



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 16, 2022 – 01:29 pm BST

PDB ID : 7PZZ  
Title : Crystal structure of serine hydroxymethyltransferase, isoform 2 from *Arabidopsis thaliana* (SHM2)  
Authors : Ruzzkowski, M.; Sekula, B.  
Deposited on : 2021-10-13  
Resolution : 1.65 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

<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 : 1.8.4, CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.29  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0267  
CCP4 : 7.1.010 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.29

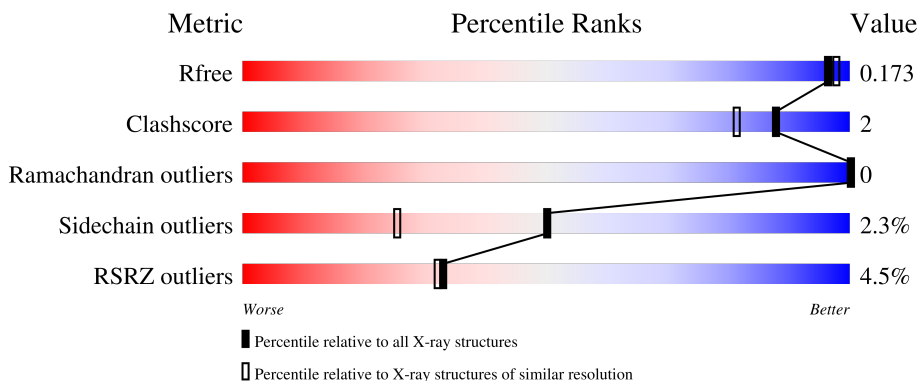
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

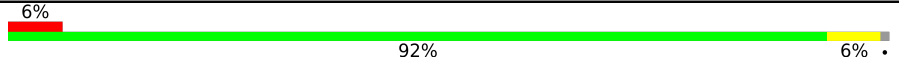
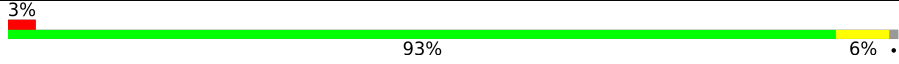
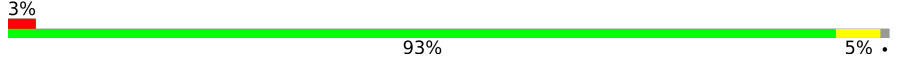
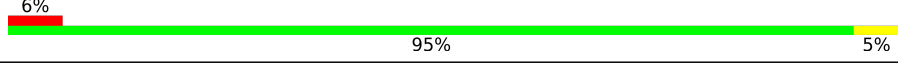
The reported resolution of this entry is 1.65 Å.

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))
$R_{free}$	130704	1827 (1.66-1.66)
Clashscore	141614	1931 (1.66-1.66)
Ramachandran outliers	138981	1891 (1.66-1.66)
Sidechain outliers	138945	1891 (1.66-1.66)
RSRZ outliers	127900	1791 (1.66-1.66)

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	480	 6% 92% 6%
1	B	480	 3% 93% 6%
1	C	480	 3% 93% 5%
1	D	480	 6% 95% 5%

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
5	PEG	C	604	-	-	X	-

## 2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 17345 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 Serine hydroxymethyltransferase 2, mitochondrial.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	P	S			
1	A	475	3746	2377	639	711	1	18	0	1	0
1	B	474	3740	2374	638	709	1	18	0	1	0
1	C	473	3767	2388	644	716	1	18	0	5	0
1	D	480	3816	2417	656	723	1	19	0	4	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	38	SER	-	expression tag	UNP Q94C74
A	39	ASN	-	expression tag	UNP Q94C74
A	40	ALA	-	expression tag	UNP Q94C74
B	38	SER	-	expression tag	UNP Q94C74
B	39	ASN	-	expression tag	UNP Q94C74
B	40	ALA	-	expression tag	UNP Q94C74
C	38	SER	-	expression tag	UNP Q94C74
C	39	ASN	-	expression tag	UNP Q94C74
C	40	ALA	-	expression tag	UNP Q94C74
D	38	SER	-	expression tag	UNP Q94C74
D	39	ASN	-	expression tag	UNP Q94C74
D	40	ALA	-	expression tag	UNP Q94C74

- Molecule 2 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
2	A	1	Total C O 4 2 2	0	0
2	A	1	Total C O 4 2 2	0	0
2	A	1	Total C O 4 2 2	0	0
2	A	1	Total C O 4 2 2	0	0
2	A	1	Total C O 4 2 2	0	0
2	A	1	Total C O 4 2 2	0	0
2	B	1	Total C O 4 2 2	0	0
2	B	1	Total C O 4 2 2	0	0
2	B	1	Total C O 4 2 2	0	0
2	B	1	Total C O 4 2 2	0	0
2	B	1	Total C O 4 2 2	0	0
2	B	1	Total C O 4 2 2	0	0
2	B	1	Total C O 4 2 2	0	0
2	C	1	Total C O 4 2 2	0	0

*Continued on next page...*

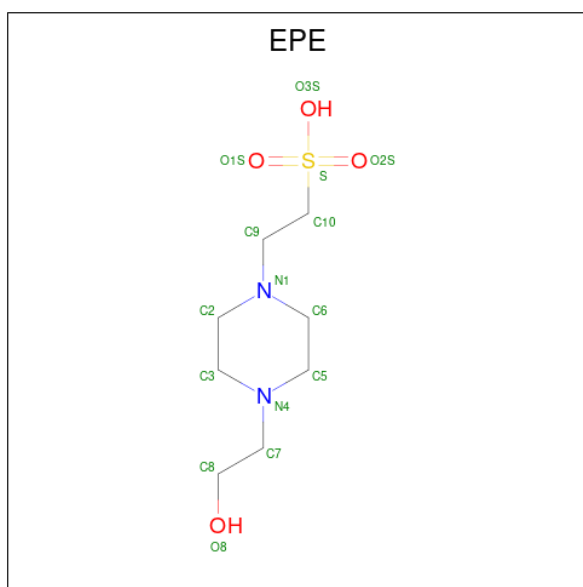
Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	C	1	Total C O 4 2 2	0	0
2	C	1	Total C O 4 2 2	0	0
2	C	1	Total C O 4 2 2	0	0
2	C	1	Total C O 4 2 2	0	0
2	C	1	Total C O 4 2 2	0	0
2	C	1	Total C O 4 2 2	0	0
2	C	1	Total C O 4 2 2	0	0
2	D	1	Total C O 4 2 2	0	0
2	D	1	Total C O 4 2 2	0	0
2	D	1	Total C O 4 2 2	0	0
2	D	1	Total C O 4 2 2	0	0
2	D	1	Total C O 4 2 2	0	0
2	D	1	Total C O 4 2 2	0	0
2	D	1	Total C O 4 2 2	0	0

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

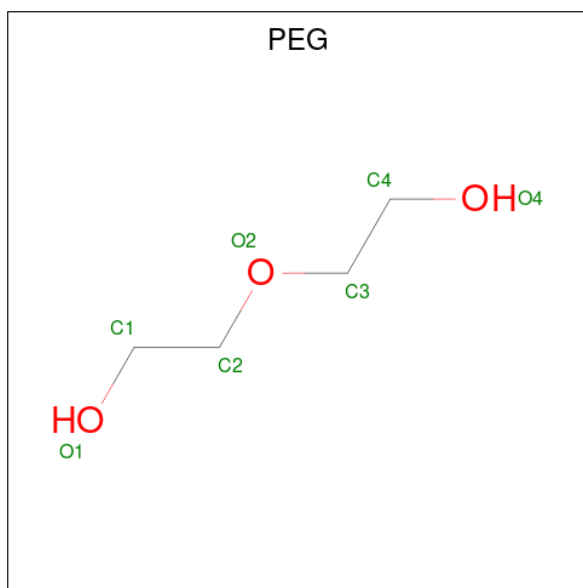
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total Cl 1 1	0	0
3	B	1	Total Cl 1 1	0	0
3	C	1	Total Cl 1 1	0	0
3	D	1	Total Cl 1 1	0	0

- Molecule 4 is 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID (three-letter code: EPE) (formula: C<sub>8</sub>H<sub>18</sub>N<sub>2</sub>O<sub>4</sub>S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
4	B	1	15	8	2	4	1	0	0

- Molecule 5 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula:  $C_4H_{10}O_3$ ).



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

- Molecule 6 is SULFATE ION (three-letter code: SO4) (formula:  $O_4S$ ).



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

- Molecule 7 is water.

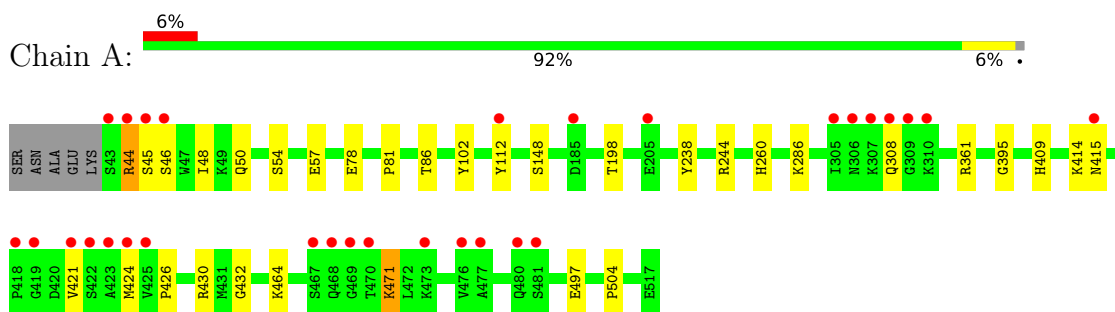
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	531	Total O 531 531	0	0
7	B	495	Total O 495 495	0	0
7	C	573	Total O 573 573	0	0
7	D	517	Total O 517 517	0	0



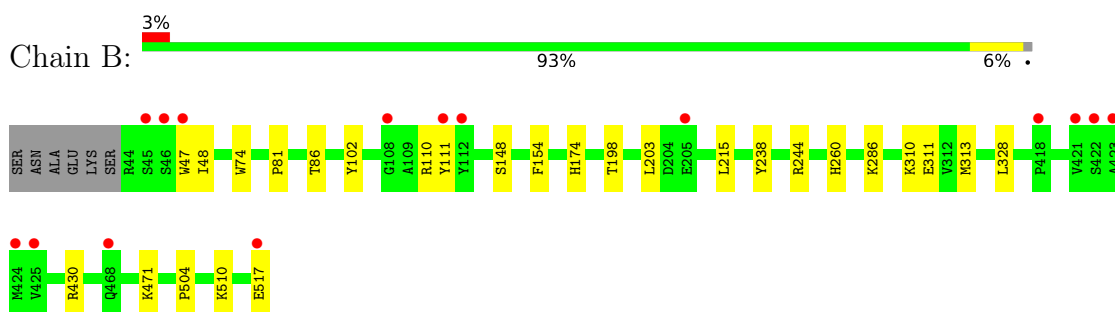
### 3 Residue-property plots [i](#)

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.

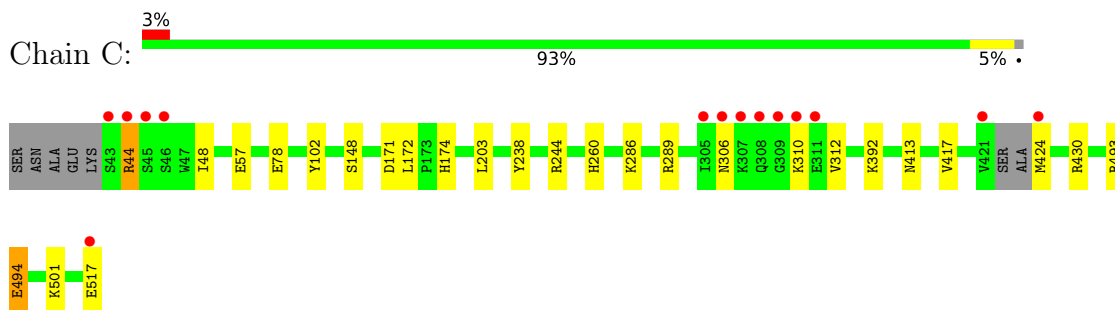
- Molecule 1: Serine hydroxymethyltransferase 2, mitochondrial



- Molecule 1: Serine hydroxymethyltransferase 2, mitochondrial

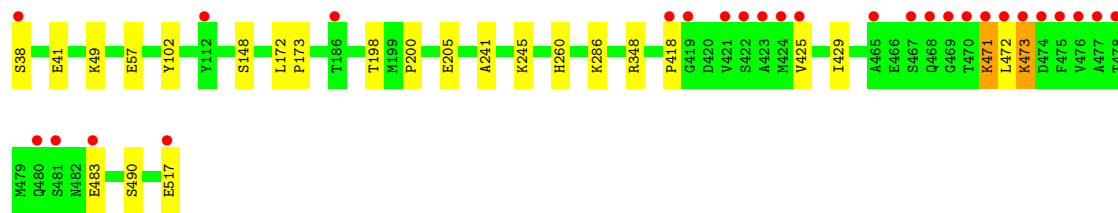


- Molecule 1: Serine hydroxymethyltransferase 2, mitochondrial



- Molecule 1: Serine hydroxymethyltransferase 2, mitochondrial





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	115.20Å 131.24Å 151.39Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	53.84 – 1.65 65.57 – 1.65	Depositor EDS
% Data completeness (in resolution range)	99.5 (53.84-1.65) 99.5 (65.57-1.65)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.08 (at 1.64Å)	Xtriage
Refinement program	PHENIX 1.19_4092	Depositor
R, $R_{free}$	0.158 , 0.178 0.157 , 0.173	Depositor DCC
$R_{free}$ test set	1238 reflections (0.45%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	18.3	Xtriage
Anisotropy	0.353	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	(Not available) , (Not available)	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	17345	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	25.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.05% 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 [i](#)

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

Bond lengths and bond angles in the following residue types are not validated in this section: EDO, CL, LLP, PEG, SO4, EPE

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	0.41	0/3795	0.60	0/5119
1	B	0.40	0/3789	0.60	0/5111
1	C	0.44	0/3815	0.63	0/5143
1	D	0.40	0/3865	0.61	0/5210
All	All	0.42	0/15264	0.61	0/20583

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3746	0	3742	17	0
1	B	3740	0	3737	16	0
1	C	3767	0	3756	16	0
1	D	3816	0	3812	13	0
2	A	24	0	36	1	0
2	B	28	0	42	3	0
2	C	32	0	48	0	0
2	D	28	0	42	2	0
3	A	1	0	0	1	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
4	B	15	0	17	0	0
5	C	7	0	10	4	0
5	D	7	0	10	0	0
6	C	15	0	0	0	0
7	A	531	0	0	5	0
7	B	495	0	0	4	0
7	C	573	0	0	8	0
7	D	517	0	0	2	0
All	All	17345	0	15252	61	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 2.

All (61) 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:172:LEU:H	5:C:604:PEG:H12	1.53	0.72
1:C:424:MET:SD	7:C:1041:HOH:O	2.52	0.68
1:D:471:LYS:HE2	1:D:472:LEU:H	1.64	0.63
1:C:172:LEU:N	5:C:604:PEG:H12	2.17	0.60
1:D:348[A]:ARG:NH1	7:D:703:HOH:O	2.36	0.58
1:A:44:ARG:NH1	1:A:45:SER:OG	2.38	0.57
1:D:205:GLU:HB3	1:D:418:PRO:HB3	1.86	0.56
1:C:493:ARG:NH2	7:C:708:HOH:O	2.39	0.55
1:D:418:PRO:HG2	7:D:714:HOH:O	2.09	0.53
1:A:50:GLN:NE2	7:A:702:HOH:O	2.36	0.52
1:B:198:THR:HB	2:B:607:EDO:H22	1.92	0.51
1:C:501:LYS:NZ	7:C:712:HOH:O	2.44	0.51
1:A:415:ASN:OD1	1:A:430:ARG:NH2	2.40	0.50
1:B:244:ARG:NH2	7:B:708:HOH:O	2.35	0.49
1:C:44:ARG:O	1:C:48:ILE:HG12	2.12	0.49
1:C:171:ASP:HA	5:C:604:PEG:H12	1.94	0.49
1:A:414:LYS:HD3	1:A:426:PRO:HG3	1.94	0.48
1:C:289:ARG:NH2	7:C:713:HOH:O	2.45	0.48
1:D:200:PRO:HA	2:D:606:EDO:H11	1.96	0.48
1:A:54:SER:OG	1:A:57:GLU:HG3	2.13	0.47
1:C:78:GLU:HG3	7:C:727:HOH:O	2.15	0.47
1:B:328:LEU:HD22	2:B:601:EDO:H22	1.96	0.47

Continued on next page...

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:430:ARG:NH1	3:A:607:CL:CL	2.86	0.46
1:D:38:SER:HB3	1:D:41:GLU:OE1	2.15	0.46
1:D:471:LYS:HA	1:D:471:LYS:HD2	1.71	0.46
1:B:471:LYS:HA	1:B:471:LYS:HD2	1.68	0.46
1:C:244:ARG:NH2	7:C:704:HOH:O	2.29	0.46
1:B:154:PHE:CD1	2:B:601:EDO:H11	2.51	0.45
1:A:409:HIS:ND1	1:A:497:GLU:OE2	2.49	0.45
1:C:392:LYS:HE2	1:C:417:VAL:HG23	1.99	0.45
5:C:604:PEG:H22	5:C:604:PEG:H42	1.58	0.45
1:B:74:TRP:NE1	1:B:510:LYS:HE2	2.32	0.44
1:A:504:PRO:HG2	1:B:48:ILE:HD13	1.98	0.44
1:B:110:ARG:HB3	7:B:712:HOH:O	2.17	0.44
1:A:48:ILE:HD13	1:B:504:PRO:HG2	2.01	0.43
1:B:311:GLU:HG3	1:B:313:MET:H	1.83	0.43
1:A:198:THR:HB	2:A:603:EDO:H11	2.00	0.43
1:A:244:ARG:NH2	7:A:719:HOH:O	2.52	0.42
1:A:395:GLY:O	1:A:464:LYS:NZ	2.48	0.42
1:A:112:TYR:HD1	7:A:723:HOH:O	2.01	0.42
1:B:517:GLU:OXT	7:B:701:HOH:O	2.22	0.42
1:D:172:LEU:HB3	1:D:173:PRO:HD3	2.01	0.42
1:D:241:ALA:O	1:D:245:LYS:HG2	2.20	0.42
1:A:361:ARG:NH2	7:A:718:HOH:O	2.52	0.42
1:A:81:PRO:HB3	1:A:432:GLY:HA3	2.02	0.41
1:D:471:LYS:HG3	1:D:473:LYS:HZ3	1.85	0.41
1:C:174:HIS:HB3	1:C:203:LEU:HG	2.02	0.41
1:C:57:GLU:OE1	7:C:701:HOH:O	2.22	0.41
1:B:74:TRP:HE1	1:B:510:LYS:HE2	1.86	0.41
1:D:198:THR:HB	2:D:606:EDO:H22	2.02	0.41
1:D:49:LYS:NZ	1:D:57:GLU:OE2	2.54	0.41
1:A:78:GLU:HG3	7:A:703:HOH:O	2.21	0.41
1:B:81:PRO:HG3	1:B:430:ARG:HG2	2.03	0.41
1:B:310:LYS:HD2	1:B:310:LYS:HA	1.84	0.41
1:C:494:GLU:HG2	7:C:836:HOH:O	2.20	0.41
1:A:471:LYS:HA	1:A:471:LYS:HD2	1.73	0.40
1:B:111:TYR:N	7:B:712:HOH:O	2.49	0.40
1:B:174:HIS:HB3	1:B:203:LEU:HG	2.04	0.40
1:C:413:ASN:HB3	1:C:430:ARG:HB3	2.03	0.40
1:D:245:LYS:HA	1:D:245:LYS:HD3	1.90	0.40
1:C:306:ASN:HD21	1:C:310:LYS:HB2	1.87	0.40

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	473/480 (98%)	459 (97%)	14 (3%)	0	100	100
1	B	472/480 (98%)	459 (97%)	13 (3%)	0	100	100
1	C	473/480 (98%)	461 (98%)	12 (2%)	0	100	100
1	D	481/480 (100%)	470 (98%)	11 (2%)	0	100	100
All	All	1899/1920 (99%)	1849 (97%)	50 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 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	397/400 (99%)	386 (97%)	11 (3%)	43	18
1	B	396/400 (99%)	389 (98%)	7 (2%)	59	36
1	C	400/400 (100%)	392 (98%)	8 (2%)	55	32
1	D	404/400 (101%)	394 (98%)	10 (2%)	47	22
All	All	1597/1600 (100%)	1561 (98%)	36 (2%)	50	25

All (36) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	44	ARG
1	A	46	SER

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>
1	A	86	THR
1	A	102	TYR
1	A	148	SER
1	A	238	TYR
1	A	260	HIS
1	A	308	GLN
1	A	421	VAL
1	A	424	MET
1	A	471	LYS
1	B	47	TRP
1	B	86	THR
1	B	102	TYR
1	B	148	SER
1	B	215	LEU
1	B	238	TYR
1	B	260	HIS
1	C	44	ARG
1	C	102	TYR
1	C	148	SER
1	C	238	TYR
1	C	260	HIS
1	C	312	VAL
1	C	494	GLU
1	C	517	GLU
1	D	102	TYR
1	D	148	SER
1	D	260	HIS
1	D	425	VAL
1	D	429	ILE
1	D	471	LYS
1	D	473	LYS
1	D	483	GLU
1	D	490	SER
1	D	517	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>
1	A	346	GLN
1	B	214	GLN
1	C	308	GLN



### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

4 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	LLP	B	286	1	23,24,25	1.58	4 (17%)	25,32,34	1.86	4 (16%)
1	LLP	A	286	1	23,24,25	1.45	5 (21%)	25,32,34	1.89	5 (20%)
1	LLP	D	286	1	23,24,25	1.55	5 (21%)	25,32,34	1.74	4 (16%)
1	LLP	C	286	1	23,24,25	1.62	5 (21%)	25,32,34	2.06	6 (24%)

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	LLP	B	286	1	-	5/16/17/19	0/1/1/1
1	LLP	A	286	1	-	5/16/17/19	0/1/1/1
1	LLP	D	286	1	-	5/16/17/19	0/1/1/1
1	LLP	C	286	1	-	5/16/17/19	0/1/1/1

All (19) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	286	LLP	C4-C4'	3.94	1.54	1.46
1	D	286	LLP	C4-C4'	3.79	1.53	1.46
1	B	286	LLP	C4-C4'	3.50	1.53	1.46
1	B	286	LLP	C4-C5	-3.43	1.37	1.42
1	A	286	LLP	C4-C4'	3.39	1.53	1.46
1	C	286	LLP	C4-C5	-3.21	1.37	1.42

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	286	LLP	C3-C2	-3.09	1.37	1.40
1	D	286	LLP	C4-C5	-2.99	1.38	1.42
1	B	286	LLP	C3-C2	-2.96	1.37	1.40
1	C	286	LLP	C4-C3	-2.79	1.36	1.40
1	A	286	LLP	C4-C5	-2.78	1.38	1.42
1	A	286	LLP	C4-C3	-2.68	1.36	1.40
1	D	286	LLP	C2'-C2	2.47	1.54	1.50
1	A	286	LLP	C3-C2	-2.25	1.38	1.40
1	A	286	LLP	C2'-C2	2.23	1.54	1.50
1	C	286	LLP	C2'-C2	2.22	1.54	1.50
1	B	286	LLP	C2'-C2	2.19	1.54	1.50
1	D	286	LLP	C4-C3	-2.10	1.37	1.40
1	D	286	LLP	C3-C2	-2.08	1.38	1.40

All (19) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	286	LLP	OP4-C5'-C5	5.80	120.39	109.35
1	B	286	LLP	OP4-C5'-C5	5.46	119.76	109.35
1	A	286	LLP	OP4-C5'-C5	5.25	119.35	109.35
1	D	286	LLP	OP4-C5'-C5	4.70	118.30	109.35
1	D	286	LLP	CE-NZ-C4'	4.61	133.06	118.90
1	C	286	LLP	CE-NZ-C4'	4.50	132.72	118.90
1	A	286	LLP	CE-NZ-C4'	4.18	131.75	118.90
1	B	286	LLP	CE-NZ-C4'	4.09	131.47	118.90
1	A	286	LLP	C4-C3-C2	3.82	122.55	120.19
1	A	286	LLP	C4-C4'-NZ	-3.36	108.88	124.31
1	B	286	LLP	C4-C4'-NZ	-3.35	108.91	124.31
1	D	286	LLP	C4-C4'-NZ	-3.34	108.98	124.31
1	C	286	LLP	C4-C4'-NZ	-3.25	109.40	124.31
1	C	286	LLP	C4-C3-C2	2.97	122.03	120.19
1	A	286	LLP	C5'-C5-C6	-2.62	115.07	119.37
1	B	286	LLP	C5'-C5-C6	-2.51	115.24	119.37
1	C	286	LLP	OP4-P-OP1	-2.31	99.99	106.47
1	C	286	LLP	C5-C6-N1	-2.11	120.30	123.82
1	D	286	LLP	C4-C3-C2	2.04	121.45	120.19

There are no chirality outliers.

All (20) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	286	LLP	O-C-CA-CB

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
1	B	286	LLP	O-C-CA-CB
1	C	286	LLP	O-C-CA-CB
1	D	286	LLP	O-C-CA-CB
1	A	286	LLP	CG-CD-CE-NZ
1	B	286	LLP	CG-CD-CE-NZ
1	C	286	LLP	CG-CD-CE-NZ
1	D	286	LLP	CG-CD-CE-NZ
1	A	286	LLP	C3-C4-C4'-NZ
1	C	286	LLP	C3-C4-C4'-NZ
1	B	286	LLP	C3-C4-C4'-NZ
1	B	286	LLP	CD-CE-NZ-C4'
1	C	286	LLP	CD-CE-NZ-C4'
1	A	286	LLP	CD-CE-NZ-C4'
1	C	286	LLP	C5-C4-C4'-NZ
1	D	286	LLP	CD-CE-NZ-C4'
1	D	286	LLP	C3-C4-C4'-NZ
1	B	286	LLP	C5-C4-C4'-NZ
1	A	286	LLP	C5-C4-C4'-NZ
1	D	286	LLP	C5-C4-C4'-NZ

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 38 ligands modelled in this entry, 4 are monoatomic - leaving 34 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
6	SO4	C	611	-	4,4,4	0.23	0	6,6,6	0.41	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	EDO	C	603	-	3,3,3	0.46	0	2,2,2	0.62	0
2	EDO	C	605	-	3,3,3	0.46	0	2,2,2	0.22	0
2	EDO	D	607	-	3,3,3	0.47	0	2,2,2	0.53	0
2	EDO	C	608	-	3,3,3	0.50	0	2,2,2	0.33	0
2	EDO	B	606	-	3,3,3	0.48	0	2,2,2	0.27	0
2	EDO	C	601	-	3,3,3	0.50	0	2,2,2	0.36	0
2	EDO	C	609	-	3,3,3	0.53	0	2,2,2	0.37	0
2	EDO	A	602	-	3,3,3	0.49	0	2,2,2	0.38	0
2	EDO	D	603	-	3,3,3	0.45	0	2,2,2	0.09	0
2	EDO	A	603	-	3,3,3	0.55	0	2,2,2	0.47	0
2	EDO	A	605	-	3,3,3	0.52	0	2,2,2	0.29	0
5	PEG	C	604	-	6,6,6	0.13	0	5,5,5	0.26	0
2	EDO	B	601	-	3,3,3	0.30	0	2,2,2	0.39	0
2	EDO	B	603	-	3,3,3	0.43	0	2,2,2	0.45	0
2	EDO	D	602	-	3,3,3	0.40	0	2,2,2	0.39	0
2	EDO	C	602	-	3,3,3	0.57	0	2,2,2	0.17	0
2	EDO	B	607	-	3,3,3	0.67	0	2,2,2	0.49	0
4	EPE	B	602	-	15,15,15	0.79	1 (6%)	18,20,20	1.68	5 (27%)
2	EDO	A	606	-	3,3,3	0.41	0	2,2,2	0.45	0
5	PEG	D	605	-	6,6,6	0.15	0	5,5,5	0.16	0
2	EDO	B	609	-	3,3,3	0.52	0	2,2,2	0.17	0
6	SO4	C	610	-	4,4,4	0.13	0	6,6,6	0.29	0
2	EDO	D	604	-	3,3,3	0.51	0	2,2,2	0.11	0
2	EDO	D	606	-	3,3,3	0.46	0	2,2,2	0.57	0
2	EDO	A	601	-	3,3,3	0.44	0	2,2,2	0.45	0
6	SO4	C	612	-	4,4,4	0.11	0	6,6,6	0.19	0
2	EDO	A	604	-	3,3,3	0.47	0	2,2,2	0.39	0
2	EDO	C	606	-	3,3,3	0.57	0	2,2,2	0.28	0
2	EDO	D	608	-	3,3,3	0.48	0	2,2,2	0.31	0
2	EDO	C	607	-	3,3,3	0.52	0	2,2,2	0.44	0
2	EDO	D	601	-	3,3,3	0.50	0	2,2,2	0.27	0
2	EDO	B	604	-	3,3,3	0.45	0	2,2,2	0.68	0
2	EDO	B	605	-	3,3,3	0.43	0	2,2,2	0.43	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
2	EDO	C	603	-	-	0/1/1/1	-
2	EDO	C	605	-	-	1/1/1/1	-

*Continued on next page...*

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	EDO	D	607	-	-	0/1/1/1	-
2	EDO	C	608	-	-	1/1/1/1	-
2	EDO	B	606	-	-	0/1/1/1	-
2	EDO	C	601	-	-	0/1/1/1	-
2	EDO	C	609	-	-	0/1/1/1	-
2	EDO	A	602	-	-	0/1/1/1	-
2	EDO	D	603	-	-	1/1/1/1	-
2	EDO	A	603	-	-	1/1/1/1	-
2	EDO	A	605	-	-	0/1/1/1	-
5	PEG	C	604	-	-	4/4/4/4	-
2	EDO	B	601	-	-	1/1/1/1	-
2	EDO	B	603	-	-	1/1/1/1	-
2	EDO	D	602	-	-	0/1/1/1	-
2	EDO	C	602	-	-	0/1/1/1	-
2	EDO	B	607	-	-	1/1/1/1	-
4	EPE	B	602	-	-	4/9/19/19	0/1/1/1
2	EDO	A	606	-	-	1/1/1/1	-
5	PEG	D	605	-	-	1/4/4/4	-
2	EDO	B	609	-	-	0/1/1/1	-
2	EDO	D	604	-	-	0/1/1/1	-
2	EDO	D	606	-	-	0/1/1/1	-
2	EDO	A	601	-	-	0/1/1/1	-
2	EDO	A	604	-	-	0/1/1/1	-
2	EDO	C	606	-	-	1/1/1/1	-
2	EDO	D	608	-	-	0/1/1/1	-
2	EDO	C	607	-	-	0/1/1/1	-
2	EDO	D	601	-	-	1/1/1/1	-
2	EDO	B	604	-	-	0/1/1/1	-
2	EDO	B	605	-	-	1/1/1/1	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	602	EPE	C10-S	2.64	1.81	1.77

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	602	EPE	C5-N4-C3	4.33	118.58	108.83
4	B	602	EPE	C7-N4-C5	2.48	117.58	111.23
4	B	602	EPE	O1S-S-C10	2.33	109.72	106.92
4	B	602	EPE	C7-N4-C3	2.19	116.83	111.23

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	602	EPE	O3S-S-C10	2.02	109.03	105.77

There are no chirality outliers.

All (20) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	B	602	EPE	C9-C10-S-O1S
4	B	602	EPE	C9-C10-S-O3S
5	C	604	PEG	C4-C3-O2-C2
2	A	606	EDO	O1-C1-C2-O2
2	B	603	EDO	O1-C1-C2-O2
2	B	607	EDO	O1-C1-C2-O2
2	C	605	EDO	O1-C1-C2-O2
2	D	601	EDO	O1-C1-C2-O2
5	C	604	PEG	O2-C3-C4-O4
5	C	604	PEG	O1-C1-C2-O2
2	C	608	EDO	O1-C1-C2-O2
5	C	604	PEG	C1-C2-O2-C3
5	D	605	PEG	O2-C3-C4-O4
2	B	605	EDO	O1-C1-C2-O2
4	B	602	EPE	C9-C10-S-O2S
4	B	602	EPE	N4-C7-C8-O8
2	A	603	EDO	O1-C1-C2-O2
2	C	606	EDO	O1-C1-C2-O2
2	D	603	EDO	O1-C1-C2-O2
2	B	601	EDO	O1-C1-C2-O2

There are no ring outliers.

5 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	603	EDO	1	0
5	C	604	PEG	4	0
2	B	601	EDO	2	0
2	B	607	EDO	1	0
2	D	606	EDO	2	0

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues

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	474/480 (98%)	0.03	30 (6%) 20 19	12, 20, 57, 96	0
1	B	473/480 (98%)	-0.07	15 (3%) 47 48	11, 22, 56, 88	0
1	C	472/480 (98%)	0.05	14 (2%) 50 51	12, 19, 40, 80	0
1	D	479/480 (99%)	0.05	27 (5%) 24 23	12, 21, 52, 95	0
All	All	1898/1920 (98%)	0.01	86 (4%) 33 32	11, 21, 52, 96	0

All (86) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	424	MET	10.9
1	A	421	VAL	8.7
1	D	421	VAL	8.4
1	C	421	VAL	7.4
1	A	422	SER	6.8
1	D	468	GLN	6.4
1	D	423	ALA	6.3
1	A	309	GLY	6.2
1	D	424	MET	5.2
1	A	470	THR	5.2
1	A	45	SER	5.1
1	B	424	MET	5.0
1	B	111	TYR	5.0
1	D	470	THR	4.9
1	A	425	VAL	4.8
1	B	112	TYR	4.6
1	C	305	ILE	4.5
1	B	425	VAL	4.5
1	D	419	GLY	4.5
1	A	112	TYR	4.5
1	A	310	LYS	4.4

*Continued on next page...*



*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
1	C	43	SER	4.4
1	A	468	GLN	4.4
1	B	468	GLN	4.4
1	D	425	VAL	4.4
1	A	418	PRO	4.3
1	A	307	LYS	4.3
1	C	310	LYS	4.2
1	A	477	ALA	4.2
1	A	46	SER	4.1
1	D	422	SER	4.1
1	B	517	GLU	4.1
1	B	46	SER	3.8
1	A	44	ARG	3.8
1	B	45	SER	3.8
1	A	308	GLN	3.6
1	B	421	VAL	3.6
1	C	45	SER	3.4
1	D	473	LYS	3.4
1	C	424	MET	3.4
1	D	465	ALA	3.4
1	C	307	LYS	3.3
1	D	477	ALA	3.3
1	B	422	SER	3.3
1	A	473	LYS	3.3
1	A	480	GLN	3.1
1	C	308	GLN	3.1
1	A	423	ALA	3.1
1	D	469	GLY	3.0
1	A	419	GLY	3.0
1	B	47	TRP	3.0
1	A	415	ASN	3.0
1	B	423	ALA	2.9
1	B	108	GLY	2.9
1	D	472	LEU	2.9
1	C	309	GLY	2.8
1	A	469	GLY	2.8
1	A	305	ILE	2.7
1	C	306	ASN	2.7
1	C	517	GLU	2.7
1	C	46	SER	2.7
1	C	44	ARG	2.7
1	D	38	SER	2.6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	D	517	GLU	2.6
1	D	471	LYS	2.6
1	A	467	SER	2.5
1	D	476	VAL	2.5
1	D	186	THR	2.4
1	D	467	SER	2.4
1	D	483	GLU	2.4
1	A	476	VAL	2.3
1	A	481	SER	2.3
1	D	418	PRO	2.3
1	D	481	SER	2.3
1	D	112	TYR	2.2
1	B	418	PRO	2.2
1	A	43	SER	2.2
1	A	205	GLU	2.2
1	A	306	ASN	2.2
1	D	475	PHE	2.1
1	B	205	GLU	2.1
1	C	311	GLU	2.1
1	A	185	ASP	2.0
1	D	478	THR	2.0
1	D	480	GLN	2.0
1	D	474	ASP	2.0

## 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. 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
1	LLP	D	286	24/25	0.96	0.09	13,16,20,26	0
1	LLP	B	286	24/25	0.97	0.08	13,16,23,28	0
1	LLP	C	286	24/25	0.97	0.10	12,14,22,24	0
1	LLP	A	286	24/25	0.97	0.09	13,15,23,25	0

## 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
5	PEG	D	605	7/7	0.67	0.16	44,51,52,53	0
2	EDO	D	608	4/4	0.78	0.31	36,36,40,46	0
2	EDO	C	607	4/4	0.80	0.13	36,38,40,42	0
4	EPE	B	602	15/15	0.81	0.24	46,54,72,74	0
2	EDO	A	605	4/4	0.82	0.26	35,36,44,46	0
2	EDO	C	602	4/4	0.82	0.14	31,32,37,41	0
2	EDO	C	605	4/4	0.83	0.24	37,38,40,51	0
2	EDO	B	607	4/4	0.85	0.21	23,25,28,31	0
2	EDO	C	601	4/4	0.85	0.19	30,32,40,40	0
2	EDO	D	607	4/4	0.86	0.13	24,32,38,40	0
5	PEG	C	604	7/7	0.86	0.26	16,18,20,22	7
2	EDO	B	605	4/4	0.86	0.15	30,31,34,44	0
2	EDO	A	603	4/4	0.87	0.27	18,22,25,32	0
2	EDO	B	609	4/4	0.87	0.32	26,32,35,46	0
2	EDO	B	606	4/4	0.87	0.15	36,39,42,45	0
2	EDO	D	604	4/4	0.87	0.17	26,28,34,40	0
2	EDO	D	606	4/4	0.87	0.18	22,22,27,30	0
6	SO4	C	611	5/5	0.87	0.26	29,31,47,50	0
2	EDO	C	606	4/4	0.89	0.12	25,26,28,31	0
2	EDO	C	608	4/4	0.89	0.18	31,32,38,46	0
2	EDO	C	609	4/4	0.90	0.15	29,35,39,41	0
2	EDO	D	602	4/4	0.91	0.12	23,24,25,27	0
6	SO4	C	612	5/5	0.91	0.34	40,44,57,60	0
2	EDO	B	601	4/4	0.92	0.27	25,25,27,32	0
2	EDO	B	604	4/4	0.92	0.18	19,22,30,36	0
2	EDO	C	603	4/4	0.92	0.15	27,30,34,35	0
2	EDO	A	604	4/4	0.92	0.17	25,33,34,45	0
2	EDO	D	601	4/4	0.92	0.12	25,29,30,37	0
6	SO4	C	610	5/5	0.93	0.16	38,46,59,63	0
2	EDO	B	603	4/4	0.94	0.10	25,26,26,33	0
2	EDO	A	606	4/4	0.95	0.11	27,30,31,37	0
2	EDO	A	602	4/4	0.95	0.08	27,31,34,38	0
2	EDO	A	601	4/4	0.96	0.07	22,24,29,30	0
2	EDO	D	603	4/4	0.96	0.12	29,31,35,35	0
3	CL	D	609	1/1	0.96	0.07	29,29,29,29	0
3	CL	B	608	1/1	0.98	0.07	27,27,27,27	0
3	CL	C	613	1/1	0.99	0.07	22,22,22,22	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	CL	A	607	1/1	0.99	0.04	27,27,27,27	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.