



# Full wwPDB X-ray Structure Validation Report

Feb 28, 2014 – 03:43 PM GMT

PDB ID : 4C7R  
Title : Inward facing conformation of the trimeric betaine transporter BetP in complex with lipids  
Authors : Koshy, C.; Yildiz, O.; Ziegler, C.  
Deposited on : 2013-09-24  
Resolution : 2.70 Å(reported)

This is a full wwPDB 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/ValidationPDFNotes.html>

---

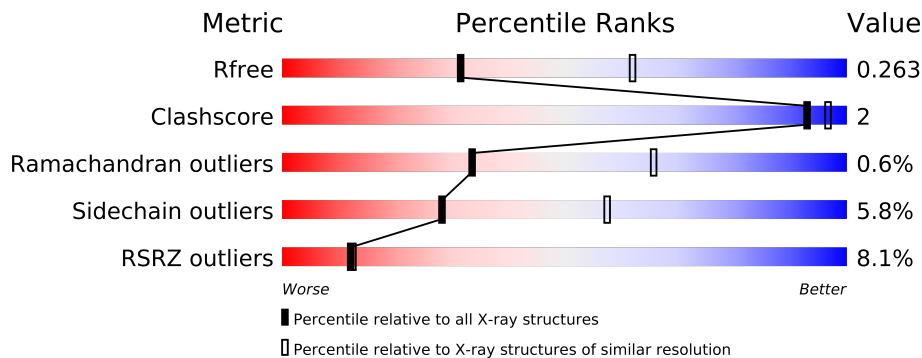
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : dev-1323  
EDS : stable22639  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 2.70 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))
$R_{free}$	66092	1557 (2.70-2.70)
Clashscore	79885	1939 (2.70-2.70)
Ramachandran outliers	78287	1905 (2.70-2.70)
Sidechain outliers	78261	1905 (2.70-2.70)
RSRZ outliers	66119	1559 (2.70-2.70)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	566	
1	B	566	
1	C	566	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
2	FLC	A	1001	-	X
2	FLC	C	1001	-	X
2	FLC	C	1569	-	X
3	CL	A	1587	-	X
4	PGT	A	1588	-	X

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Geometry	Electron density
4	PGT	A	1591	-	X
4	PGT	C	1565	-	X
4	PGT	C	1566	-	X
4	PGT	C	1567	-	X
5	PEG	B	1555	-	X

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 12208 atoms, of which 0 are hydrogen and 0 are deuterium.

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.

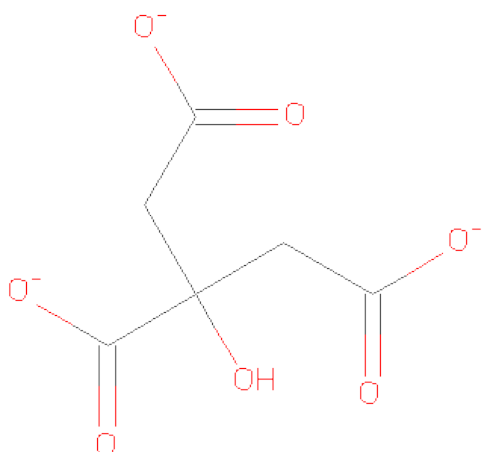
- Molecule 1 is a protein called GLYCINE BETAINES TRANSPORTER BETP.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	531	Total	C	N	O	S	0	0	0
			4057	2659	675	707	16			
1	B	497	Total	C	N	O	S	0	0	0
			3742	2473	595	658	16			
1	C	506	Total	C	N	O	S	0	0	0
			3827	2525	613	673	16			

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	44	ALA	GLU	ENGINEERED MUTATION	UNP P54582
A	45	ALA	GLU	ENGINEERED MUTATION	UNP P54582
A	46	ALA	GLU	ENGINEERED MUTATION	UNP P54582
B	44	ALA	GLU	ENGINEERED MUTATION	UNP P54582
B	45	ALA	GLU	ENGINEERED MUTATION	UNP P54582
B	46	ALA	GLU	ENGINEERED MUTATION	UNP P54582
C	44	ALA	GLU	ENGINEERED MUTATION	UNP P54582
C	45	ALA	GLU	ENGINEERED MUTATION	UNP P54582
C	46	ALA	GLU	ENGINEERED MUTATION	UNP P54582

- Molecule 2 is CITRATE ANION (three-letter code: FLC) (formula:  $C_6H_5O_7$ ).

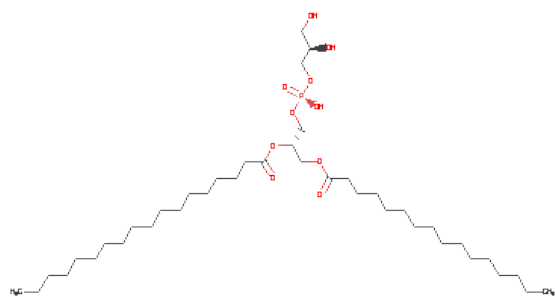


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			13	6	7		
2	B	1	Total	C	O	0	0
			13	6	7		
2	C	1	Total	C	O	0	0
			13	6	7		
2	C	1	Total	C	O	0	0
			13	6	7		

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

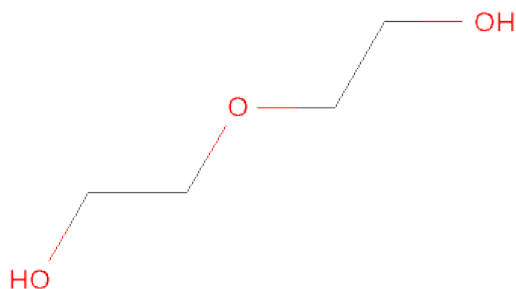
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Cl	0	0
			1	1		
3	A	1	Total	Cl	0	0
			1	1		
3	C	3	Total	Cl	0	0
			3	3		

- Molecule 4 is (1S)-2-{{[(2R)-2,3-DIHYDROXYPROPYL]OXY}(HYDROXY)PHOSPHORYL]OXY}-1-[(PALMITOYLOXY)METHYL]ETHYLSTEARATE (three-letter code: PGT) (formula: C<sub>40</sub>H<sub>79</sub>O<sub>10</sub>P).



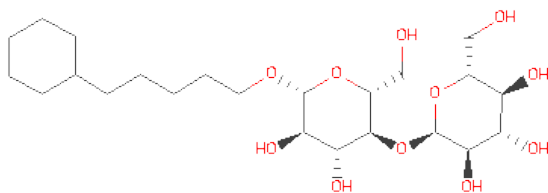
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	O	P	0	0
			51	40	10	1		
4	A	1	Total	C	O	P	0	0
			51	40	10	1		
4	A	1	Total	C	O	P	0	0
			51	40	10	1		
4	A	1	Total	C			0	0
			18	18				
4	A	1	Total	C	O	P	0	0
			51	40	10	1		
4	B	1	Total	C	O	P	0	0
			51	40	10	1		
4	C	1	Total	C	O	P	0	0
			51	40	10	1		
4	C	1	Total	C	O	P	0	0
			51	40	10	1		
4	C	1	Total	C	O	P	0	0
			51	40	10	1		

- 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	B	1	Total	C	O	0	0
			7	4	3		

- Molecule 6 is 5-CYCLOHEXYL-1-PENTYL-BETA-D-MALTOSIDE (three-letter code: CM5) (formula:  $C_{23}H_{42}O_{11}$ ).

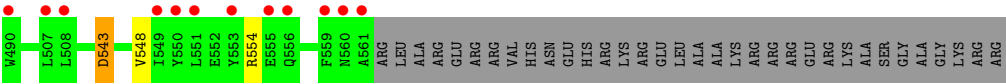


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	C	1	Total	C	O	0	0
			34	23	11		

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	13	Total 13	O 13	0	0
7	B	23	Total 23	O 23	0	0
7	C	22	Total 22	O 22	0	0





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	115.56Å 129.50Å 167.95Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.42 – 2.70 47.60 – 2.65	Depositor EDS
% Data completeness (in resolution range)	99.8 (29.42-2.70) 99.7 (47.60-2.65)	Depositor EDS
$R_{merge}$	0.20	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.77 (at 2.65Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, $R_{free}$	0.212 , 0.266 0.212 , 0.263	Depositor DCC
$R_{free}$ test set	3672 reflections (5.00%)	DCC
Wilson B-factor (Å <sup>2</sup> )	67.4	Xtriage
Anisotropy	0.592	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 80.7	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 73561 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	12208	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	105.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.96% of the height of the origin peak. No significant pseudotranslation is detected.*

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

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: PGT, PEG, CM5, FLC, 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	0.24	0/4159	0.39	0/5661
1	B	0.24	0/3840	0.40	0/5241
1	C	0.26	0/3927	0.41	0/5357
All	All	0.25	0/11926	0.40	0/16259

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	4057	0	0	7	0
1	B	3742	0	0	6	0
1	C	3827	0	0	7	0
2	A	13	0	0	0	0
2	B	13	0	0	0	0
2	C	26	0	0	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	3	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	A	222	0	0	0	0
4	B	51	0	0	0	0
4	C	153	0	0	1	0
5	B	7	0	0	0	0
6	C	34	0	0	0	0
7	A	13	0	0	0	0
7	B	23	0	0	0	0
7	C	22	0	0	0	0
All	All	12208	0	0	21	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 2.

All (21) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:430:ILE:O	1:A:431:TRP:CB	2.56	0.52
1:B:380:PHE:C	1:B:380:PHE:CD1	2.86	0.48
1:C:132:GLU:CG	1:C:133:ALA:N	2.77	0.48
1:B:98:ASN:ND2	1:B:98:ASN:N	2.62	0.48
1:B:80:PHE:CD1	1:B:80:PHE:N	2.81	0.48
1:A:430:ILE:CG2	1:A:431:TRP:N	2.78	0.47
1:B:180:VAL:O	1:B:184:THR:OG1	2.32	0.47
4:C:1567:PGT:C31	4:C:1567:PGT:C35	2.94	0.46
1:A:222:VAL:N	1:A:223:PRO:CD	2.79	0.45
1:B:80:PHE:O	1:B:81:LYS:C	2.55	0.45
1:C:213:ARG:NH2	1:C:543:ASP:OD2	2.50	0.44
1:C:118:ALA:O	1:C:395:ARG:NH1	2.51	0.43
1:C:222:VAL:N	1:C:223:PRO:CD	2.81	0.43
1:A:371:TRP:CA	1:A:371:TRP:CE3	3.01	0.43
1:C:371:TRP:N	1:C:371:TRP:CE3	2.86	0.43
1:A:310:MET:O	1:A:314:ALA:N	2.52	0.43
1:C:132:GLU:OE2	1:C:390:ARG:NE	2.53	0.42
1:C:68:VAL:O	1:C:72:ALA:N	2.53	0.42
1:B:154:LEU:O	1:B:412:TRP:NE1	2.52	0.42
1:A:460:LEU:O	1:A:464:PHE:N	2.54	0.41
1:A:371:TRP:CE3	1:A:371:TRP:N	2.89	0.41

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution. The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	529/566 (94%)	496 (94%)	28 (5%)	5 (1%)	25	55
1	B	495/566 (88%)	458 (92%)	35 (7%)	2 (0%)	43	76
1	C	504/566 (89%)	476 (94%)	26 (5%)	2 (0%)	43	76
All	All	1528/1698 (90%)	1430 (94%)	89 (6%)	9 (1%)	33	66

All (9) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	583	ARG
1	A	431	TRP
1	B	81	LYS
1	A	585	GLU
1	C	131	ASP
1	A	485	LEU
1	B	272	GLU
1	A	448	PRO
1	C	125	ILE

### 5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution. The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	415/439 (94%)	388 (94%)	27 (6%)	24	51
1	B	385/439 (88%)	370 (96%)	15 (4%)	43	76
1	C	393/439 (90%)	366 (93%)	27 (7%)	22	48
All	All	1193/1317 (91%)	1124 (94%)	69 (6%)	28	57

All (69) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	71	LEU
1	A	81	LYS
1	A	137	ARG
1	A	139	VAL
1	A	163	LEU
1	A	228	LYS
1	A	247	VAL
1	A	270	ILE
1	A	316	LEU
1	A	328	SER
1	A	358	THR
1	A	363	LEU
1	A	371	TRP
1	A	377	TRP
1	A	392	ARG
1	A	399	LEU
1	A	402	LEU
1	A	439	GLN
1	A	458	MET
1	A	479	MET
1	A	485	LEU
1	A	504	LEU
1	A	543	ASP
1	A	557	GLN
1	A	563	LEU
1	A	569	VAL
1	A	584	ARG
1	B	67	LEU
1	B	98	ASN
1	B	108	THR
1	B	131	ASP
1	B	167	ARG
1	B	228	LYS
1	B	271	ILE
1	B	311	VAL
1	B	327	VAL
1	B	346	GLN
1	B	371	TRP
1	B	377	TRP
1	B	392	ARG
1	B	430	ILE
1	B	541	VAL

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	C	67	LEU
1	C	78	ILE
1	C	81	LYS
1	C	121	LYS
1	C	127	LEU
1	C	138	THR
1	C	146	PHE
1	C	216	LEU
1	C	225	ILE
1	C	237	LEU
1	C	259	LEU
1	C	271	ILE
1	C	277	TRP
1	C	300	LYS
1	C	361	GLU
1	C	363	LEU
1	C	371	TRP
1	C	377	TRP
1	C	381	VAL
1	C	392	ARG
1	C	430	ILE
1	C	437	GLU
1	C	451	GLN
1	C	479	MET
1	C	543	ASP
1	C	548	VAL
1	C	554	ARG

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 chains 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 20 ligands modelled in this entry, 5 are monoatomic - leaving 15 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	FLC	A	1001	-	5,12,12	2.17	2 (40%)	7,17,17	1.12	1 (14%)
4	PGT	A	1588	-	50,50,50	0.88	2 (4%)	56,56,56	1.02	2 (3%)
4	PGT	A	1589	-	50,50,50	0.91	2 (4%)	56,56,56	0.93	2 (3%)
4	PGT	A	1590	-	50,50,50	0.89	2 (4%)	56,56,56	0.99	3 (5%)
4	PGT	A	1591	-	16,17,50	0.29	0	15,16,56	0.63	0
4	PGT	A	1592	-	50,50,50	0.87	2 (4%)	56,56,56	1.01	2 (3%)
2	FLC	B	1001	-	5,12,12	2.22	2 (40%)	7,17,17	1.41	2 (28%)
4	PGT	B	1554	-	50,50,50	0.88	2 (4%)	56,56,56	0.94	3 (5%)
5	PEG	B	1555	-	6,6,6	0.44	0	5,5,5	0.25	0
2	FLC	C	1001	-	5,12,12	2.06	2 (40%)	7,17,17	1.80	2 (28%)
4	PGT	C	1565	-	50,50,50	0.87	2 (4%)	56,56,56	1.05	4 (7%)
4	PGT	C	1566	-	50,50,50	0.88	2 (4%)	56,56,56	1.01	3 (5%)
4	PGT	C	1567	-	50,50,50	0.89	2 (4%)	56,56,56	0.95	3 (5%)
6	CM5	C	1568	-	36,36,36	0.39	0	49,49,49	0.80	1 (2%)
2	FLC	C	1569	-	5,12,12	2.10	2 (40%)	7,17,17	1.61	2 (28%)

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	FLC	A	1001	-	-	0/6/16/16	0/0/0/0
4	PGT	A	1588	-	-	1/55/55/55	0/0/0/0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	PGT	A	1589	-	-	0/55/55/55	0/0/0/0
4	PGT	A	1590	-	-	0/55/55/55	0/0/0/0
4	PGT	A	1591	-	-	0/15/15/55	0/0/0/0
4	PGT	A	1592	-	-	0/55/55/55	0/0/0/0
2	FLC	B	1001	-	-	0/6/16/16	0/0/0/0
4	PGT	B	1554	-	-	0/55/55/55	0/0/0/0
5	PEG	B	1555	-	-	0/4/4/4	0/0/0/0
2	FLC	C	1001	-	-	0/6/16/16	0/0/0/0
4	PGT	C	1565	-	-	2/55/55/55	0/0/0/0
4	PGT	C	1566	-	-	1/55/55/55	0/0/0/0
4	PGT	C	1567	-	-	0/55/55/55	0/0/0/0
6	CM5	C	1568	-	-	0/17/65/65	0/3/3/3
2	FLC	C	1569	-	-	0/6/16/16	0/0/0/0

All (24) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	1589	PGT	O3-C11	4.23	1.46	1.33
4	A	1590	PGT	O3-C11	4.16	1.46	1.33
4	A	1592	PGT	O3-C11	4.05	1.46	1.33
4	C	1567	PGT	O3-C11	4.03	1.45	1.33
4	A	1589	PGT	O2-C31	4.02	1.46	1.34
4	C	1565	PGT	O3-C11	3.99	1.45	1.33
4	A	1588	PGT	O2-C31	3.97	1.46	1.34
4	B	1554	PGT	O3-C11	3.97	1.45	1.33
4	C	1566	PGT	O3-C11	3.97	1.45	1.33
4	A	1588	PGT	O3-C11	3.95	1.45	1.33
4	B	1554	PGT	O2-C31	3.93	1.46	1.34
4	C	1567	PGT	O2-C31	3.89	1.46	1.34
4	C	1566	PGT	O2-C31	3.83	1.45	1.34
4	A	1592	PGT	O2-C31	3.78	1.45	1.34
4	C	1565	PGT	O2-C31	3.76	1.45	1.34
4	A	1590	PGT	O2-C31	3.73	1.45	1.34
2	B	1001	FLC	CG-CGC	3.43	1.51	1.49
2	A	1001	FLC	CG-CGC	3.26	1.51	1.49
2	B	1001	FLC	CA-CAC	3.24	1.51	1.49
2	A	1001	FLC	CA-CAC	3.23	1.51	1.49
2	C	1001	FLC	CG-CGC	3.21	1.51	1.49
2	C	1569	FLC	CG-CGC	3.17	1.51	1.49
2	C	1569	FLC	CA-CAC	3.04	1.51	1.49
2	C	1001	FLC	CA-CAC	2.83	1.51	1.49

All (30) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	1588	PGT	O2-C31-C32	4.52	121.46	111.56
4	C	1566	PGT	O2-C31-C32	4.19	120.74	111.56
4	B	1554	PGT	O2-C31-C32	3.99	120.31	111.56
4	A	1592	PGT	O2-C31-C32	3.84	119.98	111.56
4	C	1567	PGT	O2-C31-C32	3.77	119.82	111.56
4	A	1589	PGT	O2-C31-C32	3.69	119.65	111.56
4	A	1590	PGT	O2-C31-C32	3.54	119.32	111.56
4	A	1590	PGT	C2-O2-C31	-3.51	109.26	117.92
4	C	1565	PGT	C2-O2-C31	-3.47	109.35	117.92
2	C	1001	FLC	CB-CA-CAC	-3.44	109.87	115.01
4	C	1565	PGT	O2-C31-C32	3.42	119.06	111.56
2	C	1569	FLC	CB-CA-CAC	-3.02	110.50	115.01
2	C	1569	FLC	CB-CG-CGC	-2.94	110.62	115.01
4	A	1589	PGT	O3-C11-C12	2.80	120.75	111.94
2	C	1001	FLC	CB-CG-CGC	-2.78	110.86	115.01
6	C	1568	CM5	C24-O23-C16	-2.77	110.92	117.99
4	A	1590	PGT	O3-C11-C12	2.75	120.58	111.94
4	A	1592	PGT	O3-C11-C12	2.65	120.29	111.94
4	A	1588	PGT	O3-C11-C12	2.62	120.20	111.94
4	C	1566	PGT	C2-O2-C31	-2.60	111.50	117.92
4	C	1567	PGT	O3-C11-C12	2.56	120.00	111.94
4	B	1554	PGT	O3-C11-C12	2.41	119.52	111.94
2	B	1001	FLC	CB-CA-CAC	-2.34	111.51	115.01
4	C	1565	PGT	O4P-P-O3P	-2.31	98.09	104.68
2	B	1001	FLC	CB-CG-CGC	-2.29	111.58	115.01
4	C	1566	PGT	O3-C11-C12	2.28	119.11	111.94
2	A	1001	FLC	CB-CA-CAC	-2.26	111.64	115.01
4	C	1565	PGT	O3-C11-C12	2.14	118.68	111.94
4	C	1567	PGT	C2-O2-C31	-2.11	112.71	117.92
4	B	1554	PGT	C2-O2-C31	-2.03	112.92	117.92

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	C	1565	PGT	C2-O2-C31-C32
4	C	1566	PGT	C2-O2-C31-C32
4	C	1565	PGT	C2-O2-C31-O31
4	A	1588	PGT	C2-O2-C31-C32

There are no ring outliers.

## 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 ⓘ

### 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	531/566 (93%)	0.13	34 (6%) 19 21	58, 102, 170, 244	1 (0%)
1	B	497/566 (87%)	0.38	59 (11%) 5 5	48, 104, 176, 248	0
1	C	506/566 (89%)	0.17	31 (6%) 21 22	48, 86, 159, 215	0
All	All	1534/1698 (90%)	0.23	124 (8%) 12 12	48, 98, 170, 248	1 (0%)

All (124) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	277	TRP	10.8
1	B	227	GLU	7.7
1	B	81	LYS	7.4
1	B	75	VAL	7.4
1	B	550	TYR	7.3
1	B	228	LYS	7.1
1	B	80	PHE	6.4
1	B	84	PHE	6.1
1	B	76	TRP	5.8
1	B	78	ILE	5.8
1	B	71	LEU	5.8
1	B	87	PHE	5.6
1	A	233	TRP	5.5
1	C	560	ASN	5.5
1	B	82	ASP	5.4
1	C	559	PHE	5.2
1	B	490	TRP	5.2
1	B	485	LEU	5.1
1	A	274	PRO	5.1
1	B	214	LYS	5.1
1	B	74	VAL	5.0
1	C	274	PRO	5.0
1	A	446	ALA	4.9

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	279	ILE	4.8
1	B	83	SER	4.8
1	C	561	ALA	4.8
1	B	270	ILE	4.7
1	B	486	GLU	4.6
1	A	298	VAL	4.5
1	B	79	GLY	4.5
1	B	137	ARG	4.5
1	B	552	GLU	4.4
1	C	271	ILE	4.4
1	A	443	LEU	4.4
1	A	448	PRO	4.4
1	A	447	LEU	4.4
1	C	276	ASP	4.3
1	A	273	ASP	4.3
1	B	231	GLU	4.2
1	C	280	VAL	4.1
1	B	233	TRP	4.0
1	A	139	VAL	4.0
1	B	455	ILE	3.9
1	C	553	TYR	3.9
1	C	56	ALA	3.8
1	A	62	VAL	3.8
1	B	224	LEU	3.8
1	C	269	ASN	3.7
1	B	139	VAL	3.7
1	B	73	THR	3.7
1	C	550	TYR	3.7
1	C	508	LEU	3.7
1	B	58	LEU	3.6
1	B	549	ILE	3.6
1	A	297	GLY	3.5
1	B	225	ILE	3.5
1	C	272	GLU	3.4
1	B	140	SER	3.4
1	A	299	GLY	3.3
1	B	77	GLY	3.3
1	C	270	ILE	3.3
1	B	212	GLY	3.2
1	B	234	LEU	3.1
1	B	152	ILE	3.1
1	B	277	TRP	3.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	107	GLY	3.1
1	A	138	THR	3.0
1	B	215	GLN	3.0
1	A	512	ASP	3.0
1	B	482	HIS	3.0
1	B	551	LEU	2.9
1	C	288	LEU	2.9
1	A	131	ASP	2.9
1	A	302	ILE	2.9
1	A	228	LYS	2.8
1	B	452	ILE	2.8
1	B	484	GLN	2.8
1	A	445	HIS	2.8
1	C	556	GLN	2.7
1	B	70	VAL	2.7
1	A	292	PHE	2.7
1	B	86	ASN	2.6
1	A	553	TYR	2.6
1	C	507	LEU	2.6
1	B	548	VAL	2.6
1	C	273	ASP	2.6
1	B	56	ALA	2.6
1	B	377	TRP	2.6
1	B	138	THR	2.6
1	C	490	TRP	2.6
1	C	194	TRP	2.5
1	C	405	PRO	2.5
1	C	555	GLU	2.5
1	A	82	ASP	2.5
1	B	111	VAL	2.5
1	C	549	ILE	2.4
1	B	230	ALA	2.4
1	B	342	SER	2.4
1	B	60	TRP	2.4
1	C	416	PHE	2.4
1	C	190	THR	2.4
1	A	294	ALA	2.3
1	A	232	GLY	2.3
1	A	486	GLU	2.3
1	C	551	LEU	2.3
1	C	193	PRO	2.3
1	B	57	SER	2.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	572	GLU	2.3
1	B	487	ALA	2.2
1	A	225	ILE	2.2
1	B	107	GLY	2.2
1	B	454	GLY	2.2
1	B	226	GLY	2.2
1	A	323	VAL	2.2
1	A	270	ILE	2.2
1	A	549	ILE	2.1
1	A	146	PHE	2.1
1	B	269	ASN	2.1
1	B	190	THR	2.1
1	A	130	ILE	2.1
1	C	433	ASP	2.1
1	A	80	PHE	2.1
1	A	296	SER	2.1
1	A	295	ILE	2.0

## 6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

## 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 6.4 Ligands ⓘ

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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
3	CL	A	1587	1/1	0.31	8.38	105,105,105,105	0
4	PGT	A	1591	18/51	0.60	7.46	86,92,114,114	0
5	PEG	B	1555	7/7	0.48	6.38	146,152,153,155	0
4	PGT	C	1567	51/51	0.47	5.20	107,157,233,319	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
4	PGT	C	1565	51/51	0.33	3.02	106,176,230,291	0
4	PGT	C	1566	51/51	0.23	2.91	109,144,184,223	0
2	FLC	A	1001	13/13	0.30	2.50	147,163,169,170	0
2	FLC	C	1001	13/13	0.31	2.25	133,140,145,145	0
2	FLC	C	1569	13/13	0.29	2.16	188,193,208,210	0
4	PGT	A	1588	51/51	0.31	2.03	85,136,175,236	0
4	PGT	B	1554	51/51	0.47	1.92	72,136,208,240	0
2	FLC	B	1001	13/13	0.33	1.64	140,150,161,162	0
4	PGT	A	1589	51/51	0.34	1.56	74,132,232,385	0
3	CL	C	1564	1/1	0.20	1.51	85,85,85,85	0
4	PGT	A	1590	51/51	0.24	1.03	72,99,130,231	0
4	PGT	A	1592	51/51	0.29	0.91	69,150,203,314	0
3	CL	B	1553	1/1	0.20	0.25	109,109,109,109	0
6	CM5	C	1568	34/34	0.19	0.17	112,182,209,211	0
3	CL	C	1562	1/1	0.08	-1.09	74,74,74,74	0
3	CL	C	1563	1/1	0.10	-1.51	64,64,64,64	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.