



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 09:11 AM GMT

PDB ID : 3HHL  
Title : Crystal structure of methylated RPA0582 protein  
Authors : Sledz, P.; Niedzialkowska, E.; Chruszcz, M.; Porebski, P.; Yim, V.; Kudritska, M.; Zimmerman, M.D.; Evdokimova, E.; Savchenko, A.; Edwards, A.; Joachimiak, A.; Minor, W.; Midwest Center for Structural Genomics (MCSG)  
Deposited on : 2009-05-15  
Resolution : 2.65 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  
<http://wwpdb.org/validation/2016/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 : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

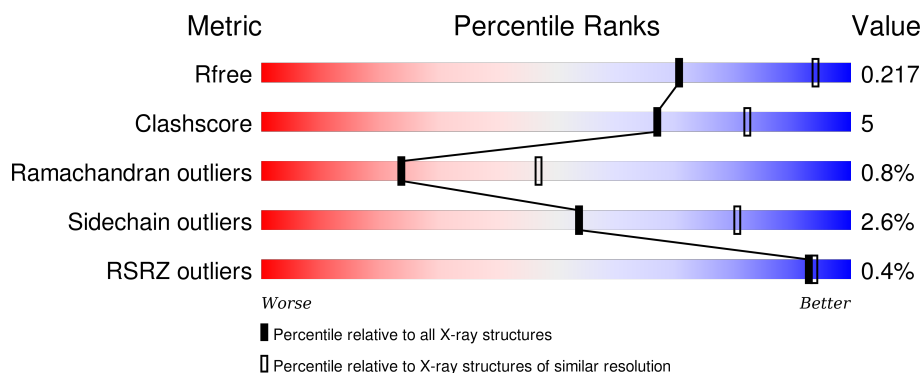
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.65 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))
$R_{free}$	91344	3152 (2.70-2.62)
Clashscore	102246	3524 (2.70-2.62)
Ramachandran outliers	100387	3469 (2.70-2.62)
Sidechain outliers	100360	3469 (2.70-2.62)
RSRZ outliers	91569	3161 (2.70-2.62)

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. 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	143	<div> <div>85%</div> <div>8%</div> <div>6%</div> </div>
2	B	143	<div> <div>78%</div> <div>16%</div> <div>6%</div> </div>
3	C	143	<div> <div>83%</div> <div>11%</div> <div>6%</div> </div>
4	D	143	<div> <div>81%</div> <div>13%</div> <div>6%</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
13	PG4	C	145	-	-	-	X
13	PG4	D	144	-	-	-	X
5	SO4	A	144	-	-	-	X
9	2PE	A	148	-	-	-	X
9	2PE	D	146	-	-	-	X

## 2 Entry composition [i](#)

There are 14 unique types of molecules in this entry. The entry contains 4944 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 RPA0582.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	135	Total	C	N	O	Se	0	2	0
			1098	696	203	195	4			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	GLY	-	leader sequence	UNP Q6NC90
A	61	SER	PHE	SEE REMARK 999	UNP Q6NC90

- Molecule 2 is a protein called RPA0582.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	135	Total	C	N	O	Se	0	3	0
			1103	698	208	194	3			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	0	GLY	-	leader sequence	UNP Q6NC90
B	61	SER	PHE	SEE REMARK 999	UNP Q6NC90

- Molecule 3 is a protein called RPA0582.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	135	Total	C	N	O	Se	0	5	0
			1119	710	210	195	4			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	0	GLY	-	leader sequence	UNP Q6NC90
C	61	SER	PHE	SEE REMARK 999	UNP Q6NC90

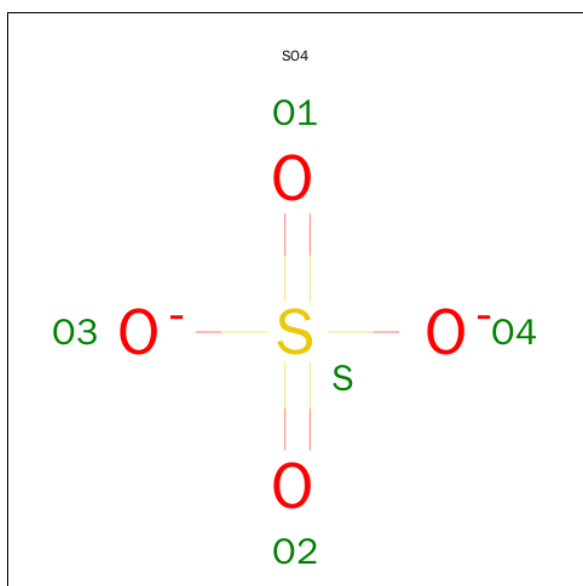
- Molecule 4 is a protein called RPA0582.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	135	Total	C	N	O	Se	0	3	0
			1103	697	205	198	3			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	0	GLY	-	leader sequence	UNP Q6NC90
D	61	SER	PHE	SEE REMARK 999	UNP Q6NC90

- Molecule 5 is SULFATE ION (three-letter code: SO<sub>4</sub>) (formula: O<sub>4</sub>S).



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

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

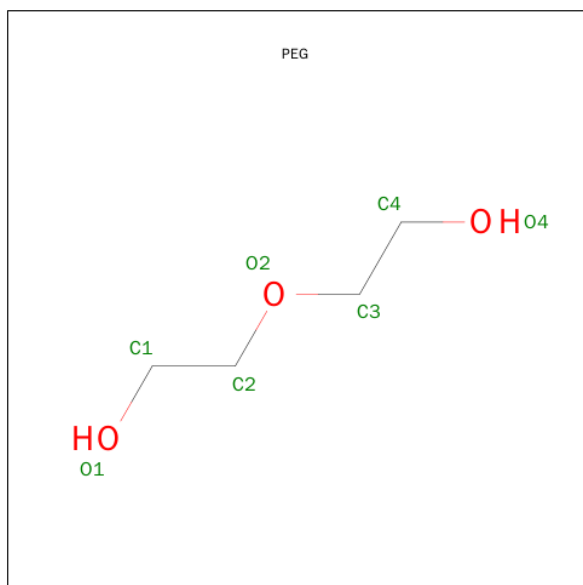
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	B	1	Total	Cl	0	0
			1	1		

*Continued on next page...*

Continued from previous page...

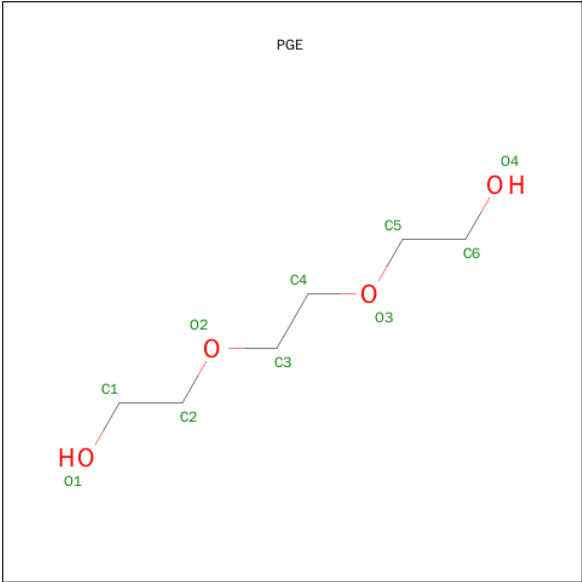
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	1	Total	Cl	0	0
			1	1		
6	D	1	Total	Cl	0	0
			1	1		
6	C	1	Total	Cl	0	0
			1	1		

- Molecule 7 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula:  $C_4H_{10}O_3$ ).



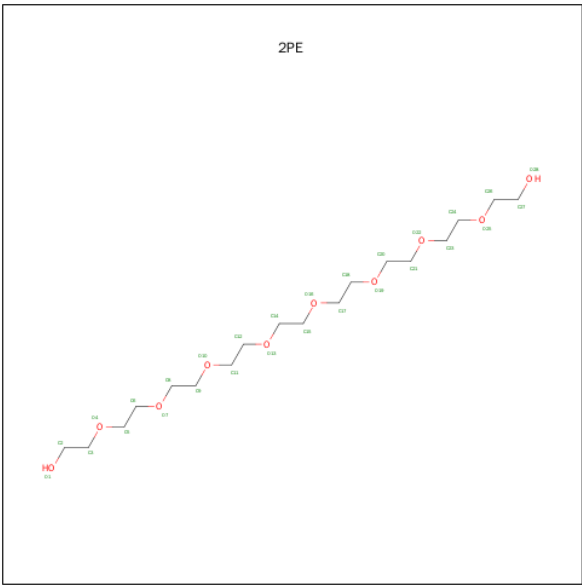
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total	C	O	0	0
			7	4	3		
7	A	1	Total	C	O	0	0
			7	4	3		
7	B	1	Total	C	O	0	0
			7	4	3		
7	B	1	Total	C	O	0	0
			7	4	3		
7	B	1	Total	C	O	0	0
			7	4	3		
7	C	1	Total	C	O	0	0
			7	4	3		
7	D	1	Total	C	O	0	0
			7	4	3		

- Molecule 8 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula:  $C_6H_{14}O_4$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	A	1	Total	C	O	0	0
			10	6	4		
8	B	1	Total	C	O	0	0
			10	6	4		
8	C	1	Total	C	O	0	0
			10	6	4		

- Molecule 9 is NONAETHYLENE GLYCOL (three-letter code: 2PE) (formula: C<sub>18</sub>H<sub>38</sub>O<sub>10</sub>).



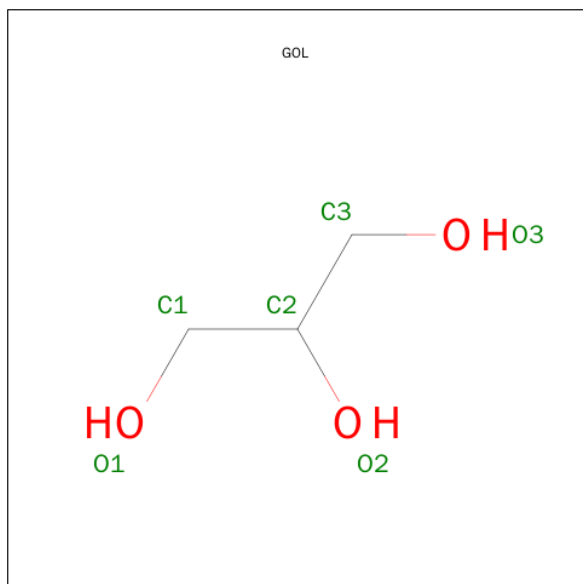
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
9	A	1	Total	C	O	0	0
			28	18	10		

Continued on next page...

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
9	D	1	Total	C	O	0	0
			28	18	10		

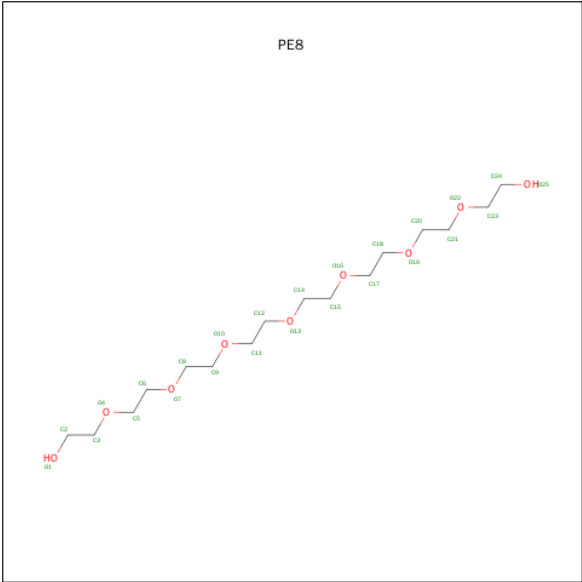
- Molecule 10 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



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

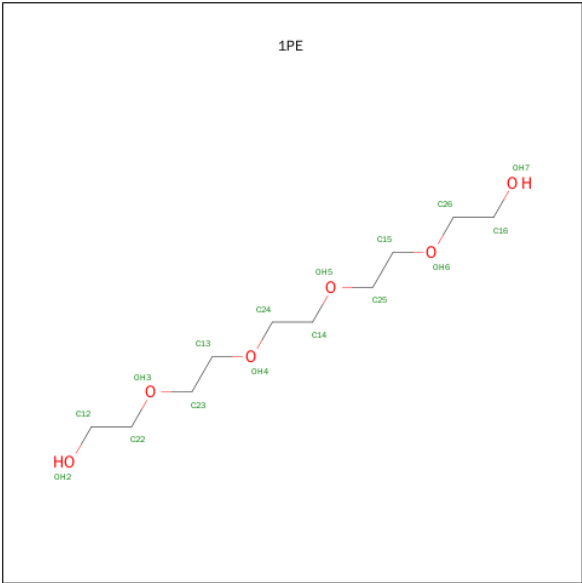
- Molecule 11 is 3,6,9,12,15,18,21-HEPTAOXATRICOSANE-1,23-DIOL (three-letter code: PE8) (formula:  $C_{16}H_{34}O_9$ ).





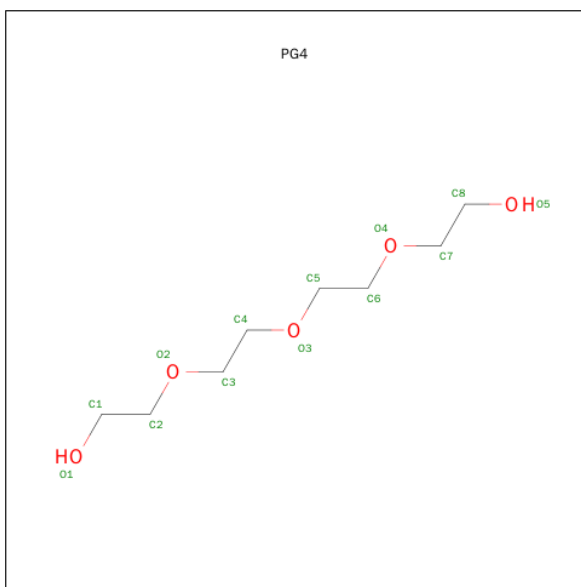
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
11	B	1	Total	C	O	0	0
			25	16	9		
11	C	1	Total	C	O	0	0
			25	16	9		

- Molecule 12 is PENTAETHYLENE GLYCOL (three-letter code: 1PE) (formula: C<sub>10</sub>H<sub>22</sub>O<sub>6</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
12	B	1	Total	C	O	0	0
			16	10	6		

- Molecule 13 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula: C<sub>8</sub>H<sub>18</sub>O<sub>5</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
13	C	1	Total	C	O	0	0
			13	8	5		
13	D	1	Total	C	O	0	0
			13	8	5		

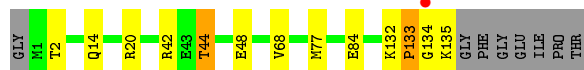
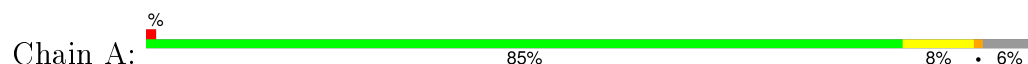
- Molecule 14 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
14	A	57	Total	O	0	0
			57	57		
14	B	66	Total	O	0	0
			66	66		
14	C	70	Total	O	0	0
			70	70		
14	D	58	Total	O	0	0
			58	58		

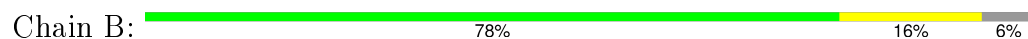
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

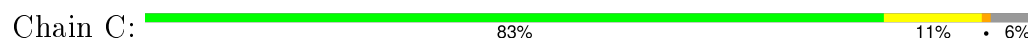
#### • Molecule 1: RPA0582



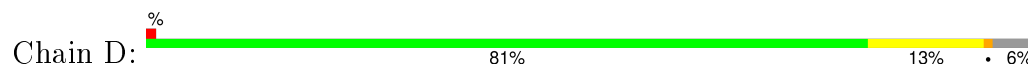
#### • Molecule 2: RPA0582



#### • Molecule 3: RPA0582



#### • Molecule 4: RPA0582



## 4 Data and refinement statistics

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	193.01Å 193.01Å 117.97Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	32.17 – 2.65 32.17 – 2.65	Depositor EDS
% Data completeness (in resolution range)	99.6 (32.17-2.65) 99.2 (32.17-2.65)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	0.09	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.31 (at 2.64Å)	Xtriage
Refinement program	REFMAC 5.5.0089	Depositor
R, $R_{free}$	0.178 , 0.206 0.190 , 0.217	Depositor DCC
$R_{free}$ test set	2400 reflections (5.33%)	DCC
Wilson B-factor (Å <sup>2</sup> )	53.8	Xtriage
Anisotropy	0.011	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 24.8	EDS
Estimated twinning fraction	0.015 for -2/3*h-1/3*k-4/3*l,-1/3*h-2/3*k+4/3*l,-1/3*h+1/3*k+1/3*l 0.012 for -h,1/3*h-1/3*k-4/3*l,-1/3*h-2/3*k+1/3*l 0.014 for -1/3*h+1/3*k+4/3*l,-k,2/3*h+1/3*k+1/3*l 0.470 for -h,2/3*h+1/3*k+4/3*l,1/3*h+2/3*k-1/3*l 0.477 for -1/3*h-2/3*k+4/3*l,-2/3*h-1/3*k-4/3*l,1/3*h-1/3*k-1/3*l 0.470 for 1/3*h+2/3*k-4/3*l,-k,-2/3*h-1/3*k-1/3*l 0.016 for h,-h-k,-l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.35$	Xtriage
Outliers	0 of 47441 reflections	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	4944	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	42.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.73% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: GOL, PGE, PE8, CL, MLY, 1PE, PG4, 2PE, SO4, MLZ, PEG

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\# Z  > 5$	RMSZ	$\# Z  > 5$
1	A	0.82	0/1090	0.78	0/1471
2	B	0.83	1/1112 (0.1%)	0.77	0/1500
3	C	0.78	0/1109	0.79	0/1494
4	D	0.84	2/1105 (0.2%)	0.82	1/1494 (0.1%)
All	All	0.82	3/4416 (0.1%)	0.79	1/5959 (0.0%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	10	GLU	CG-CD	6.83	1.62	1.51
4	D	10[A]	GLU	CG-CD	5.35	1.59	1.51
4	D	10[B]	GLU	CG-CD	5.35	1.59	1.51

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	D	33	ARG	NE-CZ-NH1	5.64	123.12	120.30

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	1098	0	1076	9	0
2	B	1103	0	1089	20	0
3	C	1119	0	1111	9	0
4	D	1103	0	1071	11	0
5	A	10	0	0	0	0
5	C	5	0	0	0	0
6	A	1	0	0	0	0
6	B	1	0	0	1	0
6	C	1	0	0	1	0
6	D	1	0	0	0	0
7	A	14	0	20	0	0
7	B	21	0	30	0	0
7	C	7	0	10	0	0
7	D	7	0	10	0	0
8	A	10	0	14	0	0
8	B	10	0	14	0	0
8	C	10	0	14	0	0
9	A	28	0	38	1	0
9	D	28	0	38	1	0
10	A	6	0	8	0	0
10	B	6	0	8	0	0
10	C	6	0	8	0	0
10	D	6	0	8	0	0
11	B	25	0	34	2	0
11	C	25	0	34	1	0
12	B	16	0	22	0	0
13	C	13	0	18	0	0
13	D	13	0	18	0	0
14	A	57	0	0	1	0
14	B	66	0	0	0	0
14	C	70	0	0	1	0
14	D	58	0	0	1	0
All	All	4944	0	4693	45	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 5.

All (45) 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)
2:B:16[B]:ARG:HG2	2:B:16[B]:ARG:HH11	1.09	1.15
4:D:16[B]:ARG:HG2	4:D:16[B]:ARG:HH11	1.20	1.02

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:16[B]:ARG:HG2	2:B:16[B]:ARG:NH1	1.90	0.85
4:D:16[B]:ARG:NH1	4:D:16[B]:ARG:HG2	1.95	0.76
2:B:16[B]:ARG:HH11	2:B:16[B]:ARG:CG	1.95	0.75
2:B:1:MSE:O	2:B:2:THR:HG22	1.89	0.73
1:A:44:THR:HG22	1:A:48:GLU:OE1	1.92	0.69
2:B:100:PHE:CE2	2:B:104:ILE:HD12	2.32	0.65
1:A:77:MSE:HE3	1:A:84:GLU:HB3	1.80	0.64
4:D:33:ARG:HG3	4:D:35:ARG:O	1.98	0.63
2:B:1:MSE:O	2:B:2:THR:CG2	2.47	0.62
2:B:68:VAL:HG23	11:B:148:PE8:H231	1.82	0.60
6:C:144:CL:CL	11:C:148:PE8:H172	2.40	0.59
4:D:55[A]:ARG:NH1	4:D:56:ASP:OD1	2.37	0.58
4:D:77:MSE:HE3	4:D:84:GLU:HB3	1.85	0.58
2:B:132:LYS:O	2:B:134:GLY:N	2.35	0.57
3:C:77:MSE:HE3	3:C:84:GLU:HB3	1.85	0.57
3:C:5:ILE:O	3:C:116:ARG:NH2	2.37	0.56
2:B:131:LEU:HD12	3:C:71:GLN:HG2	1.91	0.53
4:D:68:VAL:HG23	9:D:146:2PE:H262	1.91	0.50
2:B:135:MLY:HH11	3:C:90:PHE:CD1	2.46	0.50
2:B:77:MSE:HE3	2:B:84:GLU:HB3	1.94	0.50
3:C:1:MSE:N	14:C:233:HOH:O	2.46	0.49
3:C:55[A]:ARG:NH1	3:C:56:ASP:OD1	2.46	0.47
1:A:68:VAL:HG23	9:A:148:2PE:H261	1.96	0.47
2:B:16[B]:ARG:NH1	2:B:16[B]:ARG:CG	2.64	0.47
2:B:135:MLY:HH12	2:B:135:MLY:HD2	1.60	0.47
1:A:132:LYS:O	1:A:134:GLY:N	2.43	0.46
1:A:2:THR:HG22	14:A:238:HOH:O	2.14	0.46
4:D:134:GLY:O	4:D:135:MLY:C	2.65	0.45
1:A:2:THR:HG23	1:A:2:THR:O	2.17	0.45
2:B:90:PHE:CD1	3:C:135:MLY:HH11	2.52	0.45
2:B:5:ILE:O	2:B:116:ARG:NH2	2.50	0.45
3:C:132:LYS:O	3:C:134:GLY:N	2.43	0.44
6:B:143:CL:CL	11:B:148:PE8:H171	2.54	0.44
2:B:135:MLY:HE3	2:B:135:MLY:HB2	1.57	0.43
3:C:38:TYR:CD1	3:C:42:ARG:HD2	2.54	0.43
2:B:40:ASP:CG	2:B:40:ASP:O	2.57	0.43
1:A:133:PRO:HB3	4:D:68:VAL:HG12	2.00	0.42
4:D:29:LEU:HD23	4:D:89:VAL:HG13	2.01	0.42
2:B:100:PHE:CE2	2:B:104:ILE:CD1	3.02	0.42
1:A:135:MLY:HD2	4:D:71:GLN:HE22	1.86	0.41
1:A:14:GLN:NE2	2:B:63:ARG:HG2	2.36	0.41

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:D:1:MSE:N	14:D:189:HOH:O	2.54	0.41
2:B:35:ARG:HD2	2:B:43:GLU:OE2	2.20	0.41

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	133/143 (93%)	129 (97%)	3 (2%)	1 (1%)	24	47
2	B	135/143 (94%)	131 (97%)	3 (2%)	1 (1%)	26	51
3	C	135/143 (94%)	130 (96%)	4 (3%)	1 (1%)	26	51
4	D	135/143 (94%)	130 (96%)	4 (3%)	1 (1%)	26	51
All	All	538/572 (94%)	520 (97%)	14 (3%)	4 (1%)	24	51

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	133	PRO
3	C	133	PRO
4	D	133	PRO
1	A	133	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	107/108 (99%)	104 (97%)	3 (3%)	51	79
2	B	109/109 (100%)	107 (98%)	2 (2%)	66	88
3	C	109/107 (102%)	106 (97%)	3 (3%)	51	79
4	D	108/109 (99%)	105 (97%)	3 (3%)	51	79
All	All	433/433 (100%)	422 (98%)	11 (2%)	54	82

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	20	ARG
1	A	42	ARG
1	A	44	THR
2	B	20	ARG
2	B	44	THR
3	C	20	ARG
3	C	42	ARG
3	C	104	ILE
4	D	2	THR
4	D	20	ARG
4	D	40	ASP

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

Mol	Chain	Res	Type
1	A	4	HIS
4	D	4	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](#)

11 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected

value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# $ Z  > 2$	Counts	RMSZ	# $ Z  > 2$
1	MLZ	A	129	1	8,9,10	1.94	2 (25%)	7,9,11	1.13	0
1	MLY	A	135	1	9,10,11	1.37	1 (11%)	9,11,13	0.74	0
1	MLY	A	67	1	9,10,11	0.62	0	9,11,13	0.69	0
2	MLY	B	135	2	9,10,11	0.49	0	9,11,13	0.69	0
2	MLZ	B	67	2	8,9,10	1.68	1 (12%)	7,9,11	0.87	0
3	MLZ	C	114	3	8,9,10	1.09	1 (12%)	7,9,11	0.80	0
3	MLZ	C	129	3	8,9,10	1.07	0	7,9,11	0.98	1 (14%)
3	MLY	C	135	3	9,10,11	1.29	1 (11%)	9,11,13	0.90	0
3	MLY	C	67	3	9,10,11	0.68	0	9,11,13	0.61	0
4	MLY	D	135	4	9,10,11	0.93	0	9,11,13	0.66	0
4	MLY	D	67	4	9,10,11	0.70	0	9,11,13	1.05	1 (11%)

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
1	MLZ	A	129	1	-	0/6/8/10	0/0/0/0
1	MLY	A	135	1	-	0/7/9/11	0/0/0/0
1	MLY	A	67	1	-	0/7/9/11	0/0/0/0
2	MLY	B	135	2	-	0/7/9/11	0/0/0/0
2	MLZ	B	67	2	-	0/6/8/10	0/0/0/0
3	MLZ	C	114	3	-	0/6/8/10	0/0/0/0
3	MLZ	C	129	3	-	0/6/8/10	0/0/0/0
3	MLY	C	135	3	-	0/7/9/11	0/0/0/0
3	MLY	C	67	3	-	0/7/9/11	0/0/0/0
4	MLY	D	135	4	-	0/7/9/11	0/0/0/0
4	MLY	D	67	4	-	0/7/9/11	0/0/0/0

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	67	MLZ	CB-CA	-4.05	1.49	1.53
1	A	129	MLZ	CB-CA	-3.90	1.50	1.53
1	A	129	MLZ	CD-CE	-2.63	1.40	1.51
3	C	114	MLZ	CB-CA	-2.23	1.51	1.53
3	C	135	MLY	CB-CA	3.28	1.56	1.53

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	135	MLY	CB-CA	3.46	1.57	1.53

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	129	MLZ	O-C-CA	-2.17	119.85	125.49
4	D	67	MLY	CD-CE-NZ	2.41	120.12	113.92

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	135	MLY	1	0
2	B	135	MLY	3	0
3	C	135	MLY	1	0
4	D	135	MLY	1	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 28 ligands modelled in this entry, 4 are monoatomic - leaving 24 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	SO4	A	143	-	4,4,4	0.77	0	6,6,6	0.36	0
5	SO4	A	144	-	4,4,4	0.09	0	6,6,6	0.16	0
7	PEG	A	146	-	6,6,6	0.46	0	5,5,5	0.23	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
8	PGE	A	147	-	9,9,9	0.66	0	8,8,8	0.37	0
9	2PE	A	148	-	27,27,27	0.66	0	26,26,26	0.47	0
7	PEG	A	149	-	6,6,6	0.40	0	5,5,5	0.43	0
10	GOL	A	150	-	5,5,5	0.27	0	5,5,5	0.18	0
7	PEG	B	144	-	6,6,6	0.54	0	5,5,5	0.20	0
7	PEG	B	145	-	6,6,6	0.47	0	5,5,5	0.42	0
7	PEG	B	146	-	6,6,6	0.40	0	5,5,5	0.43	0
8	PGE	B	147	-	9,9,9	0.59	0	8,8,8	0.69	0
11	PE8	B	148	-	24,24,24	0.71	0	23,23,23	0.54	0
10	GOL	B	149	-	5,5,5	0.45	0	5,5,5	0.72	0
12	1PE	B	150	-	15,15,15	0.60	0	14,14,14	0.39	0
5	SO4	C	143	-	4,4,4	0.42	0	6,6,6	0.14	0
13	PG4	C	145	-	12,12,12	0.60	0	11,11,11	0.32	0
7	PEG	C	146	-	6,6,6	0.49	0	5,5,5	0.20	0
8	PGE	C	147	-	9,9,9	0.61	0	8,8,8	0.50	0
11	PE8	C	148	-	24,24,24	0.76	0	23,23,23	0.73	1 (4%)
10	GOL	C	149	-	5,5,5	0.34	0	5,5,5	0.38	0
13	PG4	D	144	-	12,12,12	0.70	0	11,11,11	0.54	0
7	PEG	D	145	-	6,6,6	0.50	0	5,5,5	0.23	0
9	2PE	D	146	-	27,27,27	0.62	0	26,26,26	0.43	0
10	GOL	D	147	-	5,5,5	0.27	0	5,5,5	0.34	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	SO4	A	143	-	-	0/0/0/0	0/0/0/0
5	SO4	A	144	-	-	0/0/0/0	0/0/0/0
7	PEG	A	146	-	-	0/4/4/4	0/0/0/0
8	PGE	A	147	-	-	0/7/7/7	0/0/0/0
9	2PE	A	148	-	-	0/25/25/25	0/0/0/0
7	PEG	A	149	-	-	0/4/4/4	0/0/0/0
10	GOL	A	150	-	-	0/4/4/4	0/0/0/0
7	PEG	B	144	-	-	0/4/4/4	0/0/0/0
7	PEG	B	145	-	-	0/4/4/4	0/0/0/0
7	PEG	B	146	-	-	0/4/4/4	0/0/0/0
8	PGE	B	147	-	-	0/7/7/7	0/0/0/0
11	PE8	B	148	-	-	0/22/22/22	0/0/0/0
10	GOL	B	149	-	-	0/4/4/4	0/0/0/0
12	1PE	B	150	-	-	0/13/13/13	0/0/0/0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	SO4	C	143	-	-	0/0/0/0	0/0/0/0
13	PG4	C	145	-	-	0/10/10/10	0/0/0/0
7	PEG	C	146	-	-	0/4/4/4	0/0/0/0
8	PGE	C	147	-	-	0/7/7/7	0/0/0/0
11	PE8	C	148	-	-	0/22/22/22	0/0/0/0
10	GOL	C	149	-	-	0/4/4/4	0/0/0/0
13	PG4	D	144	-	-	0/10/10/10	0/0/0/0
7	PEG	D	145	-	-	0/4/4/4	0/0/0/0
9	2PE	D	146	-	-	0/25/25/25	0/0/0/0
10	GOL	D	147	-	-	0/4/4/4	0/0/0/0

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
11	C	148	PE8	O19-C18-C17	2.21	120.21	110.36

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
9	A	148	2PE	1	0
11	B	148	PE8	2	0
11	C	148	PE8	1	0
9	D	146	2PE	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, 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	128/143 (89%)	-0.06	1 (0%) 87 87	26, 37, 55, 79	0
2	B	129/143 (90%)	-0.11	0 100 100	26, 37, 55, 79	0
3	C	127/143 (88%)	-0.07	0 100 100	25, 37, 53, 77	0
4	D	129/143 (90%)	-0.07	1 (0%) 87 87	26, 37, 55, 79	0
All	All	513/572 (89%)	-0.08	2 (0%) 93 94	25, 37, 55, 79	0

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	134	GLY	5.2
4	D	134	GLY	2.3

### 6.2 Non-standard residues in protein, DNA, RNA chains [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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
3	MLZ	C	114	10/11	0.98	0.21	-	35,37,47,50	0
1	MLY	A	135	11/12	0.71	0.31	-	74,80,82,82	0
2	MLZ	B	67	10/11	0.98	0.22	-	38,38,60,60	0
3	MLY	C	135	11/12	0.69	0.28	-	73,77,79,79	0
3	MLZ	C	129	10/11	0.98	0.20	-	33,34,59,61	0
3	MLY	C	67	11/12	0.97	0.18	-	36,38,62,64	0
4	MLY	D	135	11/12	0.65	0.25	-	72,79,82,82	0
4	MLY	D	67	11/12	0.98	0.24	-	38,39,67,68	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
1	MLZ	A	129	10/11	0.98	0.16	-	33,34,62,62	0
1	MLY	A	67	11/12	0.96	0.22	-	37,38,69,69	0
2	MLY	B	135	11/12	0.71	0.22	-	74,78,81,81	0

## 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
5	SO4	A	144	5/5	0.93	0.26	9.12	125,125,125,126	0
13	PG4	D	144	13/13	0.79	0.27	7.38	110,113,116,116	0
13	PG4	C	145	13/13	0.84	0.27	4.42	110,114,118,118	0
9	2PE	A	148	28/28	0.78	0.27	3.32	102,107,116,116	0
9	2PE	D	146	28/28	0.82	0.24	2.44	99,105,116,117	0
11	PE8	B	148	25/25	0.81	0.22	1.75	93,100,107,108	0
11	PE8	C	148	25/25	0.84	0.23	1.67	97,104,111,112	0
12	1PE	B	150	16/16	0.85	0.22	1.47	108,111,116,116	0
8	PGE	C	147	10/10	0.79	0.18	0.02	88,89,92,92	0
8	PGE	B	147	10/10	0.80	0.17	-0.35	89,96,100,100	0
5	SO4	A	143	5/5	0.99	0.14	-0.92	55,57,57,58	0
8	PGE	A	147	10/10	0.87	0.14	-2.01	84,85,87,88	0
6	CL	B	143	1/1	0.98	0.11	-3.31	54,54,54,54	0
6	CL	A	145	1/1	0.99	0.12	-4.11	54,54,54,54	0
6	CL	D	143	1/1	0.98	0.08	-4.20	54,54,54,54	0
6	CL	C	144	1/1	0.98	0.07	-5.01	55,55,55,55	0
7	PEG	C	146	7/7	0.92	0.19	-	76,77,79,80	0
10	GOL	C	149	6/6	0.89	0.20	-	77,78,79,79	0
7	PEG	B	145	7/7	0.93	0.15	-	76,76,78,78	0
7	PEG	B	146	7/7	0.94	0.15	-	80,81,84,85	0
10	GOL	D	147	6/6	0.89	0.16	-	82,86,87,88	0
10	GOL	A	150	6/6	0.92	0.22	-	77,79,80,80	0
7	PEG	D	145	7/7	0.80	0.22	-	110,111,111,112	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
10	GOL	B	149	6/6	0.90	0.17	-	73,76,80,82	0
5	SO4	C	143	5/5	0.99	0.15	-	51,52,53,53	5
7	PEG	B	144	7/7	0.79	0.23	-	127,128,129,130	0
7	PEG	A	146	7/7	0.92	0.19	-	77,78,81,82	0
7	PEG	A	149	7/7	0.65	0.24	-	105,107,108,109	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.