



# Full wwPDB X-ray Structure Validation Report ⓘ

May 14, 2020 – 08:01 am BST

PDB ID : 3B5Q  
Title : Crystal structure of a putative sulfatase (NP\_810509.1) from Bacteroides  
thetaiotaomicron VPI-5482 at 2.40 Å resolution  
Authors : Joint Center for Structural Genomics (JCSG)  
Deposited on : 2007-10-26  
Resolution : 2.40 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  
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

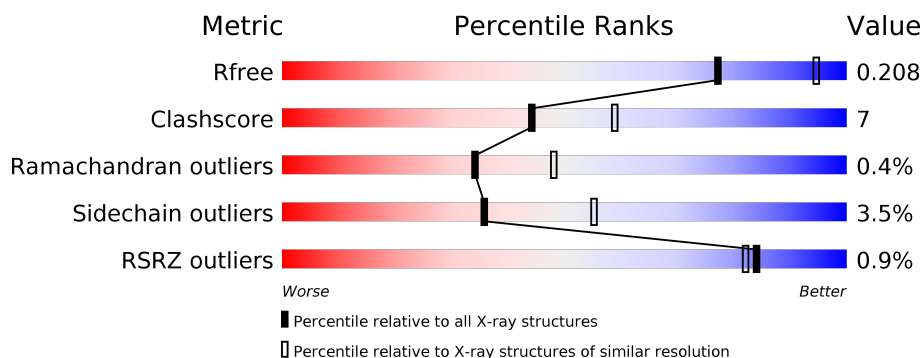
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.40 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	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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	482	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 1%, orange 1%, yellow 13%, green 82%, grey 14%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>%</span> <span>82%</span> <span>13%</span> <span>• •</span> </div> </div>
1	B	482	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 1%, orange 1%, yellow 10%, green 86%, grey 3%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>%</span> <span>86%</span> <span>10%</span> <span>• •</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 criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	EDO	B	505	-	-	X	-

## 2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 7815 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 Putative sulfatase yidJ.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	464	Total	C	N	O	S	Se	0	4	0
			3628	2296	633	678	10	11			
1	B	467	Total	C	N	O	S	Se	0	6	0
			3654	2312	637	684	10	11			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	GLY	-	LEADER SEQUENCE	UNP Q8A7C8
B	0	GLY	-	LEADER SEQUENCE	UNP Q8A7C8

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

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

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

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



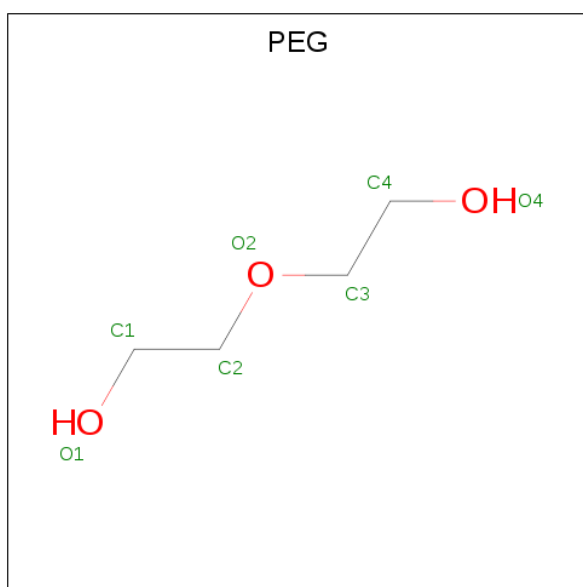
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			4	2	2		
4	A	1	Total	C	O	0	0
			4	2	2		
4	A	1	Total	C	O	0	0
			4	2	2		
4	A	1	Total	C	O	0	0
			4	2	2		
4	A	1	Total	C	O	0	0
			4	2	2		
4	A	1	Total	C	O	0	0
			4	2	2		
4	A	1	Total	C	O	0	0
			4	2	2		
4	A	1	Total	C	O	0	0
			4	2	2		
4	A	1	Total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

*Continued on next page...*

Continued from previous page...

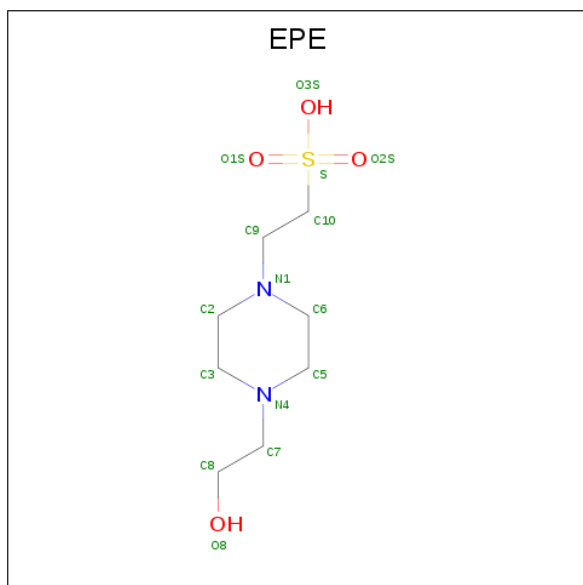
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	B	1	Total	C	O	0	0
			4	2	2		
4	B	1	Total	C	O	0	0
			4	2	2		
4	B	1	Total	C	O	0	0
			4	2	2		
4	B	1	Total	C	O	0	0
			4	2	2		
4	B	1	Total	C	O	0	0
			4	2	2		
4	B	1	Total	C	O	0	0
			4	2	2		
4	B	1	Total	C	O	0	0
			4	2	2		
4	B	1	Total	C	O	0	0
			4	2	2		

- Molecule 5 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: C<sub>4</sub>H<sub>10</sub>O<sub>3</sub>).



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

- Molecule 6 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
6	B	1	Total C N O S 15 8 2 4 1	0	0
6	B	1	Total C N O S 15 8 2 4 1	0	0
6	B	1	Total C N O S 15 8 2 4 1	0	0

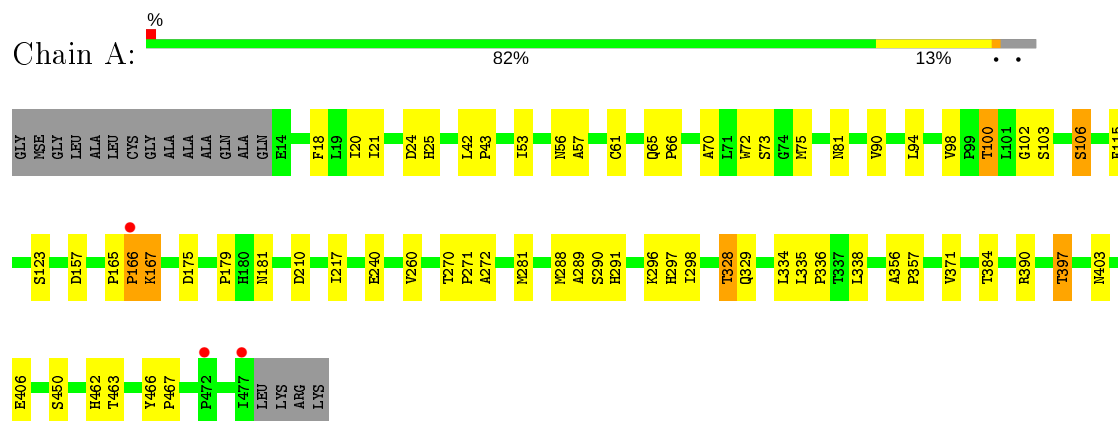
- Molecule 7 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	186	Total O 186 186	0	0
7	B	177	Total O 177 177	0	0

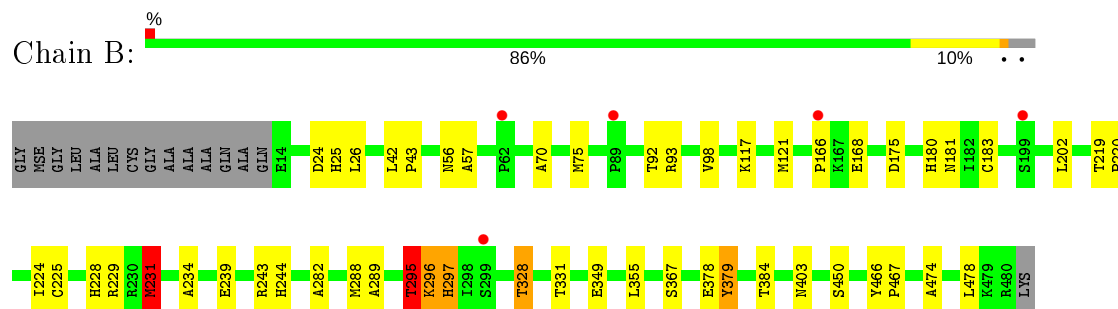
### 3 Residue-property plots [i](#)

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: Putative sulfatase yidJ



- Molecule 1: Putative sulfatase yidJ





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	106.28Å 106.28Å 111.68Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.91 – 2.40 29.91 – 2.40	Depositor EDS
% Data completeness (in resolution range)	99.2 (29.91-2.40) 99.2 (29.91-2.40)	Depositor EDS
$R_{merge}$	0.15	Depositor
$R_{sym}$	0.15	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.40 (at 2.39Å)	Xtriage
Refinement program	REFMAC 5.2.0019, PHENIX	Depositor
R, $R_{free}$	0.149 , 0.206 0.154 , 0.208	Depositor DCC
$R_{free}$ test set	2420 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	32.9	Xtriage
Anisotropy	0.280	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 55.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.017 for -h,-l,-k 0.007 for -h,l,k 0.007 for l,-k,h 0.017 for -l,-k,-h 0.039 for h,-k,-l	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	7815	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.07% of the height of the origin peak. No significant pseudotranslation is detected.*

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

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: PEG, ZN, EPE, EDO, 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	0.73	1/3733 (0.0%)	0.70	0/5062
1	B	0.72	1/3766 (0.0%)	0.69	1/5113 (0.0%)
All	All	0.72	2/7499 (0.0%)	0.69	1/10175 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
1	B	0	1
All	All	0	2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	61	CYS	CB-SG	5.58	1.91	1.82
1	B	183	CYS	CB-SG	-5.39	1.73	1.81

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	B	231	MSE	CG-SE-CE	7.26	114.88	98.90

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	165	PRO	Peptide
1	B	166	PRO	Peptide

## 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	3628	0	3477	48	0
1	B	3654	0	3465	47	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	56	0	84	1	0
4	B	44	0	66	4	0
5	A	7	0	10	0	0
5	B	14	0	20	0	0
6	B	45	0	52	6	0
7	A	186	0	0	4	0
7	B	177	0	0	3	0
All	All	7815	0	7174	97	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 7.

All (97) 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:93[B]:ARG:HG2	1:B:121:MSE:HE1	1.18	1.10
1:A:75:MSE:CE	1:A:98:VAL:HG11	1.84	1.07
6:B:504:EPE:H51	7:B:547:HOH:O	1.53	1.06
1:B:295:THR:HG22	1:B:296:LYS:H	1.28	0.97
1:B:295:THR:CG2	1:B:296:LYS:H	1.79	0.96
1:B:93[B]:ARG:HG2	1:B:121:MSE:CE	1.96	0.94
1:A:75:MSE:HE1	1:A:98:VAL:HG11	1.49	0.91
1:B:75:MSE:HE1	1:B:98:VAL:CG1	2.01	0.91
6:B:504:EPE:O3S	6:B:504:EPE:N1	2.05	0.89

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:180:HIS:NE2	6:B:504:EPE:H32	1.89	0.88
1:A:166:PRO:HA	1:A:167:LYS:HB3	1.54	0.88
1:A:100:THR:HG22	1:A:103:SER:H	1.39	0.84
1:A:75:MSE:HE2	1:A:98:VAL:HG11	1.62	0.80
1:A:90:VAL:HG23	7:A:543:HOH:O	1.82	0.79
1:B:295:THR:HG22	1:B:296:LYS:N	1.97	0.78
1:B:56:ASN:O	1:B:328:THR:HB	1.86	0.75
1:B:355:LEU:HD21	4:B:505:EDO:H12	1.70	0.74
1:B:75:MSE:HE1	1:B:98:VAL:HG11	1.68	0.74
1:B:93[B]:ARG:CG	1:B:121:MSE:HE1	2.10	0.73
1:B:75:MSE:HE1	1:B:98:VAL:HG13	1.73	0.70
1:A:57:ALA:HA	1:A:328:THR:HG22	1.73	0.69
1:A:75:MSE:HE2	1:A:98:VAL:CG1	2.24	0.68
1:A:270:THR:HG22	1:A:272:ALA:H	1.56	0.68
1:B:295:THR:HG23	1:B:296:LYS:HD3	1.75	0.68
1:B:180:HIS:NE2	6:B:504:EPE:C3	2.59	0.66
1:B:295:THR:CG2	1:B:296:LYS:N	2.46	0.66
1:A:102:GLY:O	1:A:106:SER:HB2	1.99	0.61
1:B:93[B]:ARG:HA	1:B:121:MSE:CE	2.32	0.59
1:A:270:THR:HB	7:A:531:HOH:O	2.00	0.59
1:B:93[B]:ARG:CG	1:B:121:MSE:CE	2.78	0.58
1:A:24:ASP:O	1:A:25:HIS:HB2	2.04	0.58
1:B:93[B]:ARG:HA	1:B:121:MSE:HE2	1.86	0.58
1:B:295:THR:HG21	7:B:611:HOH:O	2.04	0.57
1:B:93[A]:ARG:HA	1:B:121:MSE:CE	2.34	0.57
1:A:270:THR:HG23	1:A:271:PRO:HD2	1.86	0.56
1:B:57:ALA:HA	1:B:328:THR:HG22	1.88	0.56
1:A:397:THR:HG21	1:A:406:GLU:OE1	2.05	0.56
1:A:56:ASN:O	1:A:328:THR:HB	2.05	0.56
1:A:100:THR:HG23	7:A:538:HOH:O	2.07	0.54
1:B:24:ASP:O	1:B:25:HIS:HB2	2.07	0.54
1:B:93[A]:ARG:HA	1:B:121:MSE:HE2	1.90	0.54
1:A:166:PRO:HA	1:A:167:LYS:CB	2.32	0.53
1:A:397:THR:HB	1:A:406:GLU:HB2	1.90	0.53
1:B:224:ILE:HG22	1:B:231:MSE:HG3	1.90	0.52
1:B:367:SER:HA	4:B:505:EDO:H22	1.92	0.52
1:B:231:MSE:HE3	1:B:234:ALA:HB3	1.92	0.51
1:B:228:HIS:HB2	1:B:297:HIS:CD2	2.46	0.51
1:B:295:THR:HG23	1:B:296:LYS:H	1.68	0.51
1:A:384:THR:HG23	1:A:384:THR:O	2.12	0.49
1:A:290:SER:O	1:A:291:HIS:HB2	2.13	0.49

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:20:ILE:HD11	1:A:338:LEU:HD11	1.94	0.49
1:A:75:MSE:CE	1:A:98:VAL:CG1	2.71	0.49
1:A:75:MSE:HE3	1:A:94:LEU:HD11	1.94	0.49
1:B:288:MSE:O	1:B:289:ALA:HB3	2.12	0.49
1:B:367:SER:HA	4:B:505:EDO:C2	2.43	0.49
1:A:20:ILE:HD11	1:A:338:LEU:CD1	2.42	0.49
1:A:21:ILE:HG21	1:A:260:VAL:HG22	1.96	0.48
1:B:202:LEU:HD11	1:B:244:HIS:CE1	2.49	0.47
1:B:296:LYS:HZ1	6:B:504:EPE:H52	1.77	0.47
1:B:42[A]:LEU:HB3	1:B:43:PRO:HD3	1.96	0.47
1:B:384:THR:O	1:B:384:THR:HG23	2.14	0.47
1:A:72:TRP:HB2	1:A:123[B]:SER:HB3	1.96	0.47
1:A:270:THR:CG2	1:A:271:PRO:HD2	2.46	0.46
1:A:281:MSE:HE2	7:A:628:HOH:O	2.15	0.46
1:B:228:HIS:HB2	1:B:297:HIS:CG	2.51	0.45
1:A:217:ILE:HG23	4:A:507:EDO:H21	2.00	0.44
1:A:100:THR:HG22	1:A:103:SER:N	2.20	0.44
1:A:75:MSE:CE	1:A:94:LEU:HD11	2.48	0.44
1:A:335:LEU:HB3	1:A:336:PRO:CD	2.47	0.44
1:B:355:LEU:CD2	4:B:505:EDO:H12	2.44	0.44
1:A:42:LEU:HB3	1:A:43:PRO:HD3	2.00	0.43
1:A:70:ALA:HB1	1:A:75:MSE:O	2.18	0.43
1:B:474:ALA:HB1	1:B:478:LEU:HD12	2.00	0.43
1:A:18:PHE:CD2	1:A:338:LEU:HD22	2.54	0.43
1:A:73:SER:HB2	1:A:75:MSE:HE3	2.00	0.42
1:B:225:CYS:HA	1:B:231:MSE:HB3	2.01	0.42
1:B:93[B]:ARG:CA	1:B:121:MSE:CE	2.98	0.42
1:B:26:LEU:HB3	1:B:282:ALA:HB2	2.02	0.42
1:B:219:THR:N	1:B:220:PRO:CD	2.83	0.42
1:A:65:GLN:HB3	1:A:66:PRO:HD3	2.02	0.42
1:A:371:VAL:HG23	1:A:390:ARG:HB3	2.02	0.41
1:A:466:TYR:N	1:A:467:PRO:CD	2.84	0.41
1:B:239:GLU:O	1:B:243:ARG:HG3	2.20	0.41
1:B:295:THR:CG2	7:B:611:HOH:O	2.67	0.41
1:B:378:GLU:HG2	1:B:379[B]:TYR:CD2	2.55	0.41
1:A:288:MSE:O	1:A:289:ALA:HB3	2.20	0.41
1:A:115:PHE:CD1	1:A:157:ASP:HB3	2.56	0.41
1:A:24:ASP:O	1:A:179:PRO:HD2	2.20	0.41
1:A:281:MSE:SE	1:A:334:LEU:HD22	2.71	0.41
1:A:356:ALA:N	1:A:357:PRO:CD	2.83	0.41
1:A:297:HIS:CD2	1:A:298:ILE:HG13	2.55	0.40

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:117:LYS:NZ	6:B:504:EPE:C8	2.84	0.40
1:A:462:HIS:ND1	1:A:463:THR:O	2.48	0.40
1:B:466:TYR:N	1:B:467:PRO:CD	2.84	0.40
1:B:70:ALA:HB1	1:B:331:THR:CG2	2.52	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	466/482 (97%)	445 (96%)	20 (4%)	1 (0%)	47	62
1	B	471/482 (98%)	450 (96%)	18 (4%)	3 (1%)	25	36
All	All	937/964 (97%)	895 (96%)	38 (4%)	4 (0%)	34	48

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	166	PRO
1	B	297	HIS
1	B	295	THR
1	B	168	GLU

### 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	390/400 (98%)	374 (96%)	16 (4%)	30	48
1	B	387/400 (97%)	374 (97%)	13 (3%)	37	56
All	All	777/800 (97%)	748 (96%)	29 (4%)	36	53

All (29) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	53	ILE
1	A	81	ASN
1	A	100	THR
1	A	106	SER
1	A	167	LYS
1	A	175	ASP
1	A	181	ASN
1	A	210	ASP
1	A	240	GLU
1	A	296	LYS
1	A	328	THR
1	A	329	GLN
1	A	397	THR
1	A	403	ASN
1	A	450[A]	SER
1	A	450[B]	SER
1	B	92	THR
1	B	175	ASP
1	B	181	ASN
1	B	229	ARG
1	B	231	MSE
1	B	295	THR
1	B	296	LYS
1	B	328	THR
1	B	349	GLU
1	B	379[A]	TYR
1	B	379[B]	TYR
1	B	403	ASN
1	B	450	SER

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

### 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 35 ligands modelled in this entry, 4 are monoatomic - leaving 31 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	$\# Z  > 2$
4	EDO	B	510	-	3,3,3	0.46	0	2,2,2	0.17	0
4	EDO	B	508	-	3,3,3	0.34	0	2,2,2	0.59	0
4	EDO	A	506	-	3,3,3	0.43	0	2,2,2	0.74	0
4	EDO	A	515	-	3,3,3	0.49	0	2,2,2	0.17	0
4	EDO	A	507	-	3,3,3	0.37	0	2,2,2	0.39	0
4	EDO	A	504	-	3,3,3	0.44	0	2,2,2	0.40	0
4	EDO	B	509	-	3,3,3	0.52	0	2,2,2	0.24	0
4	EDO	A	512	-	3,3,3	0.44	0	2,2,2	0.67	0
5	PEG	A	516	-	6,6,6	0.53	0	5,5,5	0.46	0
4	EDO	B	506	-	3,3,3	0.58	0	2,2,2	0.54	0
4	EDO	B	505	-	3,3,3	0.34	0	2,2,2	0.17	0
4	EDO	A	503	-	3,3,3	0.57	0	2,2,2	0.29	0
4	EDO	A	509	-	3,3,3	0.57	0	2,2,2	0.21	0
4	EDO	B	507	-	3,3,3	0.51	0	2,2,2	0.34	0
4	EDO	B	514	-	3,3,3	0.58	0	2,2,2	0.18	0
4	EDO	A	511	-	3,3,3	0.58	0	2,2,2	0.10	0
4	EDO	B	515	-	3,3,3	0.55	0	2,2,2	0.25	0



Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
6	EPE	B	504	-	15,15,15	0.80	1 (6%)	18,20,20	1.87	4 (22%)
4	EDO	A	502	-	3,3,3	0.44	0	2,2,2	0.24	0
5	PEG	B	517	-	6,6,6	0.40	0	5,5,5	0.40	0
6	EPE	B	502	-	15,15,15	0.90	1 (6%)	18,20,20	1.19	1 (5%)
5	PEG	B	516	-	6,6,6	0.44	0	5,5,5	0.57	0
4	EDO	A	510	-	3,3,3	0.30	0	2,2,2	0.58	0
4	EDO	A	513	-	3,3,3	0.43	0	2,2,2	0.30	0
4	EDO	B	513	-	3,3,3	0.53	0	2,2,2	0.31	0
4	EDO	A	505	-	3,3,3	0.45	0	2,2,2	0.62	0
4	EDO	B	511	-	3,3,3	0.48	0	2,2,2	0.41	0
4	EDO	B	512	-	3,3,3	0.41	0	2,2,2	0.30	0
4	EDO	A	508	-	3,3,3	0.41	0	2,2,2	0.21	0
4	EDO	A	514	-	3,3,3	0.41	0	2,2,2	0.47	0
6	EPE	B	503	-	15,15,15	1.30	1 (6%)	18,20,20	1.60	4 (22%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	EDO	B	510	-	-	1/1/1/1	-
4	EDO	B	508	-	-	1/1/1/1	-
4	EDO	A	506	-	-	1/1/1/1	-
4	EDO	A	515	-	-	0/1/1/1	-
4	EDO	A	507	-	-	0/1/1/1	-
4	EDO	A	504	-	-	1/1/1/1	-
4	EDO	B	509	-	-	0/1/1/1	-
4	EDO	A	512	-	-	1/1/1/1	-
5	PEG	A	516	-	-	2/4/4/4	-
4	EDO	B	506	-	-	1/1/1/1	-
4	EDO	B	505	-	-	0/1/1/1	-
4	EDO	A	503	-	-	0/1/1/1	-
4	EDO	A	509	-	-	1/1/1/1	-
4	EDO	B	507	-	-	1/1/1/1	-
4	EDO	B	514	-	-	1/1/1/1	-
4	EDO	A	511	-	-	1/1/1/1	-
4	EDO	B	515	-	-	1/1/1/1	-
6	EPE	B	504	-	-	7/9/19/19	0/1/1/1
4	EDO	A	502	-	-	1/1/1/1	-
5	PEG	B	517	-	-	2/4/4/4	-

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	EPE	B	502	-	-	3/9/19/19	0/1/1/1
5	PEG	B	516	-	-	1/4/4/4	-
4	EDO	A	510	-	-	1/1/1/1	-
4	EDO	A	513	-	-	1/1/1/1	-
4	EDO	B	513	-	-	1/1/1/1	-
4	EDO	A	505	-	-	0/1/1/1	-
4	EDO	B	511	-	-	0/1/1/1	-
4	EDO	B	512	-	-	1/1/1/1	-
4	EDO	A	508	-	-	1/1/1/1	-
4	EDO	A	514	-	-	1/1/1/1	-
6	EPE	B	503	-	-	1/9/19/19	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	B	503	EPE	C10-S	4.69	1.84	1.77
6	B	502	EPE	C10-S	3.08	1.81	1.77
6	B	504	EPE	C10-S	2.34	1.80	1.77

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	B	504	EPE	O2S-S-C10	4.63	112.49	106.92
6	B	504	EPE	C6-N1-C2	3.97	117.77	108.83
6	B	502	EPE	O3S-S-C10	3.91	112.10	105.77
6	B	503	EPE	O2S-S-C10	3.48	111.10	106.92
6	B	503	EPE	O2S-S-O1S	-3.40	102.20	113.95
6	B	503	EPE	O3S-S-C10	2.91	110.47	105.77
6	B	503	EPE	C6-N1-C2	2.41	114.25	108.83
6	B	504	EPE	O3S-S-C10	2.28	109.45	105.77
6	B	504	EPE	C5-N4-C3	-2.06	104.18	108.83

There are no chirality outliers.

All (34) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	B	504	EPE	C9-C10-S-O1S
6	B	504	EPE	C9-C10-S-O3S
5	B	516	PEG	O1-C1-C2-O2
5	B	517	PEG	O2-C3-C4-O4
5	A	516	PEG	O2-C3-C4-O4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
6	B	504	EPE	N4-C7-C8-O8
4	A	506	EDO	O1-C1-C2-O2
4	A	509	EDO	O1-C1-C2-O2
4	A	514	EDO	O1-C1-C2-O2
5	B	517	PEG	O1-C1-C2-O2
6	B	504	EPE	S-C10-C9-N1
6	B	504	EPE	C10-C9-N1-C2
6	B	504	EPE	C10-C9-N1-C6
6	B	502	EPE	C10-C9-N1-C2
6	B	502	EPE	C10-C9-N1-C6
4	B	507	EDO	O1-C1-C2-O2
4	A	513	EDO	O1-C1-C2-O2
4	B	508	EDO	O1-C1-C2-O2
4	B	506	EDO	O1-C1-C2-O2
4	A	508	EDO	O1-C1-C2-O2
6	B	504	EPE	C9-C10-S-O2S
4	B	513	EDO	O1-C1-C2-O2
4	B	514	EDO	O1-C1-C2-O2
4	B	515	EDO	O1-C1-C2-O2
4	A	502	EDO	O1-C1-C2-O2
5	A	516	PEG	C1-C2-O2-C3
4	B	510	EDO	O1-C1-C2-O2
4	A	504	EDO	O1-C1-C2-O2
4	A	512	EDO	O1-C1-C2-O2
4	A	511	EDO	O1-C1-C2-O2
4	A	510	EDO	O1-C1-C2-O2
4	B	512	EDO	O1-C1-C2-O2
6	B	502	EPE	N4-C7-C8-O8
6	B	503	EPE	C10-C9-N1-C6

There are no ring outliers.

3 monomers are involved in 11 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	507	EDO	1	0
4	B	505	EDO	4	0
6	B	504	EPE	6	0

## 5.7 Other polymers

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 [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	453/482 (93%)	-0.49	3 (0%) 87 86	18, 31, 53, 88	0
1	B	456/482 (94%)	-0.46	5 (1%) 80 79	18, 30, 49, 78	0
All	All	909/964 (94%)	-0.48	8 (0%) 84 82	18, 30, 51, 88	0

All (8) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	166	PRO	5.3
1	A	472	PRO	2.9
1	A	477	ILE	2.6
1	B	299	SER	2.5
1	B	166	PRO	2.5
1	B	89	PRO	2.4
1	B	199	SER	2.1
1	B	62	PRO	2.1

### 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(Å <sup>2</sup> )	Q<0.9
4	EDO	B	515	4/4	0.85	0.24	55,68,72,74	0
5	PEG	A	516	7/7	0.88	0.19	46,57,65,71	0
4	EDO	A	512	4/4	0.89	0.19	42,45,57,58	0
4	EDO	B	513	4/4	0.89	0.18	44,49,55,58	0
6	EPE	B	504	15/15	0.90	0.24	29,55,78,89	15
4	EDO	B	514	4/4	0.91	0.17	52,56,65,71	0
4	EDO	A	513	4/4	0.91	0.21	40,49,54,73	0
4	EDO	B	508	4/4	0.91	0.37	67,70,71,72	0
4	EDO	B	509	4/4	0.92	0.15	48,52,53,67	0
4	EDO	B	506	4/4	0.93	0.13	35,49,53,60	0
4	EDO	A	511	4/4	0.94	0.14	56,57,62,74	0
4	EDO	A	504	4/4	0.94	0.16	50,51,60,71	0
4	EDO	B	510	4/4	0.94	0.15	42,51,59,61	0
5	PEG	B	517	7/7	0.94	0.12	30,32,44,45	7
4	EDO	A	510	4/4	0.94	0.14	46,58,61,62	0
4	EDO	A	509	4/4	0.94	0.23	46,55,61,64	0
4	EDO	A	515	4/4	0.94	0.27	50,51,57,59	0
4	EDO	A	514	4/4	0.94	0.34	51,56,59,60	0
4	EDO	A	506	4/4	0.95	0.21	33,40,62,67	0
5	PEG	B	516	7/7	0.95	0.15	21,33,71,72	0
4	EDO	A	505	4/4	0.95	0.15	35,39,42,43	0
4	EDO	B	511	4/4	0.95	0.14	51,54,73,75	0
4	EDO	A	507	4/4	0.95	0.12	32,40,49,50	0
6	EPE	B	503	15/15	0.96	0.26	43,68,102,104	0
4	EDO	B	505	4/4	0.97	0.13	25,36,42,44	0
6	EPE	B	502	15/15	0.97	0.16	39,64,81,81	0
4	EDO	B	507	4/4	0.97	0.11	41,46,54,67	0
4	EDO	B	512	4/4	0.97	0.12	39,42,54,58	0
4	EDO	A	508	4/4	0.97	0.18	42,51,59,60	0
4	EDO	A	503	4/4	0.97	0.14	24,26,29,40	0
4	EDO	A	502	4/4	0.97	0.11	24,27,31,32	0
3	CL	B	501	1/1	0.99	0.05	29,29,29,29	0
3	CL	A	501	1/1	0.99	0.06	35,35,35,35	0
2	ZN	B	500	1/1	0.99	0.07	35,35,35,35	0
2	ZN	A	500	1/1	1.00	0.05	41,41,41,41	0

## 6.5 Other polymers

There are no such residues in this entry.