



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 18, 2024 – 11:17 AM JST

PDB ID : 5Z72  
Title : Crystal structure of CcpC regulatory domain in complex with citrate from *Bacillus amyloliquefaciens*  
Authors : Chen, J.; Wang, L.; Shang, F.; Xu, Y.  
Deposited on : 2018-01-26  
Resolution : 2.40 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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 : 2.36  
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) : 2.36

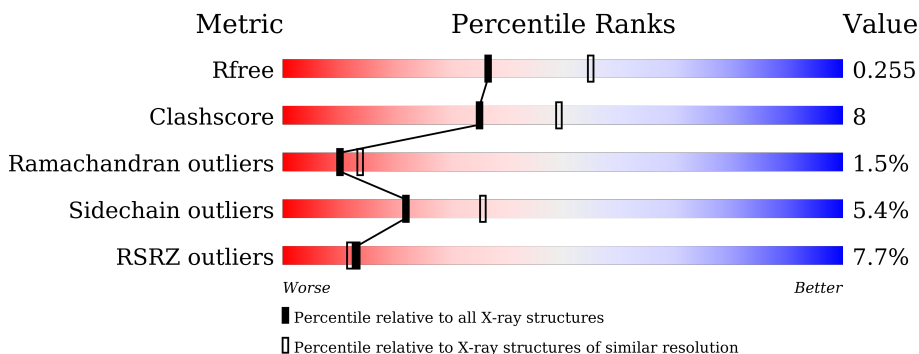
# 1 Overall quality at a glance

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 of 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	293	 % 55% 11% • 33%
1	B	293	 58% 8% • 33%
1	C	293	 3% 53% 12% • 33%
1	D	293	 9% 51% 13% • 34%
1	E	293	 12% 40% 18% • 39%

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
2	NA	E	301	-	-	-	X
3	FLC	B	304	-	X	-	-
3	FLC	D	305	-	X	-	-

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 16114 atoms, of which 7875 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.

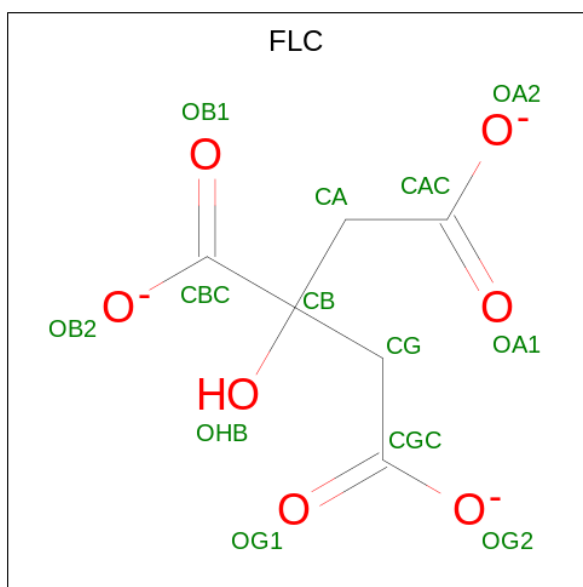
- Molecule 1 is a protein called CcpC.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	H	N	O				S
1	A	197	3230	1047	1610	268	298	7	0	1	0
1	B	197	3213	1042	1602	266	296	7	0	0	0
1	C	197	3213	1042	1602	266	296	7	0	0	0
1	D	193	3163	1027	1582	261	286	7	0	0	0
1	E	179	2901	941	1454	234	265	7	0	0	0

- Molecule 2 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	2	Total	Na	0	0
			2	2		
2	B	3	Total	Na	0	0
			3	3		
2	C	2	Total	Na	0	0
			2	2		
2	D	4	Total	Na	0	0
			4	4		
2	E	1	Total	Na	0	0
			1	1		

- Molecule 3 is CITRATE ANION (three-letter code: FLC) (formula: C<sub>6</sub>H<sub>5</sub>O<sub>7</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	H	O		
3	A	1	18	6	5	7	0	0
3	B	1	18	6	5	7	0	0
3	C	1	18	6	5	7	0	0
3	D	1	18	6	5	7	0	0
3	E	1	18	6	5	7	0	0

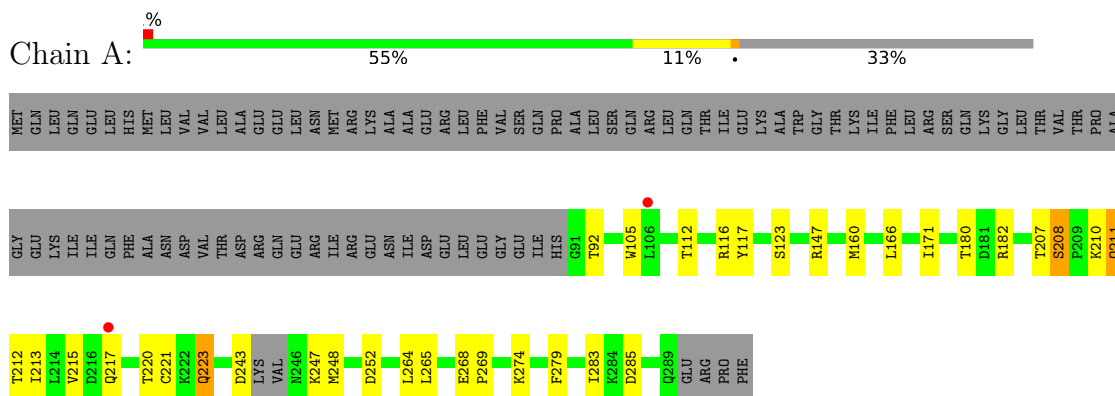
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	85	85	85	0	0
4	B	81	81	81	0	0
4	C	69	69	69	0	0
4	D	33	33	33	0	0
4	E	24	24	24	0	0

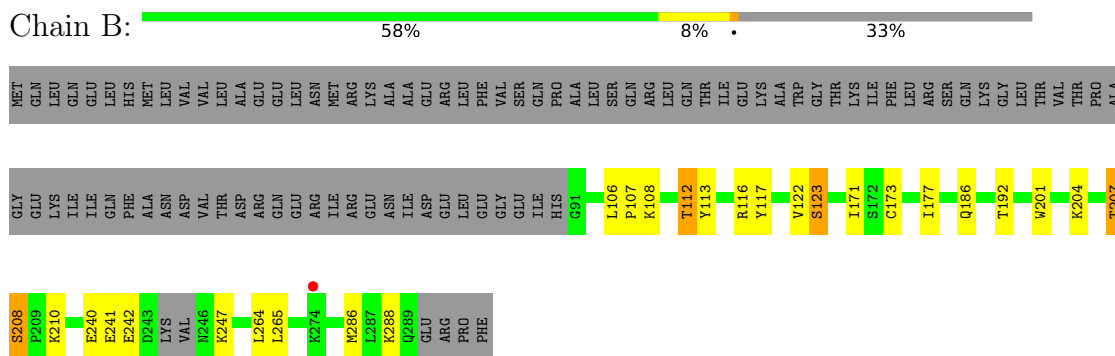
### 3 Residue-property plots [i](#)

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.

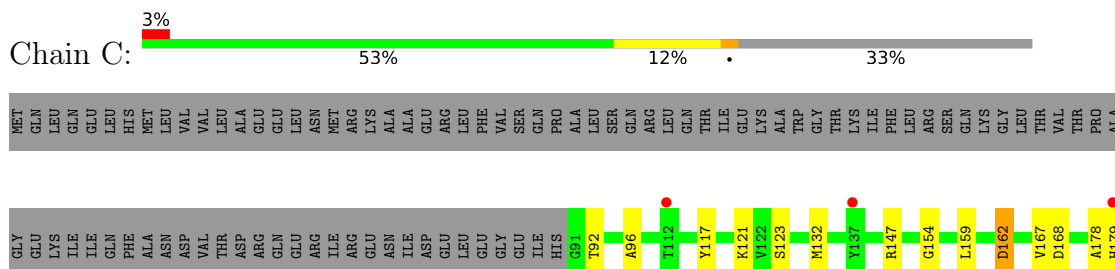
- Molecule 1: CcpC

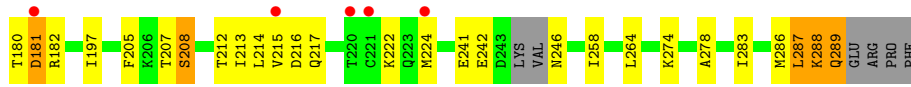


- Molecule 1: CcpC

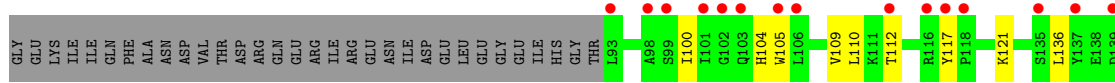


- Molecule 1: CcpC

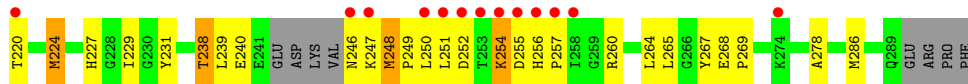
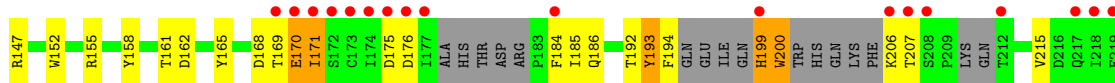
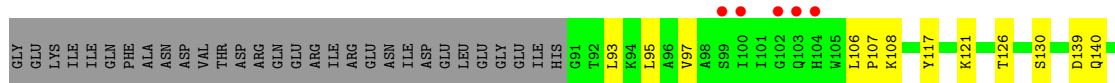




● Molecule 1: CcpC



● Molecule 1: CcpC



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	140.96Å 90.90Å 105.53Å 90.00° 106.18° 90.00°	Depositor
Resolution (Å)	29.71 – 2.40 39.34 – 2.29	Depositor EDS
% Data completeness (in resolution range)	98.1 (29.71-2.40) 96.5 (39.34-2.29)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.08	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.76 (at 2.29Å)	Xtrriage
Refinement program	PHENIX 1.8.4_1496	Depositor
R, $R_{free}$	0.185 , 0.255 0.193 , 0.255	Depositor DCC
$R_{free}$ test set	2000 reflections (3.59%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	48.6	Xtrriage
Anisotropy	0.438	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 56.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	16114	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	75.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.19% 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: FLC, NA

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	0.82	1/1661 (0.1%)	0.80	1/2251 (0.0%)
1	B	0.82	0/1652	0.80	0/2239
1	C	0.70	0/1652	0.78	1/2239 (0.0%)
1	D	0.66	0/1622	0.75	1/2199 (0.0%)
1	E	0.66	0/1479	0.73	1/1999 (0.1%)
All	All	0.74	1/8066 (0.0%)	0.77	4/10927 (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	C	0	2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	221	CYS	CB-SG	-5.06	1.73	1.81

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	D	181	ASP	CB-CG-OD1	-9.23	110.00	118.30
1	E	246	ASN	CA-C-O	-6.21	107.07	120.10
1	A	252	ASP	CB-CG-OD1	5.82	123.54	118.30
1	C	154	GLY	N-CA-C	-5.07	100.42	113.10

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	179	HIS	Peptide
1	C	258	ILE	Peptide

## 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	1620	1610	1615	21	1
1	B	1611	1602	1608	17	0
1	C	1611	1602	1608	26	1
1	D	1581	1582	1589	28	0
1	E	1447	1454	1457	37	0
2	A	2	0	0	0	0
2	B	3	0	0	0	0
2	C	2	0	0	0	0
2	D	4	0	0	0	0
2	E	1	0	0	0	0
3	A	13	5	5	0	0
3	B	13	5	5	1	0
3	C	13	5	5	2	0
3	D	13	5	5	2	0
3	E	13	5	5	0	0
4	A	85	0	0	5	0
4	B	81	0	0	5	0
4	C	69	0	0	4	1
4	D	33	0	0	2	0
4	E	24	0	0	3	1
All	All	8239	7875	7902	130	2

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 8.

The worst 5 of 130 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:E:286:MET:SD	4:E:421:HOH:O	2.22	0.97
1:B:286:MET:SD	4:B:457:HOH:O	2.45	0.74

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:185:ILE:HD13	1:E:224:MET:HB3	1.70	0.73
1:C:216:ASP:OD2	4:C:401:HOH:O	2.06	0.73
1:A:210:LYS:O	4:A:401:HOH:O	2.09	0.70

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:180:THR:O	1:C:208:SER:OG[4_455]	2.17	0.03
4:C:430:HOH:O	4:E:402:HOH:O[2_455]	2.17	0.03

## 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	194/293 (66%)	183 (94%)	10 (5%)	1 (0%)	29	41
1	B	193/293 (66%)	180 (93%)	12 (6%)	1 (0%)	29	41
1	C	193/293 (66%)	180 (93%)	10 (5%)	3 (2%)	9	13
1	D	189/293 (64%)	170 (90%)	14 (7%)	5 (3%)	5	5
1	E	167/293 (57%)	151 (90%)	12 (7%)	4 (2%)	6	6
All	All	936/1465 (64%)	864 (92%)	58 (6%)	14 (2%)	10	14

5 of 14 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	181	ASP
1	C	288	LYS
1	D	245	VAL
1	E	170	GLU
1	E	171	ILE

### 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	180/264 (68%)	171 (95%)	9 (5%)	24	40
1	B	179/264 (68%)	171 (96%)	8 (4%)	27	44
1	C	179/264 (68%)	171 (96%)	8 (4%)	27	44
1	D	176/264 (67%)	168 (96%)	8 (4%)	27	44
1	E	162/264 (61%)	148 (91%)	14 (9%)	10	16
All	All	876/1320 (66%)	829 (95%)	47 (5%)	22	36

5 of 47 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	208	SER
1	E	139	ASP
1	D	214	LEU
1	D	286	MET
1	E	186	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	186	GLN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 17 ligands modelled in this entry, 12 are monoatomic - leaving 5 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
3	FLC	E	302	2	12,12,12	1.23	1 (8%)	17,17,17	2.04	3 (17%)
3	FLC	C	303	2	12,12,12	1.25	0	17,17,17	1.30	3 (17%)
3	FLC	B	304	2	12,12,12	2.16	3 (25%)	17,17,17	3.92	7 (41%)
3	FLC	A	303	2	12,12,12	1.09	0	17,17,17	2.16	4 (23%)
3	FLC	D	305	2	12,12,12	1.27	1 (8%)	17,17,17	2.21	7 (41%)

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
3	FLC	E	302	2	-	6/16/16/16	-
3	FLC	C	303	2	-	6/16/16/16	-
3	FLC	B	304	2	-	9/16/16/16	-
3	FLC	A	303	2	-	6/16/16/16	-
3	FLC	D	305	2	-	11/16/16/16	-

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	304	FLC	CB-CBC	5.42	1.59	1.53
3	E	302	FLC	CB-CBC	-2.53	1.50	1.53
3	B	304	FLC	OHB-CB	-2.48	1.38	1.43

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	304	FLC	CA-CB	-2.42	1.50	1.53
3	D	305	FLC	CB-CBC	2.16	1.55	1.53

The worst 5 of 24 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	304	FLC	OHB-CB-CA	-10.27	85.37	109.40
3	B	304	FLC	OB2-CBC-CB	7.49	126.06	113.05
3	B	304	FLC	CA-CB-CBC	6.17	123.36	110.11
3	A	303	FLC	OB2-CBC-CB	5.77	123.07	113.05
3	E	302	FLC	OB2-CBC-CB	5.14	121.98	113.05

There are no chirality outliers.

5 of 38 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	303	FLC	CG-CB-CBC-OB1
3	A	303	FLC	CG-CB-CBC-OB2
3	A	303	FLC	OHB-CB-CBC-OB1
3	A	303	FLC	OHB-CB-CBC-OB2
3	B	304	FLC	CA-CB-CBC-OB1

There are no ring outliers.

3 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	303	FLC	2	0
3	B	304	FLC	1	0
3	D	305	FLC	2	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 [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	197/293 (67%)	0.01	2 (1%) 82 80	37, 55, 84, 109	0
1	B	197/293 (67%)	-0.07	1 (0%) 91 89	39, 52, 85, 107	0
1	C	197/293 (67%)	0.16	8 (4%) 37 36	45, 62, 94, 117	0
1	D	193/293 (65%)	0.66	27 (13%) 2 2	50, 74, 121, 152	0
1	E	179/293 (61%)	0.79	36 (20%) 1 0	46, 80, 148, 185	0
All	All	963/1465 (65%)	0.30	74 (7%) 13 12	37, 62, 117, 185	0

The worst 5 of 74 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	253	THR	6.2
1	D	153	LYS	5.8
1	D	152	TRP	4.8
1	E	174	ILE	4.7
1	E	199	HIS	4.6

### 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(Å <sup>2</sup> )	Q<0.9
2	NA	E	301	1/1	0.79	0.42	85,85,85,85	0
3	FLC	D	305	13/13	0.85	0.22	67,76,97,97	0
3	FLC	E	302	13/13	0.86	0.22	85,94,115,122	0
2	NA	D	302	1/1	0.88	0.12	84,84,84,84	0
3	FLC	B	304	13/13	0.90	0.22	45,54,75,75	0
2	NA	A	302	1/1	0.91	0.18	66,66,66,66	0
2	NA	D	301	1/1	0.92	0.24	63,63,63,63	0
2	NA	C	302	1/1	0.93	0.74	78,78,78,78	0
2	NA	D	304	1/1	0.93	0.75	93,93,93,93	0
2	NA	B	302	1/1	0.93	0.20	54,54,54,54	0
2	NA	D	303	1/1	0.94	0.17	69,69,69,69	0
2	NA	A	301	1/1	0.94	0.19	55,55,55,55	0
3	FLC	C	303	13/13	0.95	0.17	47,58,70,70	0
2	NA	C	301	1/1	0.96	0.21	53,53,53,53	0
3	FLC	A	303	13/13	0.96	0.17	42,48,60,60	0
2	NA	B	301	1/1	0.98	0.11	55,55,55,55	0
2	NA	B	303	1/1	0.98	0.07	58,58,58,58	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.