



wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 18, 2017 – 09:17 AM EDT

PDB ID : 4P02
Title : Structure of Bacterial Cellulose Synthase with cyclic-di-GMP bound.
Authors : Morgan, J.L.W.; McNamara, J.T.; Zimmer, J.
Deposited on : unknown
Resolution : 2.65 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at 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 : 1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20029824
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
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) : rb-20029824

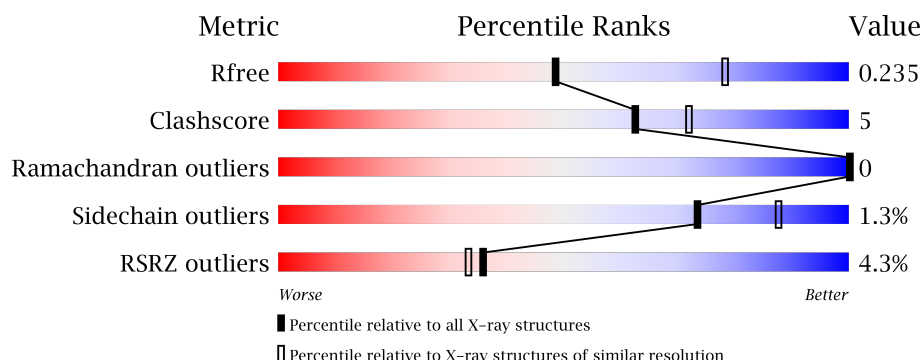
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.65 Å.

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



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	100719	3491 (2.70-2.62)
Clashscore	112137	1026 (2.68-2.64)
Ramachandran outliers	110173	1010 (2.68-2.64)
Sidechain outliers	110143	1010 (2.68-2.64)
RSRZ outliers	101464	3511 (2.70-2.62)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for ≥ 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	803	<div> <div>5%</div> <div> <div></div> <div>79%</div> <div>11%</div> <div>9%</div> </div> </div>
2	B	724	<div> <div>3%</div> <div> <div></div> <div>79%</div> <div>12%</div> <div>10%</div> </div> </div>
3	D	9	<div> <div></div> <div>100%</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	BGC	A	904	-	-	X	-
4	BGC	A	905	-	-	X	-
4	BGC	A	916	-	-	-	X
6	PLC	A	920	-	-	-	X
6	PLC	B	802	-	-	-	X
6	PLC	B	804	-	-	-	X

2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 11040 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 Cellulose Synthase subunit A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	728	Total	C	N	O	S	0	1	0
			5739	3725	1000	982	32			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	expression tag	UNP Q3J125
A	1	GLY	-	expression tag	UNP Q3J125
A	789	HIS	-	expression tag	UNP Q3J125
A	790	HIS	-	expression tag	UNP Q3J125
A	791	HIS	-	expression tag	UNP Q3J125
A	792	HIS	-	expression tag	UNP Q3J125
A	793	HIS	-	expression tag	UNP Q3J125
A	794	HIS	-	expression tag	UNP Q3J125
A	795	LYS	-	expression tag	UNP Q3J125
A	796	LEU	-	expression tag	UNP Q3J125
A	797	HIS	-	expression tag	UNP Q3J125
A	798	HIS	-	expression tag	UNP Q3J125
A	799	HIS	-	expression tag	UNP Q3J125
A	800	HIS	-	expression tag	UNP Q3J125
A	801	HIS	-	expression tag	UNP Q3J125
A	802	HIS	-	expression tag	UNP Q3J125

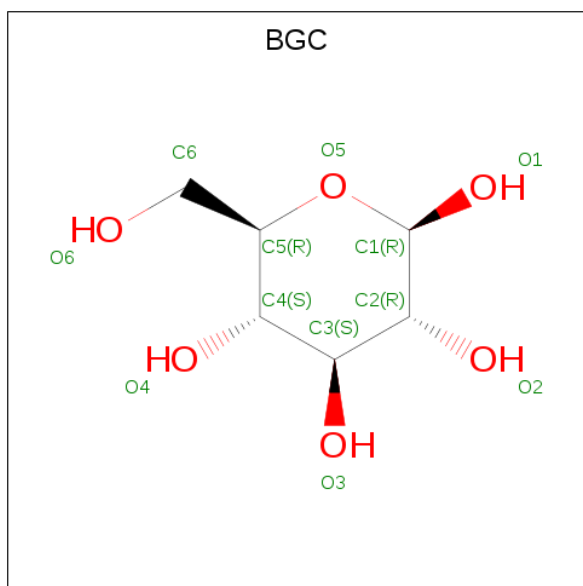
- Molecule 2 is a protein called Cellulose Synthase subunit B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	655	Total	C	N	O	S	0	0	0
			4887	3100	864	907	16			

- Molecule 3 is a protein called unidentified peptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	D	9	Total	C	N	O	0	0	0
			45	27	9	9			

- Molecule 4 is BETA-D-GLUCOSE (three-letter code: BGC) (formula: $C_6H_{12}O_6$).



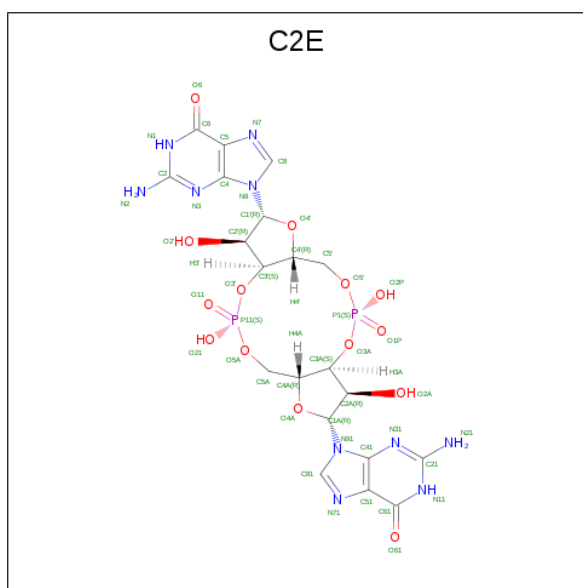
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			11	6	5		
4	A	1	Total	C	O	0	0
			11	6	5		
4	A	1	Total	C	O	0	0
			11	6	5		
4	A	1	Total	C	O	0	0
			11	6	5		
4	A	1	Total	C	O	0	0
			11	6	5		
4	A	1	Total	C	O	0	0
			11	6	5		
4	A	1	Total	C	O	0	0
			11	6	5		
4	A	1	Total	C	O	0	0
			11	6	5		

Continued on next page...

Continued from previous page...

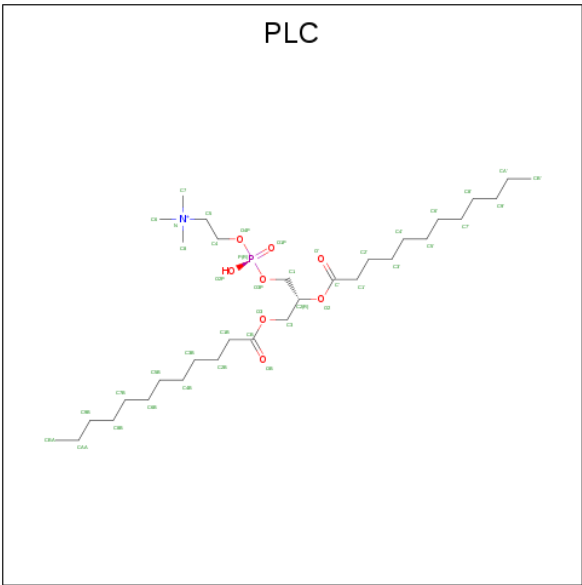
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			11	6	5		
4	A	1	Total	C	O	0	0
			11	6	5		
4	A	1	Total	C	O	0	0
			11	6	5		
4	A	1	Total	C	O	0	0
			11	6	5		
4	A	1	Total	C	O	0	0
			11	6	5		

- Molecule 5 is 9,9'-[(2R,3R,3aS,5S,7aR,9R,10R,10aS,12S,14aR)-3,5,10,12-tetrahydroxy-5,12-dioxidoctahydro-2H,7H-difuro[3,2-d:3',2'-j][1,3,7,9,2,8]tetraoxadiphosphacyclododecine-2,9-diyl]bis(2-amino-1,9-dihydro-6H-purin-6-one) (three-letter code: C2E) (formula: $C_{20}H_{24}N_{10}O_{14}P_2$).



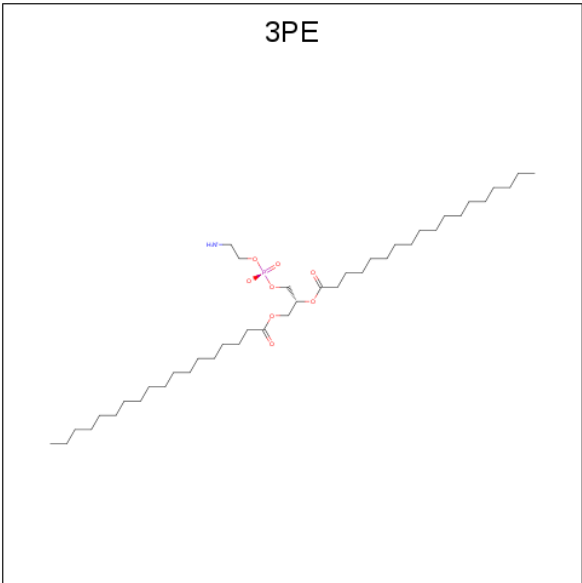
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	A	1	Total	C	N	O	P	0	0
			46	20	10	14	2		
5	A	1	Total	C	N	O	P	0	0
			46	20	10	14	2		

- Molecule 6 is DIUNDECYL PHOSPHATIDYL CHOLINE (three-letter code: PLC) (formula: $C_{32}H_{65}NO_8P$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
6	A	1	Total	C	N	O	P	0	0
			38	28	1	8	1		
6	B	1	Total	C				0	0
			11	11					
6	B	1	Total	C				0	0
			9	9					
6	B	1	Total	C				0	0
			11	11					

- Molecule 7 is 1,2-DIACYL-SN-GLYCERO-3-PHOSPHOETHANOLAMINE (three-letter code: 3PE) (formula: C₄₁H₈₂NO₈P).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
7	A	1	Total	C	N	O	P	0	0
			20	10	1	8	1		

- Molecule 8 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	B	1	Total	Mg	0	0
			1	1		

T719	ARG
R720	LYS
	GLY
	LEU

- Molecule 3: unidentified peptide

Chain D:  100%

There are no outlier residues recorded for this chain.

GLOBAL-STATISTICS INFOmissingINFO

4 Model quality

4.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: MG, BGC, PLC, 3PE, C2E

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	0.22	0/5888	0.39	0/8007
2	B	0.21	0/5006	0.41	0/6865
All	All	0.21	0/10894	0.40	0/14872

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

4.2 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	5739	0	5855	52	0
2	B	4887	0	4966	55	0
3	D	45	0	12	0	0
4	A	187	0	154	17	0
5	A	92	0	44	0	0
6	A	38	0	53	0	0
6	B	31	0	56	0	0
7	A	20	0	14	0	0
8	B	1	0	0	0	0
All	All	11040	0	11154	111	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including

hydrogen atoms). The all-atom clashscore for this structure is 5.

The worst 5 of 111 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)
4:A:904:BGC:O3	4:A:905:BGC:O5	1.85	0.94
4:A:905:BGC:O3	4:A:906:BGC:O5	1.90	0.88
4:A:904:BGC:HC	4:A:905:BGC:C5	1.86	0.88
4:A:905:BGC:O6	2:B:390:ARG:NH2	2.07	0.84
4:A:901:BGC:H6C1	4:A:902:BGC:O2	1.78	0.84

There are no symmetry-related clashes.

4.3 Torsion angles [i](#)

4.3.1 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	727/803 (90%)	706 (97%)	21 (3%)	0	100	100
2	B	651/724 (90%)	635 (98%)	16 (2%)	0	100	100
All	All	1378/1527 (90%)	1341 (97%)	37 (3%)	0	100	100

There are no Ramachandran outliers to report.

4.3.2 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	599/661 (91%)	588 (98%)	11 (2%)	64	82
2	B	520/572 (91%)	516 (99%)	4 (1%)	85	93
All	All	1119/1233 (91%)	1104 (99%)	15 (1%)	73	88

5 of 15 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	568	PHE
1	A	573	VAL
2	B	271	ARG
1	A	512	LEU
2	B	158	ARG

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

Mol	Chain	Res	Type
2	B	77	GLN
2	B	129	GLN

4.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

4.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

4.6 Ligand geometry [i](#)

Of 25 ligands modelled in this entry, 1 is monoatomic - leaving 24 for Mogul analysis.

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

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# $ Z > 2$	Counts	RMSZ	# $ Z > 2$
4	BGC	A	901	4	11,11,12	0.30	0	13,15,17	0.69	0
4	BGC	A	902	4	11,11,12	0.35	0	13,15,17	1.00	1 (7%)
4	BGC	A	903	4	11,11,12	0.30	0	13,15,17	0.67	0
4	BGC	A	904	4	11,11,12	0.38	0	13,15,17	1.16	1 (7%)
4	BGC	A	905	4	11,11,12	0.53	0	13,15,17	1.79	3 (23%)
4	BGC	A	906	4	11,11,12	0.26	0	13,15,17	1.08	2 (15%)
4	BGC	A	907	4	11,11,12	0.21	0	13,15,17	0.73	0
4	BGC	A	908	4	11,11,12	0.43	0	13,15,17	1.59	3 (23%)
4	BGC	A	909	4	11,11,12	0.30	0	13,15,17	0.77	1 (7%)
4	BGC	A	910	4	11,11,12	0.27	0	13,15,17	0.90	1 (7%)
4	BGC	A	911	4	11,11,12	0.30	0	13,15,17	0.71	0
4	BGC	A	912	4	11,11,12	0.29	0	13,15,17	0.78	1 (7%)
4	BGC	A	913	4	11,11,12	0.36	0	13,15,17	0.78	0
4	BGC	A	914	4	11,11,12	0.38	0	13,15,17	1.12	1 (7%)
4	BGC	A	915	4	11,11,12	0.28	0	13,15,17	0.92	1 (7%)
4	BGC	A	916	4	11,11,12	0.35	0	13,15,17	1.17	1 (7%)
4	BGC	A	917	4	11,11,12	0.29	0	13,15,17	0.66	0
5	C2E	A	918	-	44,52,52	1.20	4 (9%)	50,82,82	1.99	13 (26%)
5	C2E	A	919	-	44,52,52	1.22	4 (9%)	50,82,82	1.94	13 (26%)
6	PLC	A	920	-	37,37,41	1.11	2 (5%)	42,45,49	1.14	3 (7%)
7	3PE	A	921	-	19,19,50	1.59	2 (10%)	21,24,55	1.66	3 (14%)
6	PLC	B	802	-	10,10,41	0.26	0	9,9,49	0.53	0
6	PLC	B	803	-	8,8,41	0.25	0	7,7,49	0.49	0
6	PLC	B	804	-	10,10,41	0.25	0	9,9,49	0.53	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
4	BGC	A	901	4	-	0/2/19/22	0/1/1/1
4	BGC	A	902	4	-	0/2/19/22	0/1/1/1
4	BGC	A	903	4	-	0/2/19/22	0/1/1/1
4	BGC	A	904	4	-	0/2/19/22	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	BGC	A	905	4	-	0/2/19/22	0/1/1/1
4	BGC	A	906	4	-	0/2/19/22	0/1/1/1
4	BGC	A	907	4	-	0/2/19/22	0/1/1/1
4	BGC	A	908	4	-	0/2/19/22	0/1/1/1
4	BGC	A	909	4	-	0/2/19/22	0/1/1/1
4	BGC	A	910	4	-	0/2/19/22	0/1/1/1
4	BGC	A	911	4	-	0/2/19/22	0/1/1/1
4	BGC	A	912	4	-	0/2/19/22	0/1/1/1
4	BGC	A	913	4	-	0/2/19/22	0/1/1/1
4	BGC	A	914	4	-	0/2/19/22	0/1/1/1
4	BGC	A	915	4	-	0/2/19/22	0/1/1/1
4	BGC	A	916	4	-	0/2/19/22	0/1/1/1
4	BGC	A	917	4	-	0/2/19/22	0/1/1/1
5	C2E	A	918	-	-	0/22/62/62	0/6/7/7
5	C2E	A	919	-	-	0/22/62/62	0/6/7/7
6	PLC	A	920	-	-	0/41/41/45	0/0/0/0
7	3PE	A	921	-	-	0/22/22/54	0/0/0/0
6	PLC	B	802	-	-	0/8/8/45	0/0/0/0
6	PLC	B	803	-	-	0/6/6/45	0/0/0/0
6	PLC	B	804	-	-	0/8/8/45	0/0/0/0

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	918	C2E	C51-C41	2.91	1.47	1.40
5	A	919	C2E	C51-C41	3.03	1.47	1.40
5	A	918	C2E	C5-C4	3.04	1.47	1.40
5	A	919	C2E	C5-C4	3.09	1.47	1.40
5	A	919	C2E	C61-C51	3.76	1.48	1.41

The worst 5 of 48 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	919	C2E	C5-C6-N1	-3.84	118.02	123.48
5	A	918	C2E	C61-C51-C41	-3.75	117.11	120.84
5	A	919	C2E	C51-C61-N11	-3.75	118.15	123.48
5	A	918	C2E	C51-C61-N11	-3.73	118.17	123.48
5	A	918	C2E	C5-C6-N1	-3.73	118.17	123.48

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

8 monomers are involved in 17 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	901	BGC	1	0
4	A	902	BGC	2	0
4	A	903	BGC	1	0
4	A	904	BGC	6	0
4	A	905	BGC	10	0
4	A	906	BGC	3	0
4	A	907	BGC	3	0
4	A	914	BGC	1	0

4.7 Other polymers [i](#)

There are no such residues in this entry.

4.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

5 Fit of model and data [i](#)

5.1 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, 95th 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 ²)	Q<0.9
1	A	728/803 (90%)	0.16	39 (5%) 26 24	45, 69, 112, 162	0
2	B	655/724 (90%)	-0.06	21 (3%) 48 46	43, 61, 103, 148	0
3	D	0/9	-	-	-	-
All	All	1383/1536 (90%)	0.06	60 (4%) 36 33	43, 66, 109, 162	0

The worst 5 of 60 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	508	LYS	8.3
1	A	13	VAL	5.3
2	B	593	LEU	5.3
1	A	506	THR	5.1
1	A	509	ASP	5.0

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

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

5.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.4 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, 95th 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(\AA^2)	Q<0.9
6	PLC	B	804	11/42	0.91	0.37	8.72	65,74,86,87	0
6	PLC	A	920	38/42	0.88	0.38	5.49	77,105,134,161	0
6	PLC	B	802	11/42	0.89	0.25	2.51	56,68,80,82	0
4	BGC	A	916	11/12	0.95	0.33	2.38	48,62,87,117	0
7	3PE	A	921	20/51	0.93	0.20	1.45	65,78,125,128	0
4	BGC	A	917	11/12	0.96	0.32	0.82	50,60,70,85	0
6	PLC	B	803	9/42	0.84	0.20	0.82	50,78,86,89	0
4	BGC	A	910	11/12	0.96	0.18	-0.21	53,63,80,85	0
4	BGC	A	911	11/12	0.98	0.16	-0.72	42,57,75,83	0
4	BGC	A	915	11/12	0.97	0.20	-0.83	43,45,70,71	0
4	BGC	A	906	11/12	0.95	0.14	-1.01	20,38,84,108	0
4	BGC	A	912	11/12	0.96	0.17	-1.03	43,52,63,73	0
4	BGC	A	908	11/12	0.95	0.15	-1.06	37,54,91,98	0
4	BGC	A	914	11/12	0.94	0.17	-1.08	49,58,75,77	0
4	BGC	A	905	11/12	0.94	0.14	-1.12	34,48,64,66	0
4	BGC	A	913	11/12	0.96	0.17	-1.39	49,55,64,65	0
4	BGC	A	909	11/12	0.96	0.13	-1.48	45,59,76,92	0
4	BGC	A	907	11/12	0.95	0.10	-1.73	32,40,62,67	0
5	C2E	A	919	46/46	0.97	0.13	-2.15	36,51,63,75	0
5	C2E	A	918	46/46	0.97	0.11	-2.23	37,46,58,65	0
8	MG	B	801	1/1	0.91	0.12	-3.61	38,38,38,38	0
4	BGC	A	901	11/12	0.73	0.54	-	106,136,161,167	0
4	BGC	A	904	11/12	0.93	0.18	-	45,56,132,152	0
4	BGC	A	902	11/12	0.84	0.37	-	107,118,143,154	0
4	BGC	A	903	11/12	0.91	0.20	-	66,93,110,122	0

5.5 Other polymers [i](#)

There are no such residues in this entry.