A	504	3	-	0/8/33/33	0/3/3/3
3	SF4	A	503	1,4	-	-	0/6/5/5
5	GOL	A	507	-	-	1/4/4/4	-
5	GOL	B	507	-	-	2/4/4/4	-
5	GOL	A	512	-	-	0/4/4/4	-
5	GOL	A	509	-	-	0/4/4/4	-
5	GOL	B	506	-	-	2/4/4/4	-
3	SF4	A	502	1	-	-	0/6/5/5
3	SF4	B	503	1,4	-	-	0/6/5/5
3	SF4	B	502	1	-	-	0/6/5/5
5	GOL	B	509	-	-	1/4/4/4	-
5	GOL	A	506	-	-	2/4/4/4	-
3	SF4	B	501	1	-	-	0/6/5/5
5	GOL	B	508	-	-	2/4/4/4	-
5	GOL	A	510	-	-	2/4/4/4	-
5	GOL	B	510	-	-	0/4/4/4	-
5	GOL	A	505	-	-	3/4/4/4	-
5	GOL	B	511	-	-	2/4/4/4	-
4	SAM	B	504	3	-	1/8/33/33	0/3/3/3

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	504	SAM	C2-N3	4.63	1.39	1.32
4	B	504	SAM	C2-N3	4.18	1.38	1.32
4	A	504	SAM	C2-N1	2.57	1.38	1.33
4	B	504	SAM	C2-N1	2.25	1.38	1.33

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	504	SAM	N3-C2-N1	-5.16	120.61	128.68
4	A	504	SAM	N3-C2-N1	-5.14	120.65	128.68
4	A	504	SAM	C4-C5-N7	-2.00	107.31	109.40

There are no chirality outliers.

All (24) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	511	GOL	O1-C1-C2-C3
5	A	511	GOL	C1-C2-C3-O3
5	A	511	GOL	O2-C2-C3-O3
5	B	507	GOL	O1-C1-C2-C3
5	B	506	GOL	C1-C2-C3-O3
5	A	506	GOL	C1-C2-C3-O3
5	A	506	GOL	O2-C2-C3-O3
5	B	508	GOL	C1-C2-C3-O3
5	B	508	GOL	O2-C2-C3-O3
4	B	504	SAM	C3'-C4'-C5'-SD
5	A	508	GOL	C1-C2-C3-O3
5	B	509	GOL	O1-C1-C2-C3
5	A	505	GOL	O1-C1-C2-C3
5	A	511	GOL	O1-C1-C2-O2
5	B	507	GOL	O1-C1-C2-O2
5	A	508	GOL	O2-C2-C3-O3
5	B	511	GOL	O1-C1-C2-O2
5	A	510	GOL	O1-C1-C2-C3
5	B	506	GOL	O2-C2-C3-O3
5	A	510	GOL	O1-C1-C2-O2
5	A	505	GOL	O1-C1-C2-O2
5	A	507	GOL	O1-C1-C2-C3
5	A	505	GOL	C1-C2-C3-O3
5	B	511	GOL	O1-C1-C2-C3

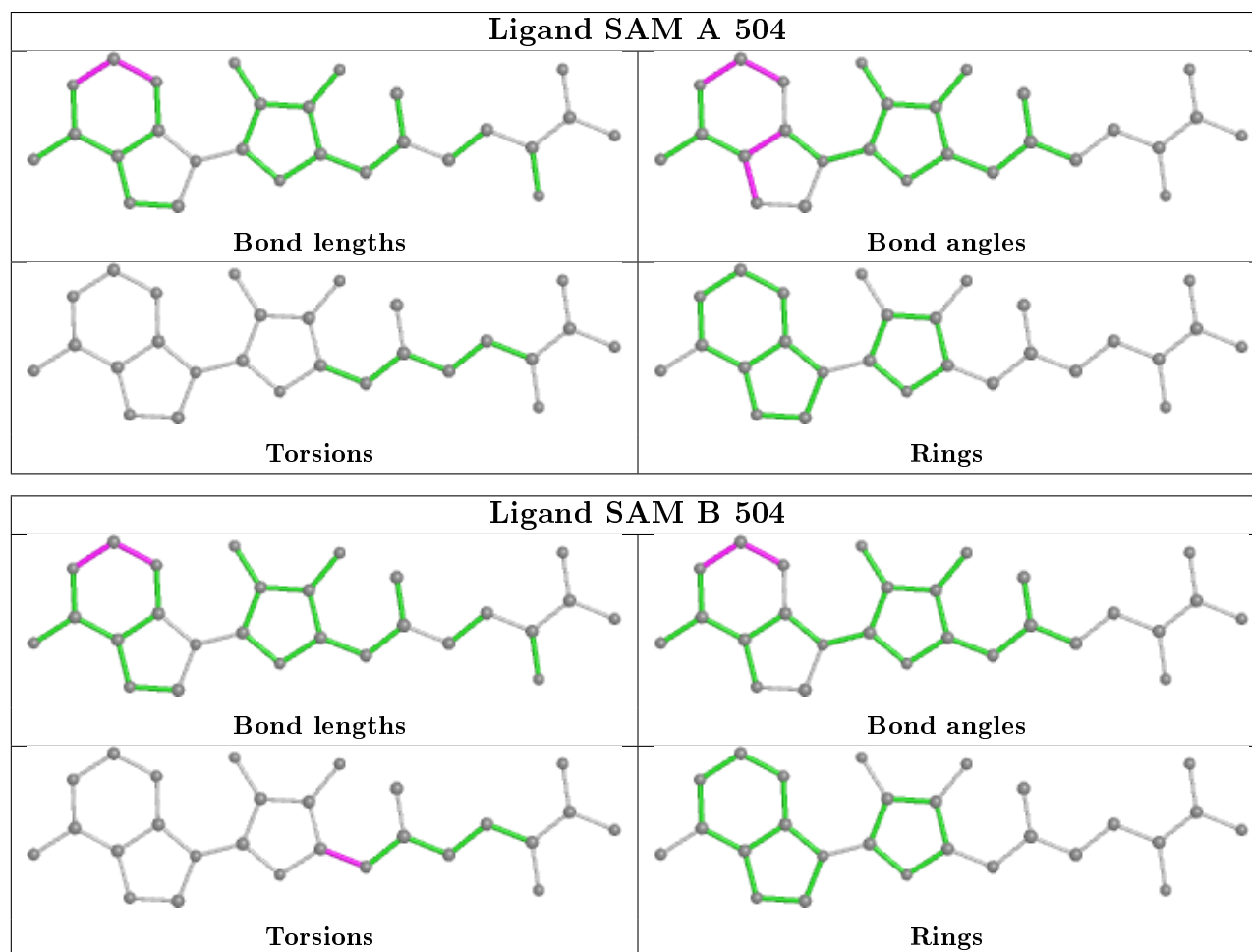
There are no ring outliers.

10 monomers are involved in 21 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	508	GOL	1	0
4	A	504	SAM	1	0
5	B	506	GOL	3	0
5	B	509	GOL	1	0
5	A	506	GOL	2	0
5	B	508	GOL	3	0
5	A	510	GOL	3	0
5	B	510	GOL	5	0
5	A	505	GOL	1	0
4	B	504	SAM	1	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will

also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.



## 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	369/370 (99%)	-0.14	5 (1%) 75 75	13, 20, 36, 49	0
1	B	362/370 (97%)	0.01	16 (4%) 34 31	16, 25, 43, 67	0
2	C	11/18 (61%)	1.50	3 (27%) 0 0	29, 38, 45, 50	0
2	D	10/18 (55%)	1.01	1 (10%) 7 6	28, 35, 41, 45	0
All	All	752/776 (96%)	-0.03	25 (3%) 46 43	13, 23, 41, 67	0

All (25) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	2	PRO	6.2
1	B	23	PHE	5.3
1	A	2	PRO	4.8
1	B	24	TYR	4.7
1	B	339	ASP	4.6
1	B	33	VAL	4.4
2	C	3	THR	3.7
1	B	340	SER	3.7
1	B	341	ALA	3.6
2	C	13	MET	3.2
1	A	29	ASP	3.2
1	B	201	GLU	3.1
1	B	203	GLY	3.0
1	B	200	GLU	3.0
1	A	324	ARG	2.8
2	C	5	PRO	2.7
1	B	204	LYS	2.5
1	B	34	LYS	2.4
1	B	26	SER	2.3
1	B	234	ARG	2.3
1	A	369	ILE	2.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	D	3	THR	2.2
1	B	61	PHE	2.2
1	B	262	THR	2.1
1	A	251	LYS	2.0

## 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 carbohydrates 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.

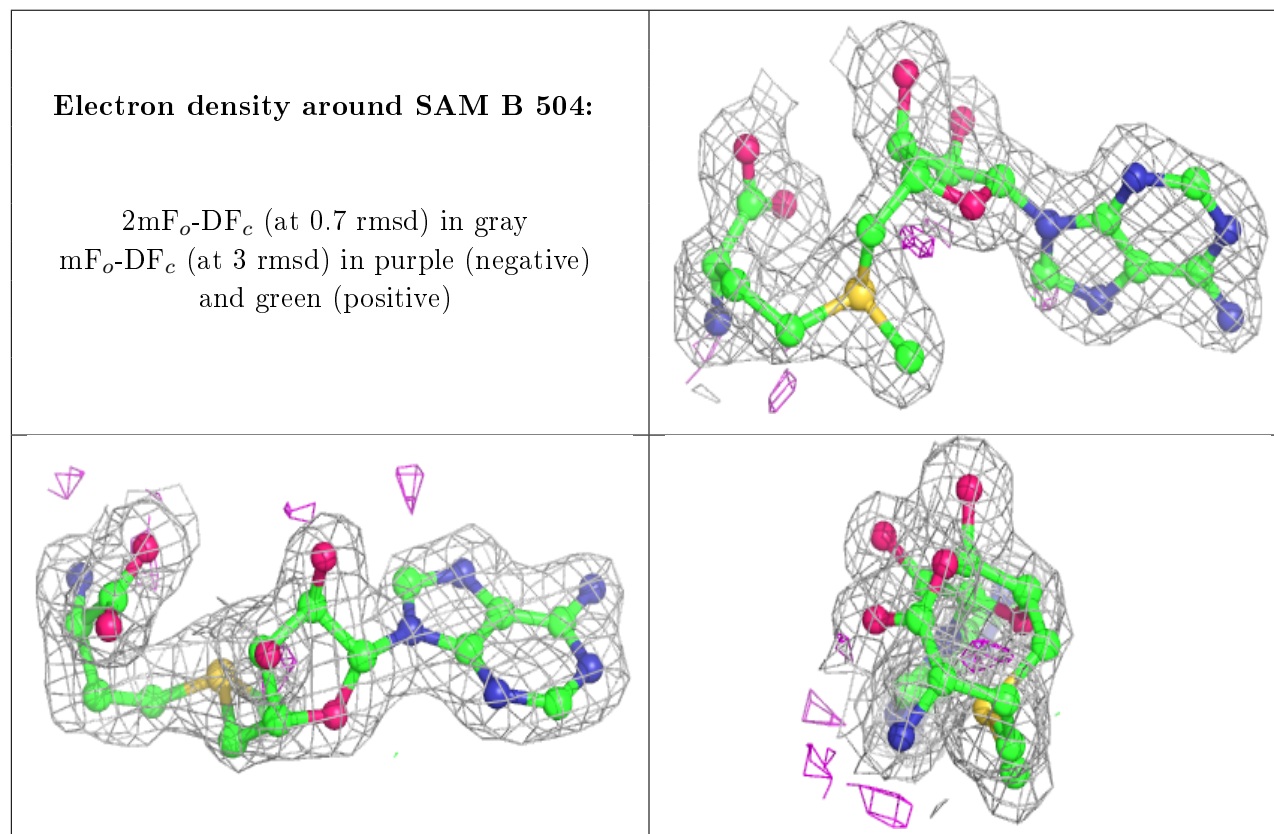
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
5	GOL	A	506	6/6	0.78	0.21	28,31,34,34	0
5	GOL	B	511	6/6	0.79	0.26	33,38,45,54	0
5	GOL	A	510	6/6	0.82	0.24	33,37,41,46	0
5	GOL	B	506	6/6	0.86	0.19	28,33,34,36	0
5	GOL	A	508	6/6	0.90	0.13	25,29,34,34	0
5	GOL	B	510	6/6	0.90	0.20	37,39,40,44	0
5	GOL	A	511	6/6	0.90	0.25	22,29,30,36	0
5	GOL	B	508	6/6	0.91	0.14	35,36,38,38	0
5	GOL	B	509	6/6	0.92	0.26	27,31,33,34	0
5	GOL	B	507	6/6	0.92	0.14	21,28,30,30	0
5	GOL	A	512	6/6	0.93	0.12	33,38,40,41	0
5	GOL	A	505	6/6	0.94	0.13	21,25,27,29	0
5	GOL	B	505	6/6	0.95	0.11	22,24,24,26	0
4	SAM	B	504	27/27	0.95	0.09	21,25,32,32	0
5	GOL	A	509	6/6	0.96	0.08	16,18,19,20	0
5	GOL	A	507	6/6	0.96	0.10	23,24,26,29	0
3	SF4	B	503	8/8	0.97	0.09	19,23,26,27	0
6	CL	A	513	1/1	0.97	0.07	26,26,26,26	0
4	SAM	A	504	27/27	0.97	0.09	13,16,18,20	0

*Continued on next page...*

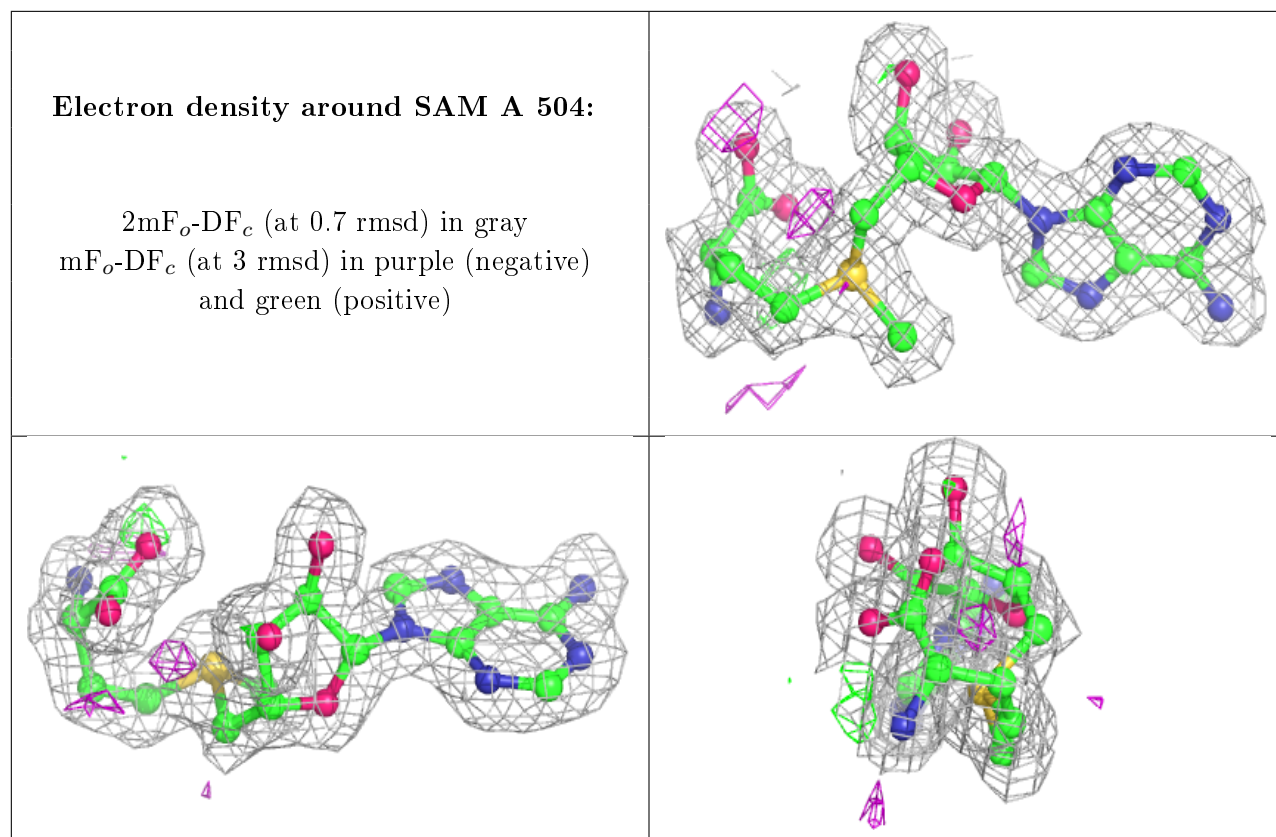
*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	SF4	A	502	8/8	0.99	0.06	19,20,21,22	0
3	SF4	A	503	8/8	0.99	0.09	15,15,15,16	0
3	SF4	B	502	8/8	0.99	0.07	19,20,21,24	0
6	CL	B	512	1/1	0.99	0.06	23,23,23,23	0
3	SF4	A	501	8/8	0.99	0.07	16,17,17,17	0
6	CL	A	514	1/1	0.99	0.04	23,23,23,23	0
3	SF4	B	501	8/8	0.99	0.06	19,19,21,22	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.







## 6.5 Other polymers [i](#)

There are no such residues in this entry.