



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 9, 2020 – 07:45 PM BST

PDB ID : 1J11  
Title : beta-amylase from *Bacillus cereus* var. *mycoides* in complex with alpha-EPG  
Authors : Oyama, T.; Miyake, H.; Kusunoki, M.; Nitta, Y.  
Deposited on : 2002-11-25  
Resolution : 2.00 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](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 : 2.13.1  
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) : 2.13.1

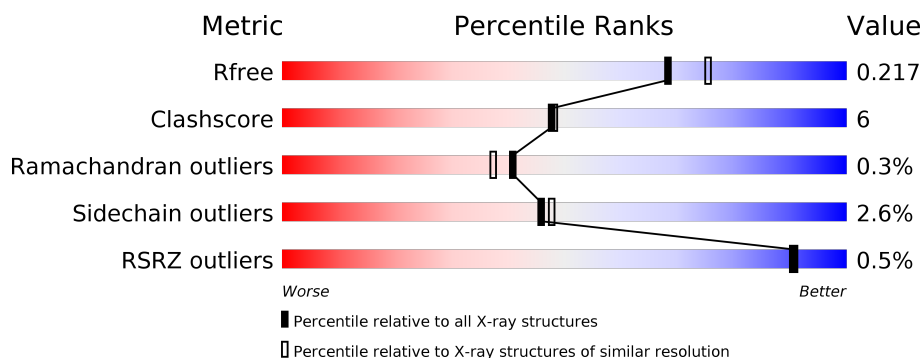
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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	516	 85% 15% .
1	B	516	 82% 17% .
1	C	516	 83% 16% .
1	D	516	 82% 16% .

## 2 Entry composition [i](#)

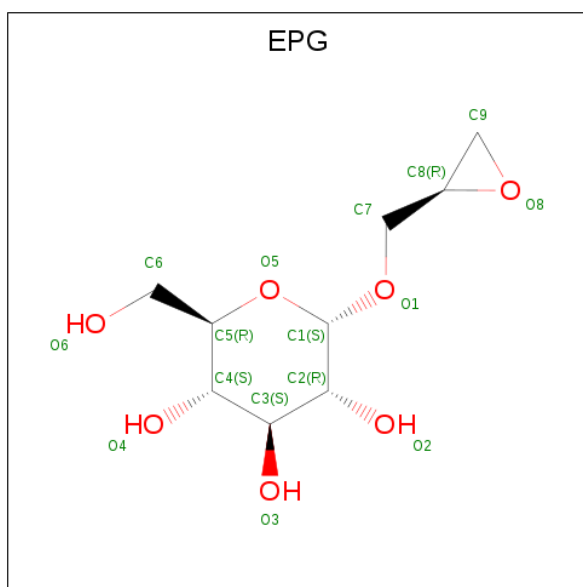
There are 4 unique types of molecules in this entry. The entry contains 17070 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Beta-amylase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	516	Total	C	N	O	S	0	0	0
			4119	2645	676	781	17			
1	B	516	Total	C	N	O	S	0	0	0
			4119	2645	676	781	17			
1	C	516	Total	C	N	O	S	0	0	0
			4119	2645	676	781	17			
1	D	516	Total	C	N	O	S	0	0	0
			4119	2645	676	781	17			

- Molecule 2 is (2R)-oxiran-2-ylmethyl alpha-D-glucopyranoside (three-letter code: EPG) (formula: C<sub>9</sub>H<sub>16</sub>O<sub>7</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			16	9	7		
2	B	1	Total	C	O	0	0
			16	9	7		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	C	1	Total	C	O	0	0
			16	9	7		
2	D	1	Total	C	O	0	0
			16	9	7		

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

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

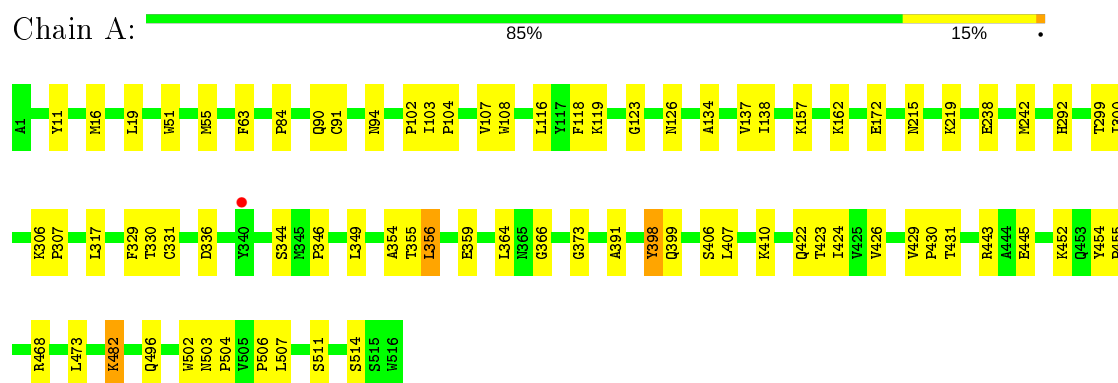
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	160	Total	O	0	0
			160	160		
4	B	103	Total	O	0	0
			103	103		
4	C	138	Total	O	0	0
			138	138		
4	D	125	Total	O	0	0
			125	125		

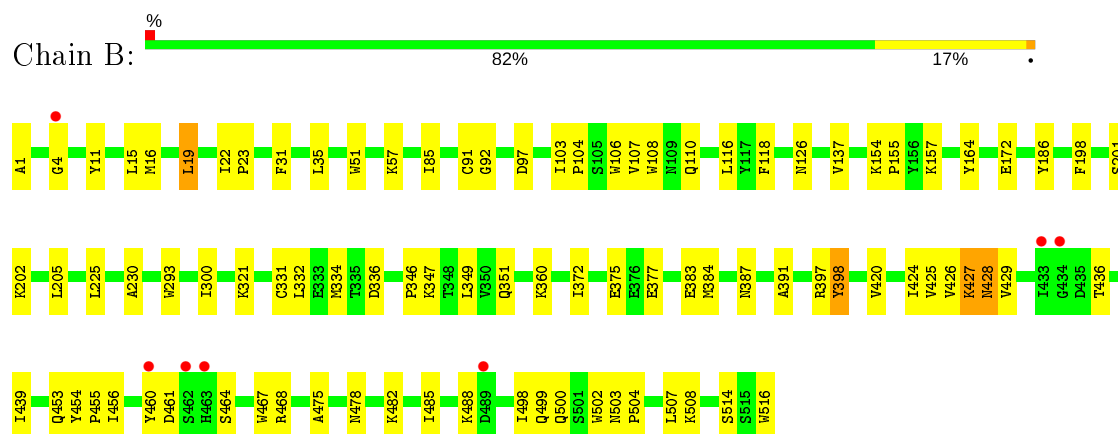
### 3 Residue-property plots

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

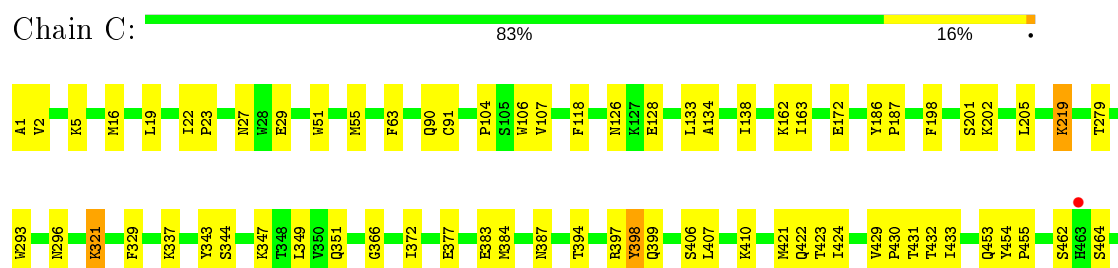
- Molecule 1: Beta-amylase

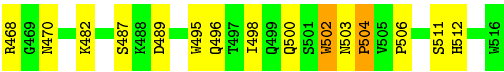


- Molecule 1: Beta-amylase

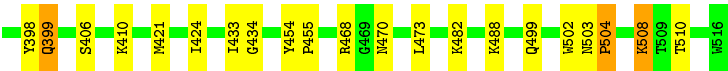
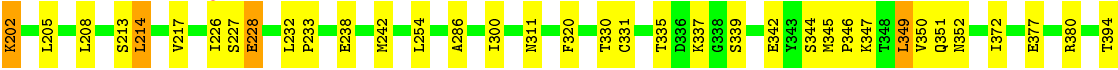
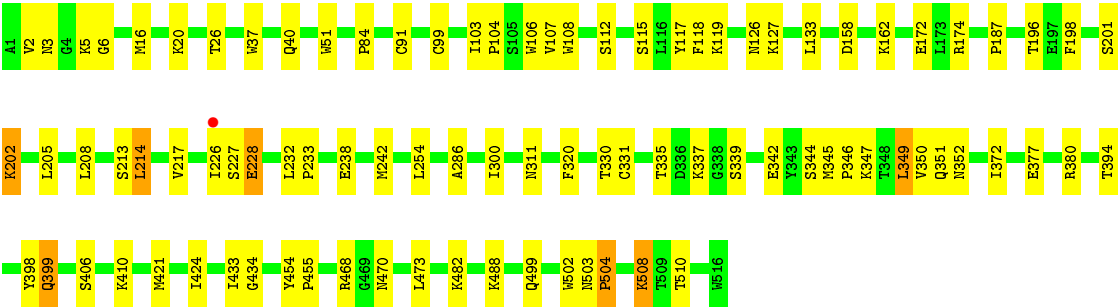
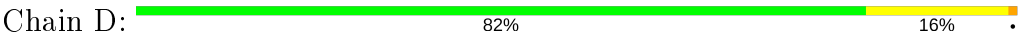


- Molecule 1: Beta-amylase





● Molecule 1: Beta-amylase



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	177.90 Å 112.90 Å 146.20 Å 90.00° 105.80° 90.00°	Depositor
Resolution (Å)	8.00 – 2.00 83.97 – 1.90	Depositor EDS
% Data completeness (in resolution range)	(Not available) (8.00-2.00) 65.8 (83.97-1.90)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.72 (at 1.90 Å)	Xtriage
Refinement program	X-PLOR 3.1	Depositor
R, $R_{free}$	0.200 , 0.246 0.180 , 0.217	Depositor DCC
$R_{free}$ test set	7107 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	13.1	Xtriage
Anisotropy	0.670	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 48.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	17070	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	17.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.75% of the height of the origin peak. No significant pseudotranslation is detected.*

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

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

## 5 Model quality [i](#)

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

Bond lengths and bond angles in the following residue types are not validated in this section: EPG, CA

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	0.43	1/4234 (0.0%)	0.61	0/5751
1	B	0.43	1/4234 (0.0%)	0.60	0/5751
1	C	0.43	1/4234 (0.0%)	0.61	0/5751
1	D	0.43	1/4234 (0.0%)	0.61	0/5751
All	All	0.43	4/16936 (0.0%)	0.61	0/23004

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	172	GLU	CD-OE2	13.35	1.40	1.25
1	B	172	GLU	CD-OE2	13.26	1.40	1.25
1	D	172	GLU	CD-OE2	12.89	1.39	1.25
1	C	172	GLU	CD-OE2	12.41	1.39	1.25

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	4119	0	3984	41	0
1	B	4119	0	3984	55	0
1	C	4119	0	3984	47	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	D	4119	0	3984	56	0
2	A	16	0	15	0	0
2	B	16	0	15	0	0
2	C	16	0	15	0	0
2	D	16	0	15	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
4	A	160	0	0	1	0
4	B	103	0	0	1	0
4	C	138	0	0	0	0
4	D	125	0	0	1	0
All	All	17070	0	15996	199	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 6.

The worst 5 of 199 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:A:482:LYS:HG2	1:A:496:GLN:HB3	1.68	0.74
1:B:347:LYS:HE3	1:B:384:MET:SD	2.28	0.74
1:D:2:VAL:O	1:D:5:LYS:HG2	1.90	0.71
1:B:375:GLU:HG3	4:B:801:HOH:O	1.90	0.71
1:C:372:ILE:HG23	1:C:377:GLU:HB2	1.75	0.68

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	514/516 (100%)	488 (95%)	25 (5%)	1 (0%)	47	44
1	B	514/516 (100%)	492 (96%)	20 (4%)	2 (0%)	34	30
1	C	514/516 (100%)	490 (95%)	23 (4%)	1 (0%)	47	44
1	D	514/516 (100%)	492 (96%)	20 (4%)	2 (0%)	34	30
All	All	2056/2064 (100%)	1962 (95%)	88 (4%)	6 (0%)	41	37

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	398	TYR
1	C	398	TYR
1	D	398	TYR
1	B	398	TYR
1	D	26	THR

### 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	440/440 (100%)	431 (98%)	9 (2%)	55	58
1	B	440/440 (100%)	431 (98%)	9 (2%)	55	58
1	C	440/440 (100%)	428 (97%)	12 (3%)	44	46
1	D	440/440 (100%)	425 (97%)	15 (3%)	37	36
All	All	1760/1760 (100%)	1715 (97%)	45 (3%)	46	48

5 of 45 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	187	PRO
1	C	399	GLN
1	D	502	TRP
1	C	279	THR
1	C	462	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 25 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	478	ASN
1	C	40	GLN
1	D	463	HIS
1	B	500	GLN
1	C	311	ASN

### 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 monosaccharides in this entry.

## 5.6 Ligand geometry ⓘ

Of 8 ligands modelled in this entry, 4 are monoatomic - leaving 4 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	EPG	A	601	1	17,17,17	5.91	2 (11%)	24,24,24	5.03	5 (20%)
2	EPG	B	601	1	17,17,17	5.98	2 (11%)	24,24,24	5.07	5 (20%)
2	EPG	C	601	1	17,17,17	6.17	2 (11%)	24,24,24	5.01	6 (25%)
2	EPG	D	601	1	17,17,17	5.94	2 (11%)	24,24,24	5.04	5 (20%)

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	EPG	A	601	1	-	1/7/29/29	0/2/2/2
2	EPG	B	601	1	-	1/7/29/29	0/2/2/2
2	EPG	C	601	1	-	2/7/29/29	0/2/2/2
2	EPG	D	601	1	-	1/7/29/29	0/2/2/2

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	601	EPG	O8-C9	24.93	2.46	1.43
2	B	601	EPG	O8-C9	24.40	2.43	1.43
2	D	601	EPG	O8-C9	24.24	2.43	1.43
2	A	601	EPG	O8-C9	24.16	2.42	1.43
2	C	601	EPG	C9-C8	3.56	1.56	1.44

The worst 5 of 21 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	601	EPG	O8-C8-C9	19.75	110.63	60.04
2	D	601	EPG	O8-C8-C9	19.65	110.36	60.04
2	A	601	EPG	O8-C8-C9	19.64	110.35	60.04
2	C	601	EPG	O8-C8-C9	19.39	109.70	60.04
2	C	601	EPG	O8-C9-C8	-10.60	33.53	59.62

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	601	EPG	O1-C7-C8-C9
2	A	601	EPG	O1-C7-C8-C9
2	D	601	EPG	O1-C7-C8-C9
2	C	601	EPG	O1-C7-C8-C9
2	C	601	EPG	C4-C5-C6-O6

There are no ring outliers.

No monomer is involved in short contacts.

## 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 [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	516/516 (100%)	-0.85	1 (0%) 95 94	2, 13, 33, 53	0
1	B	516/516 (100%)	-0.62	7 (1%) 75 74	3, 16, 42, 61	0
1	C	516/516 (100%)	-0.71	1 (0%) 95 94	3, 15, 38, 56	0
1	D	516/516 (100%)	-0.72	1 (0%) 95 94	3, 15, 38, 48	0
All	All	2064/2064 (100%)	-0.72	10 (0%) 91 90	2, 15, 39, 61	0

The worst 5 of 10 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	463	HIS	3.5
1	B	433	ILE	3.5
1	B	462	SER	3.0
1	C	463	HIS	2.8
1	B	489	ASP	2.4

### 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 monosaccharides in this entry.

### 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	EPG	B	601	16/16	0.94	0.09	9,16,28,28	0
2	EPG	C	601	16/16	0.97	0.08	7,12,25,26	0
2	EPG	A	601	16/16	0.97	0.08	4,12,21,22	0
2	EPG	D	601	16/16	0.97	0.07	9,13,23,24	0
3	CA	D	701	1/1	0.99	0.09	12,12,12,12	0
3	CA	B	701	1/1	1.00	0.09	3,3,3,3	0
3	CA	C	701	1/1	1.00	0.12	6,6,6,6	0
3	CA	A	701	1/1	1.00	0.09	7,7,7,7	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.