



wwPDB X-ray Structure Validation Summary Report ⓘ

May 25, 2020 – 11:09 am BST

PDB ID : 2GIF
Title : Asymmetric structure of trimeric AcrB from Escherichia coli
Authors : Seeger, M.A.; Schiefner, A.; Eicher, T.; Verrey, F.; Diederichs, K.; Pos, K.M.
Deposited on : 2006-03-28
Resolution : 2.90 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

<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	:	2.11
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)	:	2.11

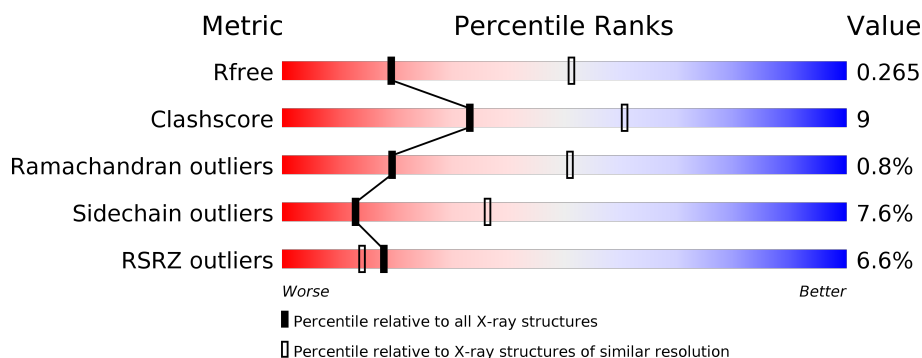
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.90 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	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (2.90-2.90)

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 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	1057	<div> <div>7%</div> <div> <div></div> <div>74%</div> <div>21%</div> <div>••</div> </div> </div>
1	B	1057	<div> <div>7%</div> <div> <div></div> <div>77%</div> <div>20%</div> <div>••</div> </div> </div>
1	C	1057	<div> <div>5%</div> <div> <div></div> <div>73%</div> <div>22%</div> <div>••</div> </div> </div>

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 23650 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 Acriflavine resistance protein B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1032	Total	C	N	O	S	0	0	0
			7841	5047	1294	1457	43			
1	B	1044	Total	C	N	O	S	0	0	0
			7942	5105	1315	1479	43			
1	C	1032	Total	C	N	O	S	0	0	0
			7841	5047	1294	1457	43			

There are 24 discrepancies between the modelled and reference sequences:

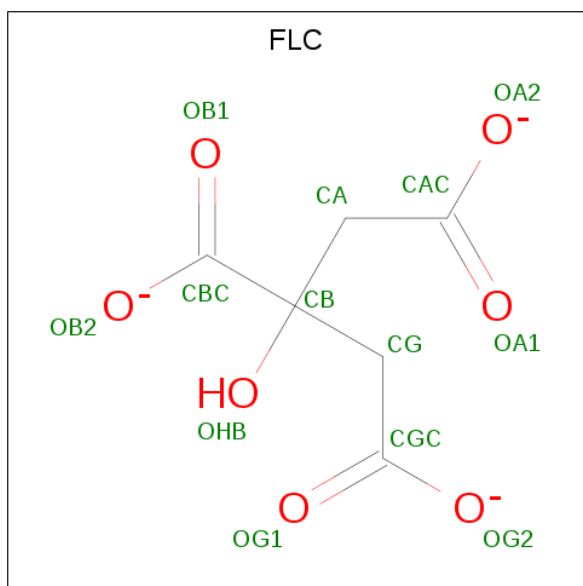
Chain	Residue	Modelled	Actual	Comment	Reference
A	1050	LEU	-	CLONING ARTIFACT	UNP P31224
A	1051	GLU	-	CLONING ARTIFACT	UNP P31224
A	1052	HIS	-	EXPRESSION TAG	UNP P31224
A	1053	HIS	-	EXPRESSION TAG	UNP P31224
A	1054	HIS	-	EXPRESSION TAG	UNP P31224
A	1055	HIS	-	EXPRESSION TAG	UNP P31224
A	1056	HIS	-	EXPRESSION TAG	UNP P31224
A	1057	HIS	-	EXPRESSION TAG	UNP P31224
B	1050	LEU	-	CLONING ARTIFACT	UNP P31224
B	1051	GLU	-	CLONING ARTIFACT	UNP P31224
B	1052	HIS	-	EXPRESSION TAG	UNP P31224
B	1053	HIS	-	EXPRESSION TAG	UNP P31224
B	1054	HIS	-	EXPRESSION TAG	UNP P31224
B	1055	HIS	-	EXPRESSION TAG	UNP P31224
B	1056	HIS	-	EXPRESSION TAG	UNP P31224
B	1057	HIS	-	EXPRESSION TAG	UNP P31224
C	1050	LEU	-	CLONING ARTIFACT	UNP P31224
C	1051	GLU	-	CLONING ARTIFACT	UNP P31224
C	1052	HIS	-	EXPRESSION TAG	UNP P31224
C	1053	HIS	-	EXPRESSION TAG	UNP P31224
C	1054	HIS	-	EXPRESSION TAG	UNP P31224
C	1055	HIS	-	EXPRESSION TAG	UNP P31224
C	1056	HIS	-	EXPRESSION TAG	UNP P31224

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	1057	HIS	-	EXPRESSION TAG	UNP P31224

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

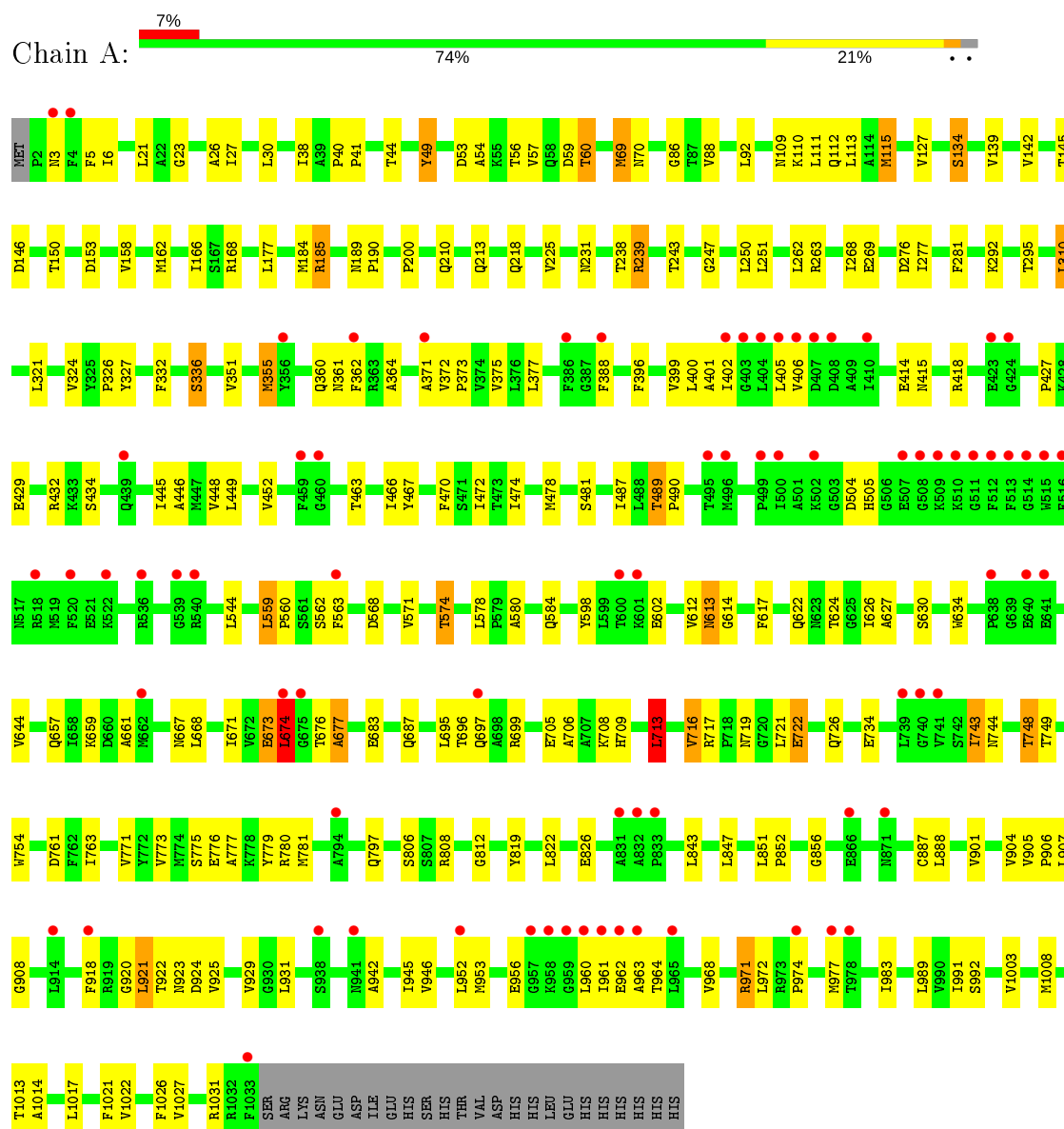


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			13	6	7		
2	B	1	Total	C	O	0	0
			13	6	7		

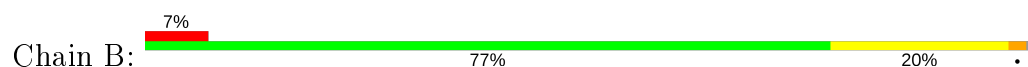
3 Residue-property plots [i](#)

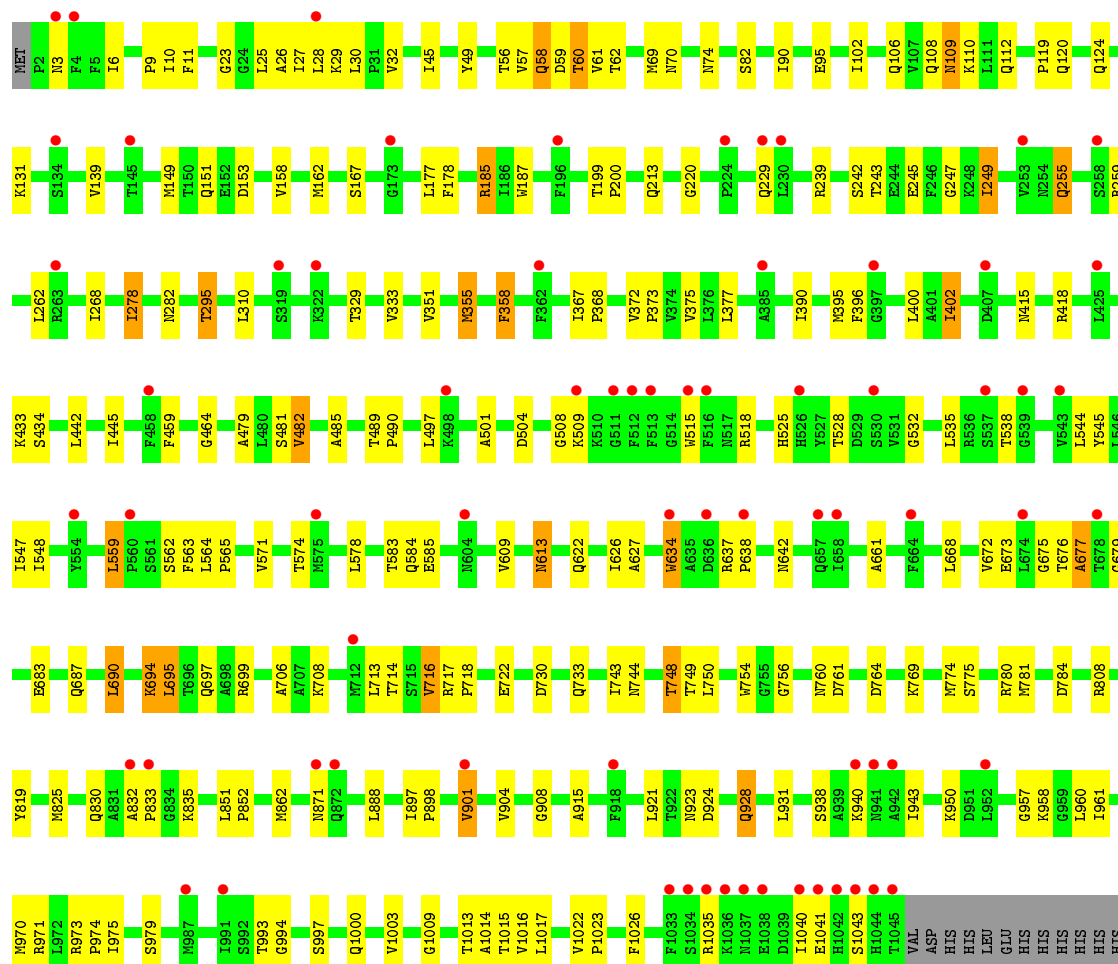
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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: Acriflavine resistance protein B

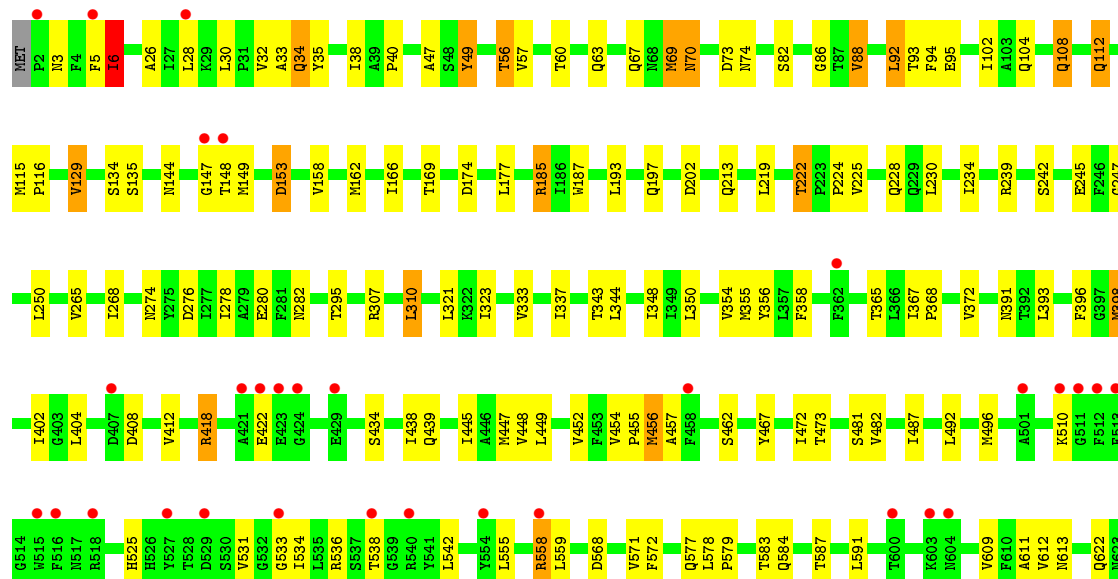
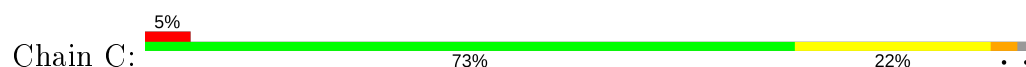


• Molecule 1: Acriflavine resistance protein B





• Molecule 1: Acriflavine resistance protein B





4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	222.80Å 134.10Å 161.01Å 90.00° 98.21° 90.00°	Depositor
Resolution (Å)	29.51 – 2.90 29.22 – 2.90	Depositor EDS
% Data completeness (in resolution range)	99.8 (29.51-2.90) 99.8 (29.22-2.90)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.70 (at 2.90Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.226 , 0.267 0.226 , 0.265	Depositor DCC
R_{free} test set	5177 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	89.1	Xtriage
Anisotropy	0.101	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 87.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	23650	wwPDB-VP
Average B, all atoms (Å ²)	73.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.79% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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: FLC

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	0.37	0/7991	0.54	1/10852 (0.0%)
1	B	0.37	0/8094	0.53	0/10990
1	C	0.36	1/7991 (0.0%)	0.53	1/10852 (0.0%)
All	All	0.37	1/24076 (0.0%)	0.53	2/32694 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	820	ASN	C-N	5.23	1.42	1.33

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	713	LEU	CA-CB-CG	5.82	128.69	115.30
1	C	92	LEU	CA-CB-CG	5.06	126.94	115.30

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	7841	0	7990	152	0
1	B	7942	0	8080	124	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	7841	0	7990	154	0
2	A	13	0	5	0	0
2	B	13	0	5	0	0
All	All	23650	0	24070	410	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 9.

The worst 5 of 410 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:781:MET:HE2	1:C:225:VAL:H	1.06	1.12
1:A:239:ARG:HH11	1:A:239:ARG:HG3	1.24	1.03
1:B:185:ARG:HH11	1:B:185:ARG:HG3	1.24	1.00
1:A:225:VAL:H	1:B:781:MET:HE2	1.27	0.95
1:C:456:MET:HG3	1:C:467:TYR:HB3	1.49	0.94

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	1030/1057 (97%)	963 (94%)	57 (6%)	10 (1%)	15	45
1	B	1042/1057 (99%)	984 (94%)	50 (5%)	8 (1%)	19	51
1	C	1030/1057 (97%)	965 (94%)	58 (6%)	7 (1%)	22	54
All	All	3102/3171 (98%)	2912 (94%)	165 (5%)	25 (1%)	19	51

5 of 25 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	115	MET
1	A	674	LEU
1	A	677	ALA
1	A	992	SER
1	B	634	TRP

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	838/863 (97%)	781 (93%)	57 (7%)	16	42
1	B	850/863 (98%)	787 (93%)	63 (7%)	13	38
1	C	838/863 (97%)	766 (91%)	72 (9%)	10	30
All	All	2526/2589 (98%)	2334 (92%)	192 (8%)	13	36

5 of 192 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	402	ILE
1	B	722	GLU
1	C	837	THR
1	B	515	TRP
1	B	673	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 66 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	197	GLN
1	B	605	ASN
1	C	622	GLN
1	B	213	GLN
1	B	284	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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

2 ligands are modelled in this entry.

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	1058	-	3,12,12	0.59	0	3,17,17	0.70	0
2	FLC	B	1058	-	3,12,12	0.50	0	3,17,17	1.75	1 (33%)

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	1058	-	-	2/6/16/16	-
2	FLC	B	1058	-	-	2/6/16/16	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
2	B	1058	FLC	CB-CA-CAC	-2.80	110.51	114.98

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	1058	FLC	OHB-CB-CG-CGC
2	A	1058	FLC	CA-CB-CG-CGC
2	B	1058	FLC	CAC-CA-CB-CBC
2	B	1058	FLC	CAC-CA-CB-OHB

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, 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	1032/1057 (97%)	0.26	77 (7%) 14 11	59, 73, 83, 88	0
1	B	1044/1057 (98%)	0.30	70 (6%) 17 13	66, 74, 82, 87	0
1	C	1032/1057 (97%)	0.24	58 (5%) 24 20	65, 74, 79, 82	0
All	All	3108/3171 (98%)	0.27	205 (6%) 18 14	59, 73, 81, 88	0

The worst 5 of 205 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	513	PHE	7.8
1	C	511	GLY	7.6
1	C	512	PHE	6.6
1	C	604	ASN	6.4
1	A	515	TRP	6.1

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. 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	B-factors(\AA^2)	Q<0.9
2	FLC	A	1058	13/13	0.88	0.29	64,65,66,67	0
2	FLC	B	1058	13/13	0.94	0.26	48,51,52,53	0

6.5 Other polymers [i](#)

There are no such residues in this entry.