



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 07:41 PM GMT

PDB ID : 4POV  
Title : ThiT with LMG135 bound  
Authors : Swier, L.J.Y.M.; Guskov, A.; Slotboom, D.J.  
Deposited on : 2014-02-26  
Resolution : 2.20 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

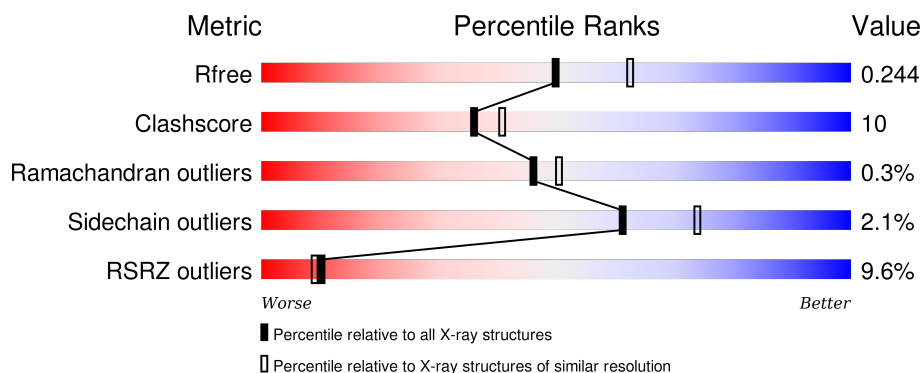
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.20 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	91344	3774 (2.20-2.20)
Clashscore	102246	4477 (2.20-2.20)
Ramachandran outliers	100387	4404 (2.20-2.20)
Sidechain outliers	100360	4405 (2.20-2.20)
RSRZ outliers	91569	3781 (2.20-2.20)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments 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	192	<div> <div>7%</div> <div>79% 14% • 7%</div> </div>
1	B	192	<div> <div>10%</div> <div>84% 8% 8%</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	PG4	A	205	-	-	-	X
4	PG4	A	206	-	-	-	X
4	PG4	A	207	-	-	-	X
4	PG4	B	203	-	-	-	X
4	PG4	B	204	-	-	-	X
4	PG4	B	205	-	-	-	X
5	BNG	A	209	-	-	-	X
5	BNG	A	210	-	-	-	X
5	BNG	B	207	-	-	-	X
5	BNG	B	209	-	-	-	X
6	PEG	A	211	-	-	-	X
6	PEG	A	215	-	-	X	X
6	PEG	B	211	-	-	-	X
6	PEG	B	214	-	-	-	X
7	1PE	B	215	-	-	-	X
7	1PE	B	216	-	-	-	X

## 2 Entry composition

There are 11 unique types of molecules in this entry. The entry contains 3276 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.

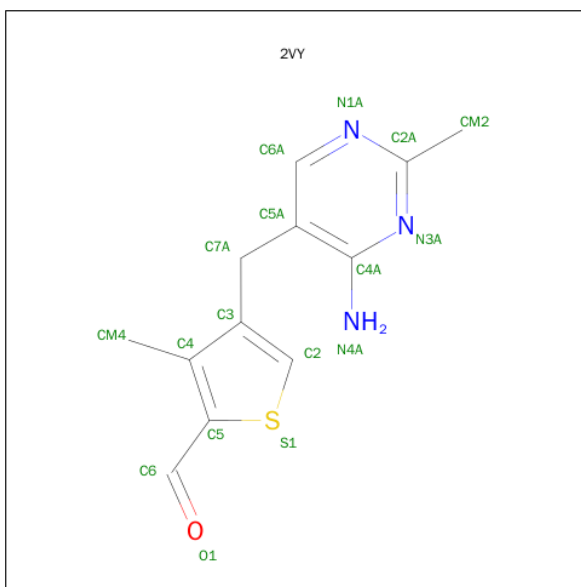
- Molecule 1 is a protein called Thiamine transporter ThiT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	178	Total	C	N	O	S	0	0	0
			1386	951	216	216	3			
1	B	176	Total	C	N	O	S	0	0	0
			1366	936	213	214	3			

There are 20 discrepancies between the modelled and reference sequences:

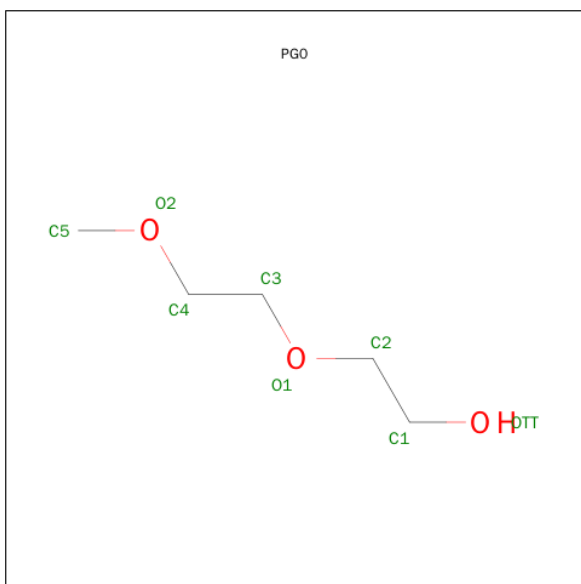
Chain	Residue	Modelled	Actual	Comment	Reference
A	-9	MET	-	INITIATING METHIONINE	UNP A2RI47
A	-8	HIS	-	EXPRESSION TAG	UNP A2RI47
A	-7	HIS	-	EXPRESSION TAG	UNP A2RI47
A	-6	HIS	-	EXPRESSION TAG	UNP A2RI47
A	-5	HIS	-	EXPRESSION TAG	UNP A2RI47
A	-4	HIS	-	EXPRESSION TAG	UNP A2RI47
A	-3	HIS	-	EXPRESSION TAG	UNP A2RI47
A	-2	HIS	-	EXPRESSION TAG	UNP A2RI47
A	-1	HIS	-	EXPRESSION TAG	UNP A2RI47
A	0	ALA	-	EXPRESSION TAG	UNP A2RI47
B	-9	MET	-	INITIATING METHIONINE	UNP A2RI47
B	-8	HIS	-	EXPRESSION TAG	UNP A2RI47
B	-7	HIS	-	EXPRESSION TAG	UNP A2RI47
B	-6	HIS	-	EXPRESSION TAG	UNP A2RI47
B	-5	HIS	-	EXPRESSION TAG	UNP A2RI47
B	-4	HIS	-	EXPRESSION TAG	UNP A2RI47
B	-3	HIS	-	EXPRESSION TAG	UNP A2RI47
B	-2	HIS	-	EXPRESSION TAG	UNP A2RI47
B	-1	HIS	-	EXPRESSION TAG	UNP A2RI47
B	0	ALA	-	EXPRESSION TAG	UNP A2RI47

- Molecule 2 is 4-[(4-AMINO-2-METHYLPYRIMIDIN-5-YL)METHYL]-3-METHYLTHIOPHENE-2-CARBALDEHYDE (three-letter code: 2VY) (formula: C<sub>12</sub>H<sub>13</sub>N<sub>3</sub>OS).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	S	0	0
			17	12	3	1	1		
2	B	1	Total	C	N	O	S	0	0
			17	12	3	1	1		

- Molecule 3 is 2-(2-METHOXYETHOXY)ETHANOL (three-letter code: PG0) (formula: C<sub>5</sub>H<sub>12</sub>O<sub>3</sub>).



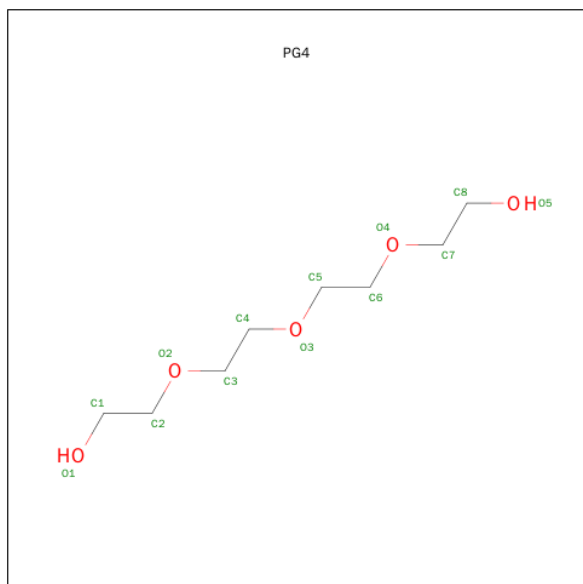
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			8	5	3		

*Continued on next page...*

*Continued from previous page...*

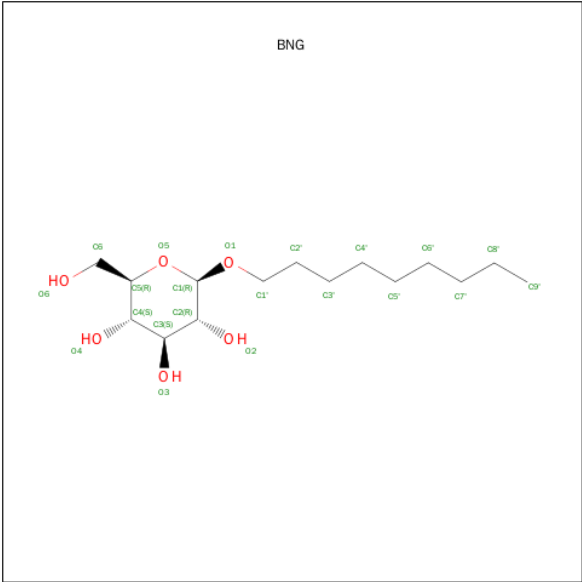
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	B	1	Total	C	O	0	0
			8	5	3		

- Molecule 4 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula:  $C_8H_{18}O_5$ ).



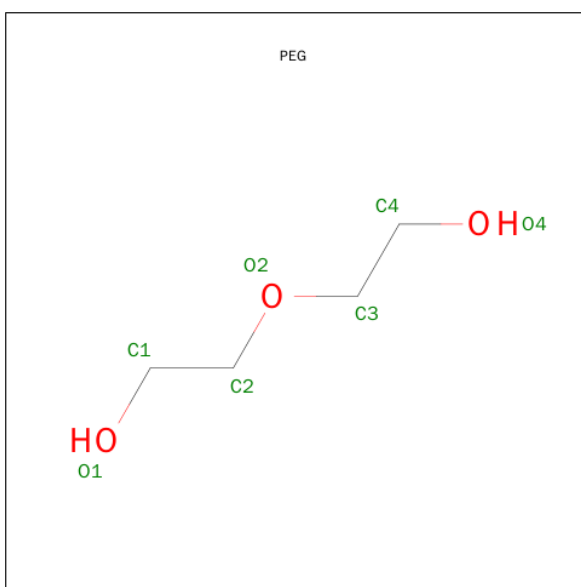
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			13	8	5		
4	A	1	Total	C	O	0	0
			13	8	5		
4	A	1	Total	C	O	0	0
			13	8	5		
4	A	1	Total	C	O	0	0
			13	8	5		
4	A	1	Total	C	O	0	0
			13	8	5		
4	B	1	Total	C	O	0	0
			13	8	5		
4	B	1	Total	C	O	0	0
			13	8	5		
4	B	1	Total	C	O	0	0
			13	8	5		
4	B	1	Total	C	O	0	0
			13	8	5		

- Molecule 5 is SUGAR (B-NONYLGLUCOSIDE) (three-letter code: BNG) (formula:  $C_{15}H_{30}O_6$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			21	15	6		
5	A	1	Total	C	O	0	0
			21	15	6		
5	A	1	Total	C	O	0	0
			21	15	6		
5	B	1	Total	C	O	0	0
			21	15	6		
5	B	1	Total	C	O	0	0
			21	15	6		
5	B	1	Total	C	O	0	0
			21	15	6		

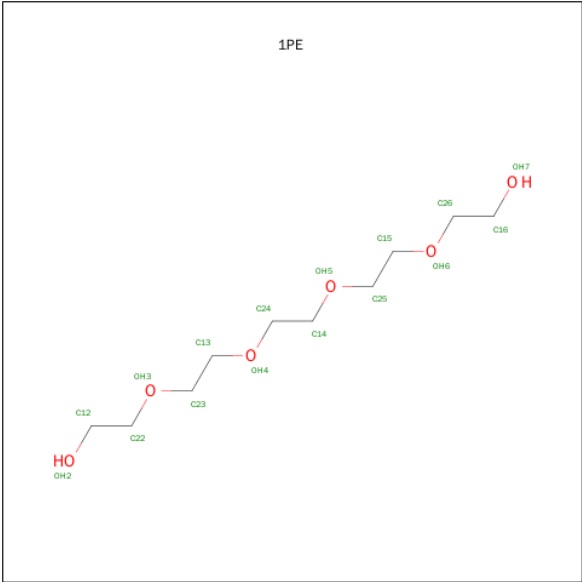
- Molecule 6 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: C<sub>4</sub>H<sub>10</sub>O<sub>3</sub>).



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

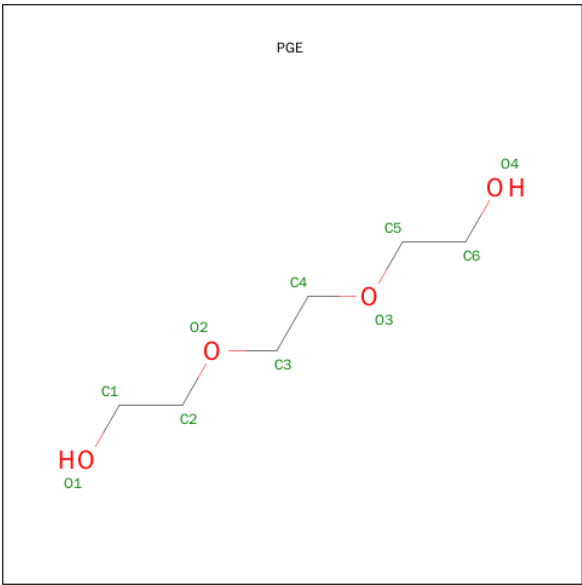
- Molecule 7 is PENTAETHYLENE GLYCOL (three-letter code: 1PE) (formula: C<sub>10</sub>H<sub>22</sub>O<sub>6</sub>).





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

- Molecule 8 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula: C<sub>6</sub>H<sub>14</sub>O<sub>4</sub>).



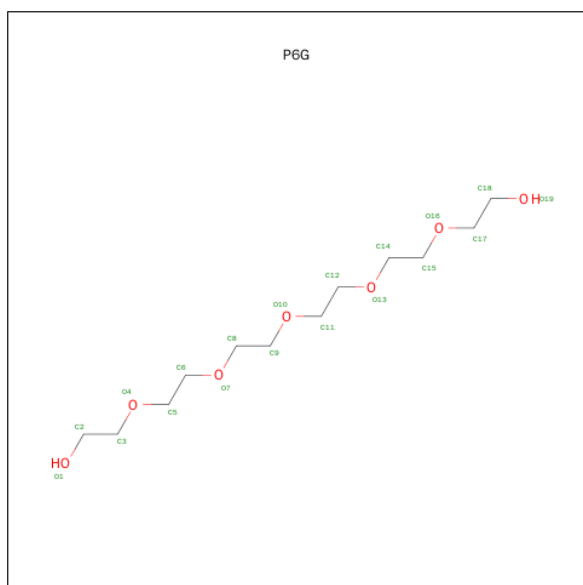
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	A	1	Total	C	O	0	0
			10	6	4		

Continued on next page...

*Continued from previous page...*

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

- Molecule 9 is HEXAETHYLENE GLYCOL (three-letter code: P6G) (formula: C<sub>12</sub>H<sub>26</sub>O<sub>7</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
9	A	1	Total	C	O	0	0
			19	12	7		

- Molecule 10 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

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

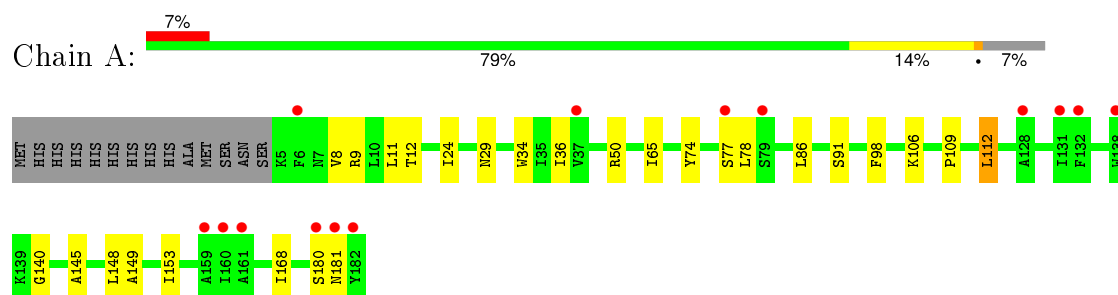
- Molecule 11 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
11	A	33	Total	O	0	0
			33	33		
11	B	30	Total	O	0	0
			30	30		

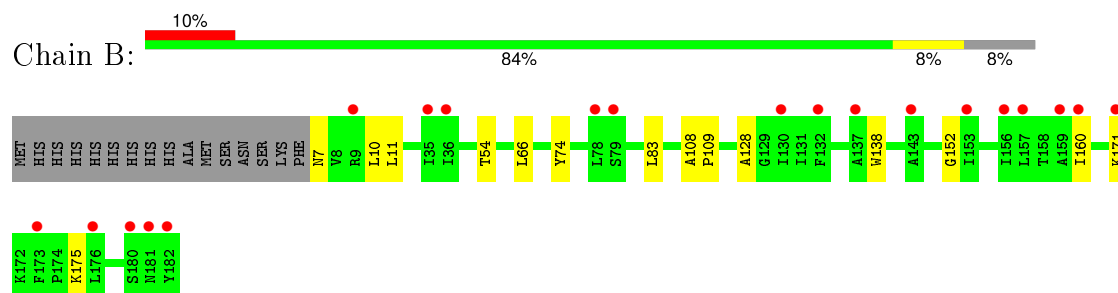
### 3 Residue-property plots [i](#)

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: Thiamine transporter ThiT



#### • Molecule 1: Thiamine transporter ThiT



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	62.72Å 84.35Å 127.25Å 90.00° 96.08° 90.00°	Depositor
Resolution (Å)	48.04 – 2.20 48.04 – 2.20	Depositor EDS
% Data completeness (in resolution range)	(Not available) (48.04-2.20) 99.4 (48.04-2.20)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.19 (at 2.20Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.2_1309)	Depositor
R, $R_{free}$	0.185 , 0.237 0.193 , 0.244	Depositor DCC
$R_{free}$ test set	1668 reflections (5.26%)	DCC
Wilson B-factor (Å <sup>2</sup> )	45.0	Xtriage
Anisotropy	0.377	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 66.5	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	1 of 33360 reflections (0.003%)	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	3276	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	60.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.20% 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: PGE, P6G, CL, 1PE, PG4, PG0, 2VY, PEG, BNG

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	0.67	0/1425	0.63	0/1941
1	B	0.62	0/1404	0.62	0/1914
All	All	0.64	0/2829	0.62	0/3855

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1386	0	1484	29	0
1	B	1366	0	1462	11	0
2	A	17	0	13	2	0
2	B	17	0	13	0	0
3	A	8	0	12	1	0
3	B	8	0	12	2	0
4	A	65	0	90	11	0
4	B	52	0	72	12	0
5	A	63	0	90	12	0
5	B	63	0	90	7	0
6	A	35	0	46	6	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	B	35	0	45	5	0
7	A	16	0	22	4	0
7	B	32	0	44	0	0
8	A	30	0	42	3	0
9	A	19	0	26	3	0
10	B	1	0	0	0	0
11	A	33	0	0	1	0
11	B	30	0	0	2	0
All	All	3276	0	3563	65	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 10.

All (65) 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)
9:A:220:P6G:H32	9:A:220:P6G:H82	1.35	1.09
4:B:206:PG4:H51	5:B:209:BNG:H3'1	1.52	0.88
7:A:216:1PE:H252	4:B:204:PG4:H22	1.58	0.85
5:A:208:BNG:H9'3	5:A:208:BNG:H5'1	1.61	0.81
1:B:160:ILE:HG22	6:B:210:PEG:H12	1.65	0.79
9:A:220:P6G:C3	9:A:220:P6G:H82	2.12	0.76
4:B:206:PG4:C5	5:B:209:BNG:H3'1	2.15	0.76
1:A:109:PRO:HG3	8:A:217:PGE:H5	1.68	0.75
1:A:74:TYR:CE2	4:A:207:PG4:H22	2.22	0.74
4:A:204:PG4:H31	4:B:206:PG4:H52	1.71	0.72
4:B:206:PG4:H51	5:B:209:BNG:C3'	2.21	0.70
1:A:78:LEU:CD2	5:A:209:BNG:H4'2	2.24	0.67
1:A:98:PHE:HA	8:A:219:PGE:H52	1.78	0.64
5:B:207:BNG:H2'2	5:B:207:BNG:O5	1.99	0.63
4:A:204:PG4:C3	4:B:206:PG4:H52	2.33	0.58
1:A:77:SER:HA	5:A:209:BNG:H61	1.87	0.57
1:A:112:LEU:HD12	8:A:217:PGE:H12	1.85	0.56
1:B:160:ILE:CG2	6:B:210:PEG:H12	2.33	0.56
1:A:78:LEU:HD22	5:A:209:BNG:H4'2	1.87	0.55
1:A:153:ILE:HD13	6:B:211:PEG:H11	1.88	0.55
1:B:175:LYS:HB2	6:B:214:PEG:H32	1.88	0.55
1:B:74:TYR:CE2	3:B:202:PG0:H42	2.41	0.55
4:B:203:PG4:H62	11:B:311:HOH:O	2.08	0.54
1:A:65:ILE:HG23	5:A:210:BNG:H9'3	1.89	0.53
1:A:168:ILE:HG12	6:A:212:PEG:H21	1.91	0.53

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:140:GLY:H	3:A:202:PG0:H21	1.73	0.53
4:B:206:PG4:H51	5:B:209:BNG:C2'	2.38	0.53
1:A:36:ILE:HG21	4:A:207:PG4:H62	1.92	0.52
5:A:208:BNG:H1'1	5:A:208:BNG:O2	2.10	0.52
1:A:149:ALA:HA	7:A:216:1PE:H131	1.92	0.51
1:A:24:ILE:CD1	4:A:203:PG4:H61	2.40	0.51
1:A:106:LYS:HB3	6:A:215:PEG:H21	1.92	0.50
1:A:78:LEU:HD23	5:A:209:BNG:H4'2	1.93	0.49
3:B:202:PG0:H12	11:B:305:HOH:O	2.12	0.49
4:A:206:PG4:H71	4:A:206:PG4:H51	1.66	0.47
1:A:153:ILE:HG21	6:B:211:PEG:H11	1.96	0.47
4:A:207:PG4:H82	11:A:311:HOH:O	2.15	0.47
1:A:91:SER:HA	4:A:205:PG4:H22	1.97	0.47
1:A:148:LEU:HD23	7:A:216:1PE:H222	1.97	0.47
1:B:83:LEU:HD23	1:B:128:ALA:HB2	1.95	0.47
1:B:152:GLY:HA2	4:B:204:PG4:H52	1.98	0.46
1:B:10:LEU:HD21	4:B:203:PG4:H61	1.98	0.46
1:A:86:LEU:HD12	5:A:209:BNG:H7'1	1.97	0.45
5:A:208:BNG:H6'2	5:A:208:BNG:H3'2	1.71	0.45
1:A:8:VAL:O	1:A:12:THR:HG23	2.16	0.45
7:A:216:1PE:H241	7:A:216:1PE:H232	1.71	0.44
1:B:66:LEU:HD11	5:B:207:BNG:H1'1	1.99	0.44
6:A:215:PEG:C4	6:A:215:PEG:H12	2.39	0.44
4:B:204:PG4:H62	4:B:204:PG4:H41	1.65	0.44
9:A:220:P6G:H181	9:A:220:P6G:H142	2.01	0.43
1:A:34:TRP:CG	2:A:201:2VY:H15	2.54	0.43
5:A:210:BNG:H9'2	5:A:210:BNG:H6'2	1.66	0.42
1:A:24:ILE:HD13	4:A:203:PG4:H61	2.01	0.42
1:A:36:ILE:CG2	4:A:207:PG4:H62	2.49	0.42
5:A:208:BNG:H5'1	5:A:208:BNG:C9'	2.30	0.42
1:A:50:ARG:NH2	6:A:213:PEG:H32	2.35	0.41
1:A:29:ASN:HA	4:A:204:PG4:H71	2.01	0.41
1:B:54:THR:HB	4:B:203:PG4:H72	2.02	0.41
6:A:215:PEG:H32	6:A:215:PEG:H12	1.36	0.41
6:A:215:PEG:H42	6:A:215:PEG:H12	2.03	0.41
1:B:108:ALA:CB	5:B:209:BNG:H2'2	2.51	0.41
1:A:145:ALA:HB2	5:A:208:BNG:H3	2.03	0.41
1:A:180:SER:OG	1:A:181:ASN:O	2.37	0.41
1:A:34:TRP:CD1	2:A:201:2VY:H15	2.56	0.40
1:B:108:ALA:HB3	1:B:109:PRO:HD3	2.04	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	176/192 (92%)	173 (98%)	3 (2%)	0	100	100
1	B	174/192 (91%)	170 (98%)	3 (2%)	1 (1%)	30	29
All	All	350/384 (91%)	343 (98%)	6 (2%)	1 (0%)	46	50

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	138	TRP

### 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	142/155 (92%)	139 (98%)	3 (2%)	61	74
1	B	140/155 (90%)	137 (98%)	3 (2%)	61	74
All	All	282/310 (91%)	276 (98%)	6 (2%)	61	74

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	9	ARG
1	A	11	LEU
1	A	112	LEU
1	B	7	ASN
1	B	11	LEU

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	B	171	LYS

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 37 ligands modelled in this entry, 1 is monoatomic - leaving 36 for Mogul analysis.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# $ Z  > 2$	Counts	RMSZ	# $ Z  > 2$
2	2VY	A	201	-	16,18,18	3.27	5 (31%)	17,25,25	3.10	7 (41%)
3	PG0	A	202	-	7,7,7	0.38	0	6,6,6	0.96	1 (16%)
4	PG4	A	203	-	12,12,12	0.72	0	11,11,11	0.38	0
4	PG4	A	204	-	12,12,12	0.81	0	11,11,11	0.24	0
4	PG4	A	205	-	12,12,12	0.78	0	11,11,11	0.42	0
4	PG4	A	206	-	12,12,12	0.74	0	11,11,11	0.44	0
4	PG4	A	207	-	12,12,12	0.83	0	11,11,11	0.54	0
5	BNG	A	208	-	21,21,21	1.11	2 (9%)	26,26,26	1.15	1 (3%)
5	BNG	A	209	-	21,21,21	1.14	3 (14%)	26,26,26	1.03	1 (3%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
5	BNG	A	210	-	21,21,21	1.10	2 (9%)	26,26,26	0.78	0
6	PEG	A	211	-	6,6,6	0.60	0	5,5,5	0.31	0
6	PEG	A	212	-	6,6,6	0.56	0	5,5,5	0.25	0
6	PEG	A	213	-	6,6,6	0.55	0	5,5,5	0.25	0
6	PEG	A	214	-	6,6,6	0.56	0	5,5,5	0.33	0
6	PEG	A	215	-	6,6,6	0.68	0	5,5,5	0.42	0
7	1PE	A	216	-	15,15,15	0.80	0	14,14,14	0.41	0
8	PGE	A	217	-	9,9,9	0.37	0	8,8,8	0.29	0
8	PGE	A	218	-	9,9,9	0.30	0	8,8,8	0.34	0
8	PGE	A	219	-	9,9,9	0.33	0	8,8,8	0.39	0
9	P6G	A	220	-	18,18,18	0.86	0	17,17,17	0.37	0
2	2VY	B	201	-	16,18,18	3.53	4 (25%)	17,25,25	2.42	7 (41%)
3	PG0	B	202	-	7,7,7	0.24	0	6,6,6	0.67	0
4	PG4	B	203	-	12,12,12	0.71	0	11,11,11	0.33	0
4	PG4	B	204	-	12,12,12	0.76	0	11,11,11	0.34	0
4	PG4	B	205	-	12,12,12	0.71	0	11,11,11	0.41	0
4	PG4	B	206	-	12,12,12	0.78	0	11,11,11	0.51	0
5	BNG	B	207	-	21,21,21	1.10	1 (4%)	26,26,26	0.91	1 (3%)
5	BNG	B	208	-	21,21,21	1.02	1 (4%)	26,26,26	1.12	2 (7%)
5	BNG	B	209	-	21,21,21	1.16	2 (9%)	26,26,26	1.44	3 (11%)
6	PEG	B	210	-	6,6,6	0.58	0	5,5,5	0.48	0
6	PEG	B	211	-	6,6,6	0.56	0	5,5,5	0.42	0
6	PEG	B	212	-	6,6,6	0.55	0	5,5,5	0.29	0
6	PEG	B	213	-	6,6,6	0.55	0	5,5,5	0.31	0
6	PEG	B	214	-	6,6,6	0.57	0	5,5,5	0.22	0
7	1PE	B	215	-	15,15,15	0.77	0	14,14,14	0.38	0
7	1PE	B	216	-	15,15,15	0.77	0	14,14,14	0.35	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	2VY	A	201	-	-	0/3/6/6	0/2/2/2
3	PG0	A	202	-	-	0/5/5/5	0/0/0/0
4	PG4	A	203	-	-	0/10/10/10	0/0/0/0
4	PG4	A	204	-	-	0/10/10/10	0/0/0/0
4	PG4	A	205	-	-	0/10/10/10	0/0/0/0
4	PG4	A	206	-	-	0/10/10/10	0/0/0/0
4	PG4	A	207	-	-	0/10/10/10	0/0/0/0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	BNG	A	208	-	-	0/12/32/32	0/1/1/1
5	BNG	A	209	-	-	0/12/32/32	0/1/1/1
5	BNG	A	210	-	-	0/12/32/32	0/1/1/1
6	PEG	A	211	-	-	0/4/4/4	0/0/0/0
6	PEG	A	212	-	-	0/4/4/4	0/0/0/0
6	PEG	A	213	-	-	0/4/4/4	0/0/0/0
6	PEG	A	214	-	-	0/4/4/4	0/0/0/0
6	PEG	A	215	-	-	0/4/4/4	0/0/0/0
7	1PE	A	216	-	-	0/13/13/13	0/0/0/0
8	PGE	A	217	-	-	0/7/7/7	0/0/0/0
8	PGE	A	218	-	-	0/7/7/7	0/0/0/0
8	PGE	A	219	-	-	0/7/7/7	0/0/0/0
9	P6G	A	220	-	-	0/16/16/16	0/0/0/0
2	2VY	B	201	-	-	0/3/6/6	0/2/2/2
3	PG0	B	202	-	-	0/5/5/5	0/0/0/0
4	PG4	B	203	-	-	0/10/10/10	0/0/0/0
4	PG4	B	204	-	-	0/10/10/10	0/0/0/0
4	PG4	B	205	-	-	0/10/10/10	0/0/0/0
4	PG4	B	206	-	-	0/10/10/10	0/0/0/0
5	BNG	B	207	-	-	0/12/32/32	0/1/1/1
5	BNG	B	208	-	-	0/12/32/32	0/1/1/1
5	BNG	B	209	-	-	0/12/32/32	0/1/1/1
6	PEG	B	210	-	-	0/4/4/4	0/0/0/0
6	PEG	B	211	-	-	0/4/4/4	0/0/0/0
6	PEG	B	212	-	-	0/4/4/4	0/0/0/0
6	PEG	B	213	-	-	0/4/4/4	0/0/0/0
6	PEG	B	214	-	-	0/4/4/4	0/0/0/0
7	1PE	B	215	-	-	0/13/13/13	0/0/0/0
7	1PE	B	216	-	-	0/13/13/13	0/0/0/0

All (20) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	209	BNG	O1-C1	2.10	1.43	1.40
5	A	209	BNG	O5-C5	2.10	1.49	1.44
5	A	208	BNG	O5-C5	2.11	1.49	1.44
5	A	210	BNG	O1-C1	2.26	1.44	1.40
2	A	201	2VY	C6A-C5A	2.27	1.42	1.37
5	B	209	BNG	O1-C1	2.56	1.44	1.40
5	B	207	BNG	O5-C1	2.58	1.48	1.41
5	B	208	BNG	O5-C1	2.64	1.48	1.41
2	A	201	2VY	C4A-N4A	2.90	1.41	1.34

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	210	BNG	O5-C1	2.92	1.49	1.41
2	A	201	2VY	C5-C6	2.96	1.52	1.48
5	A	208	BNG	O5-C1	2.98	1.49	1.41
5	A	209	BNG	O5-C1	3.00	1.49	1.41
5	B	209	BNG	O5-C1	3.10	1.49	1.41
2	B	201	2VY	C4A-N4A	3.11	1.41	1.34
2	B	201	2VY	C5-C6	3.86	1.53	1.48
2	A	201	2VY	O1-C6	5.50	1.38	1.21
2	B	201	2VY	O1-C6	5.85	1.40	1.21
2	A	201	2VY	C2-C3	10.77	1.43	1.37
2	B	201	2VY	C2-C3	11.32	1.44	1.37

All (23) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	201	2VY	O1-C6-C5	-9.00	106.90	125.13
2	B	201	2VY	O1-C6-C5	-6.07	112.83	125.13
2	A	201	2VY	C3-C2-S1	-5.00	106.93	112.26
2	B	201	2VY	C5A-C6A-N1A	-3.79	117.29	123.86
2	A	201	2VY	C5A-C6A-N1A	-3.55	117.70	123.86
2	B	201	2VY	CM4-C4-C3	-2.74	119.51	125.24
5	A	209	BNG	C6-C5-C4	-2.39	107.13	113.02
2	A	201	2VY	CM4-C4-C3	-2.37	120.27	125.24
2	A	201	2VY	C7A-C5A-C6A	-2.10	118.60	121.72
2	B	201	2VY	N1A-C2A-N3A	-2.07	121.77	125.60
3	A	202	PG0	OTT-C1-C2	-2.04	99.43	112.03
2	B	201	2VY	C6A-C5A-C4A	2.01	118.61	115.72
5	B	208	BNG	C1-C2-C3	2.05	114.02	109.97
5	B	208	BNG	O5-C1-C2	2.05	114.49	110.28
5	B	207	BNG	O1-C1-C2	2.06	110.64	108.04
2	A	201	2VY	C6A-C5A-C4A	2.39	119.15	115.72
2	B	201	2VY	CM2-C2A-N1A	2.42	119.93	117.03
5	B	209	BNG	C1-C2-C3	2.77	115.44	109.97
5	B	209	BNG	C1-O5-C5	2.90	119.38	113.75
5	A	208	BNG	O5-C5-C4	3.06	115.42	109.68
2	A	201	2VY	C6A-N1A-C2A	3.55	121.97	115.77
2	B	201	2VY	C6A-N1A-C2A	3.58	122.03	115.77
5	B	209	BNG	O5-C1-C2	3.65	117.77	110.28

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

26 monomers are involved in 61 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	201	2VY	2	0
3	A	202	PG0	1	0
4	A	203	PG4	2	0
4	A	204	PG4	3	0
4	A	205	PG4	1	0
4	A	206	PG4	1	0
4	A	207	PG4	4	0
5	A	208	BNG	5	0
5	A	209	BNG	5	0
5	A	210	BNG	2	0
6	A	212	PEG	1	0
6	A	213	PEG	1	0
6	A	215	PEG	4	0
7	A	216	1PE	4	0
8	A	217	PGE	2	0
8	A	219	PGE	1	0
9	A	220	P6G	3	0
3	B	202	PG0	2	0
4	B	203	PG4	3	0
4	B	204	PG4	3	0
4	B	206	PG4	6	0
5	B	207	BNG	2	0
5	B	209	BNG	5	0
6	B	210	PEG	2	0
6	B	211	PEG	2	0
6	B	214	PEG	1	0

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data

### 6.1 Protein, DNA and RNA chains

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	178/192 (92%)	0.61	14 (7%) 15 15	34, 48, 78, 111	0
1	B	176/192 (91%)	0.67	20 (11%) 7 6	35, 52, 83, 140	0
All	All	354/384 (92%)	0.64	34 (9%) 10 9	34, 50, 82, 140	0

All (34) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	182	TYR	6.2
1	B	182	TYR	5.9
1	A	79	SER	5.1
1	B	180	SER	4.6
1	B	79	SER	4.5
1	B	9	ARG	4.4
1	A	181	ASN	4.3
1	A	131	ILE	4.3
1	A	132	PHE	4.0
1	B	78	LEU	3.9
1	A	180	SER	3.9
1	B	132	PHE	3.6
1	B	35	ILE	3.2
1	B	130	ILE	3.2
1	A	159	ALA	3.1
1	A	160	ILE	3.0
1	B	156	ILE	2.9
1	B	159	ALA	2.9
1	B	181	ASN	2.8
1	A	77	SER	2.6
1	B	137	ALA	2.6
1	B	176	LEU	2.5
1	A	128	ALA	2.5
1	B	171	LYS	2.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	153	ILE	2.4
1	B	160	ILE	2.4
1	B	157	LEU	2.3
1	A	6	PHE	2.3
1	A	138	TRP	2.2
1	B	173	PHE	2.2
1	B	143	ALA	2.2
1	A	37	VAL	2.2
1	B	36	ILE	2.1
1	A	161	ALA	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. 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
4	PG4	A	205	13/13	0.73	0.24	9.92	73,91,102,103	0
5	BNG	B	207	21/21	0.61	0.32	8.37	64,110,154,163	0
7	1PE	B	215	16/16	0.76	0.36	5.95	62,95,104,105	0
6	PEG	A	215	7/7	0.85	0.28	5.79	47,54,80,101	0
4	PG4	A	206	13/13	0.82	0.26	4.82	53,93,117,119	0
5	BNG	B	209	21/21	0.68	0.22	4.09	70,102,126,149	0
5	BNG	A	210	21/21	0.67	0.24	3.86	58,126,137,164	0
4	PG4	B	204	13/13	0.80	0.37	3.29	60,72,168,171	0
5	BNG	A	209	21/21	0.73	0.36	3.06	60,114,150,169	0
4	PG4	B	205	13/13	0.74	0.21	2.83	75,97,110,119	0
6	PEG	A	211	7/7	0.82	0.21	2.76	66,79,93,100	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
4	PG4	A	207	13/13	0.74	0.23	2.53	65,81,107,124	0
6	PEG	B	214	7/7	0.63	0.37	2.34	75,90,99,104	0
6	PEG	B	211	7/7	0.84	0.28	2.32	79,87,102,114	0
4	PG4	B	203	13/13	0.80	0.17	2.27	88,102,121,128	0
7	1PE	B	216	16/16	0.70	0.25	2.12	85,104,123,128	0
10	CL	B	217	1/1	0.95	0.29	1.97	74,74,74,74	0
6	PEG	B	210	7/7	0.87	0.30	1.93	65,80,89,98	0
5	BNG	B	208	21/21	0.84	0.26	1.85	53,65,118,128	0
8	PGE	A	217	10/10	0.80	0.19	1.57	66,81,87,91	0
6	PEG	A	212	7/7	0.73	0.25	1.47	69,81,96,98	0
8	PGE	A	218	10/10	0.66	0.18	1.31	76,92,108,114	0
4	PG4	A	204	13/13	0.60	0.24	1.17	69,97,117,117	0
4	PG4	A	203	13/13	0.85	0.20	0.89	71,86,98,107	0
4	PG4	B	206	13/13	0.68	0.19	0.87	68,81,115,120	0
6	PEG	B	213	7/7	0.83	0.23	0.66	80,90,94,98	0
3	PG0	A	202	8/8	0.88	0.19	0.45	32,62,79,81	0
7	1PE	A	216	16/16	0.77	0.20	0.42	59,71,82,95	0
3	PG0	B	202	8/8	0.89	0.17	0.28	53,59,76,80	0
2	2VY	B	201	17/17	0.94	0.15	-0.08	32,39,57,61	0
2	2VY	A	201	17/17	0.95	0.13	-0.45	34,40,49,51	0
6	PEG	A	213	7/7	0.57	0.48	-	71,100,116,123	0
8	PGE	A	219	10/10	0.73	0.29	-	71,93,104,107	0
6	PEG	A	214	7/7	0.83	0.21	-	70,89,96,104	0
6	PEG	B	212	7/7	0.90	0.14	-	69,74,89,91	0
9	P6G	A	220	19/19	0.56	0.24	-	63,97,140,142	19
5	BNG	A	208	21/21	0.50	0.39	-	84,117,148,166	0

## 6.5 Other polymers

There are no such residues in this entry.