wwPDB X-ray Structure Validation Summary Report (i)

Aug 9, 2020 – 10:23 AM BST

PDB ID	:	7C3M
Title	:	Structure of FERM protein
Authors	:	Bu, W.; Loh, Z.Y.; Jin, S.; Basu, S.; Ero, R.; Park, J.E.; Yan, X.; Wang, M.; Sze, S.K.; Tan, S.M.; Gao, Y.G.
Deposited on	:	2020-05-13
Resolution	:	3.60 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 (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

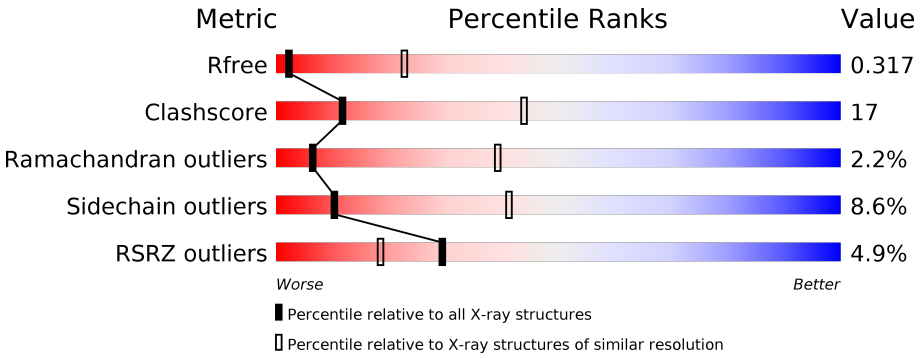
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.13.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.13.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.60 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	1257 (3.70-3.50)
Clashscore	141614	1353 (3.70-3.50)
Ramachandran outliers	138981	1307 (3.70-3.50)
Sidechain outliers	138945	1307 (3.70-3.50)
RSRZ outliers	127900	1161 (3.70-3.50)

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 <=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	675	3% 56% 26% • 15%
1	B	675	4% 53% 27% • 16%
1	C	675	6% 52% 27% 5% 16%

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 13538 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 Fermitin family homolog 3,Fermitin family homolog 3,Fermitin family homolog 3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	577	Total	C	N	O	S	0	0	0
			4575	2909	825	824	17			
1	B	570	Total	C	N	O	S	0	0	0
			4473	2837	805	811	20			
1	C	570	Total	C	N	O	S	0	0	0
			4490	2852	807	811	20			

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	668	LEU	-	expression tag	UNP Q86UX7
A	669	GLU	-	expression tag	UNP Q86UX7
A	670	HIS	-	expression tag	UNP Q86UX7
A	671	HIS	-	expression tag	UNP Q86UX7
A	672	HIS	-	expression tag	UNP Q86UX7
A	673	HIS	-	expression tag	UNP Q86UX7
A	674	HIS	-	expression tag	UNP Q86UX7
A	675	HIS	-	expression tag	UNP Q86UX7
B	668	LEU	-	expression tag	UNP Q86UX7
B	669	GLU	-	expression tag	UNP Q86UX7
B	670	HIS	-	expression tag	UNP Q86UX7
B	671	HIS	-	expression tag	UNP Q86UX7
B	672	HIS	-	expression tag	UNP Q86UX7
B	673	HIS	-	expression tag	UNP Q86UX7
B	674	HIS	-	expression tag	UNP Q86UX7
B	675	HIS	-	expression tag	UNP Q86UX7
C	668	LEU	-	expression tag	UNP Q86UX7
C	669	GLU	-	expression tag	UNP Q86UX7
C	670	HIS	-	expression tag	UNP Q86UX7
C	671	HIS	-	expression tag	UNP Q86UX7
C	672	HIS	-	expression tag	UNP Q86UX7
C	673	HIS	-	expression tag	UNP Q86UX7

Continued on next page...

Continued from previous page...

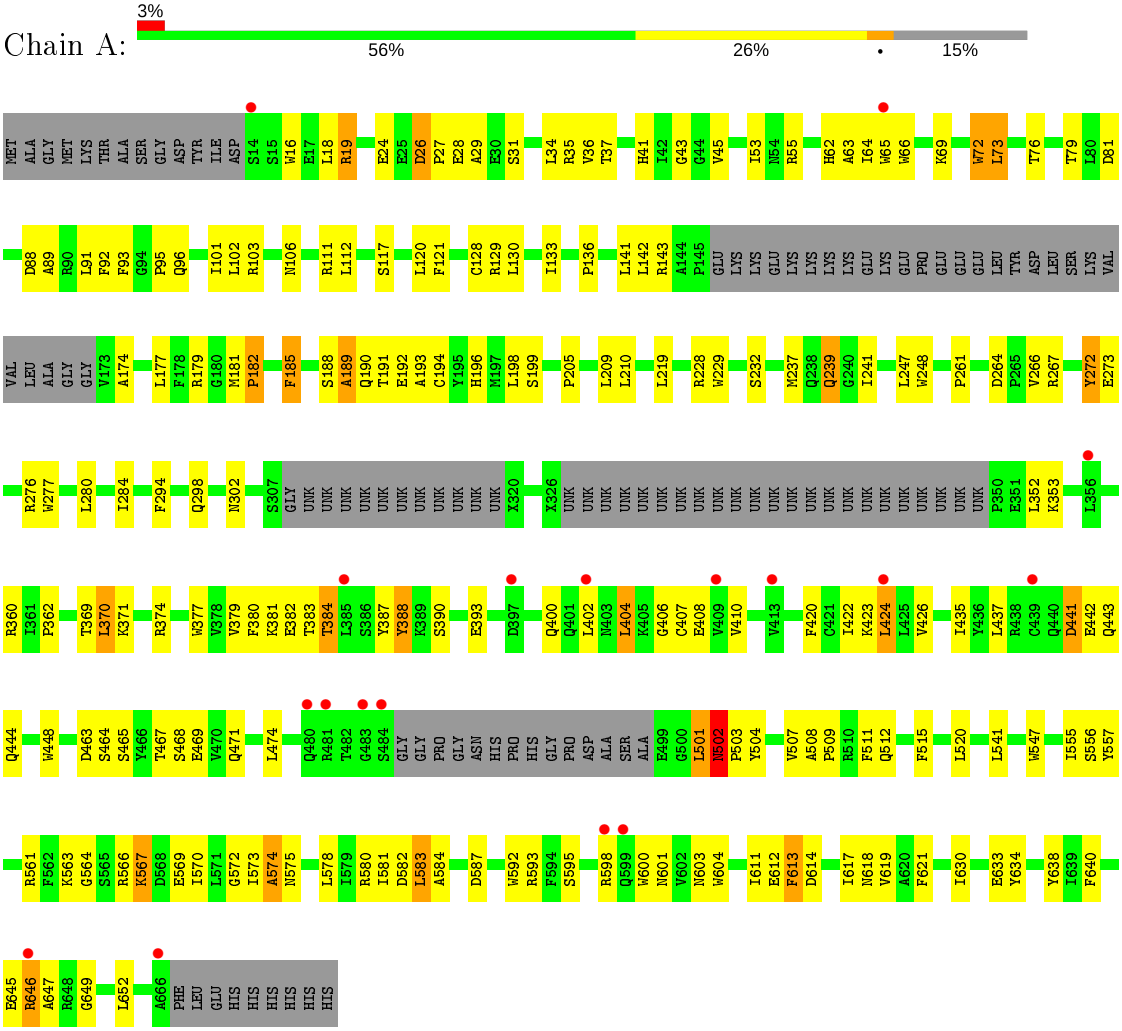
Chain	Residue	Modelled	Actual	Comment	Reference
C	674	HIS	-	expression tag	UNP Q86UX7
C	675	HIS	-	expression tag	UNP Q86UX7

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.

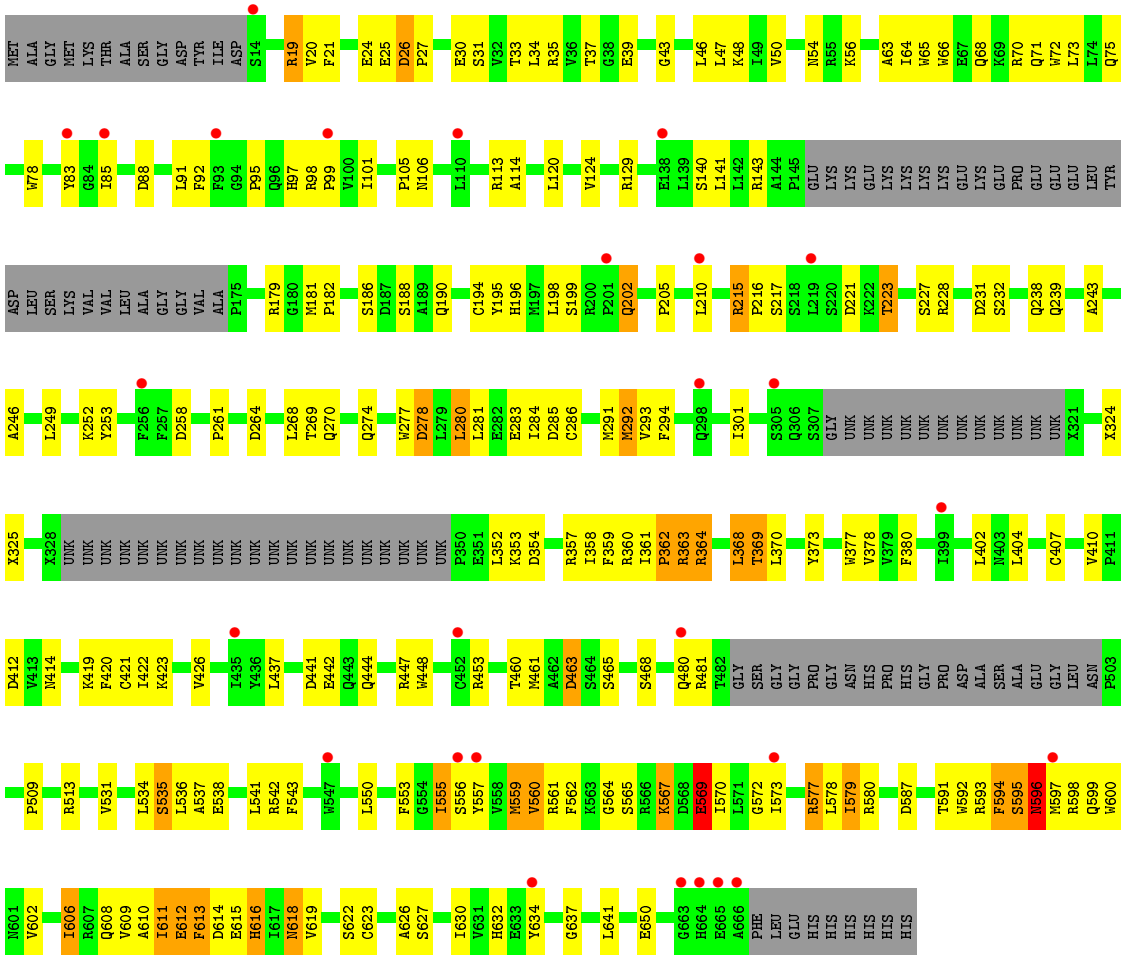
• Molecule 1: Fermitin family homolog 3,Fermitin family homolog 3,Fermitin family homolog 3

Chain A:



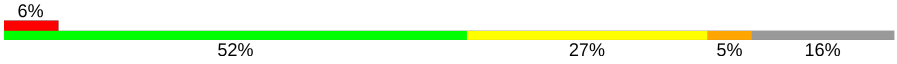
• Molecule 1: Fermitin family homolog 3,Fermitin family homolog 3,Fermitin family homolog 3

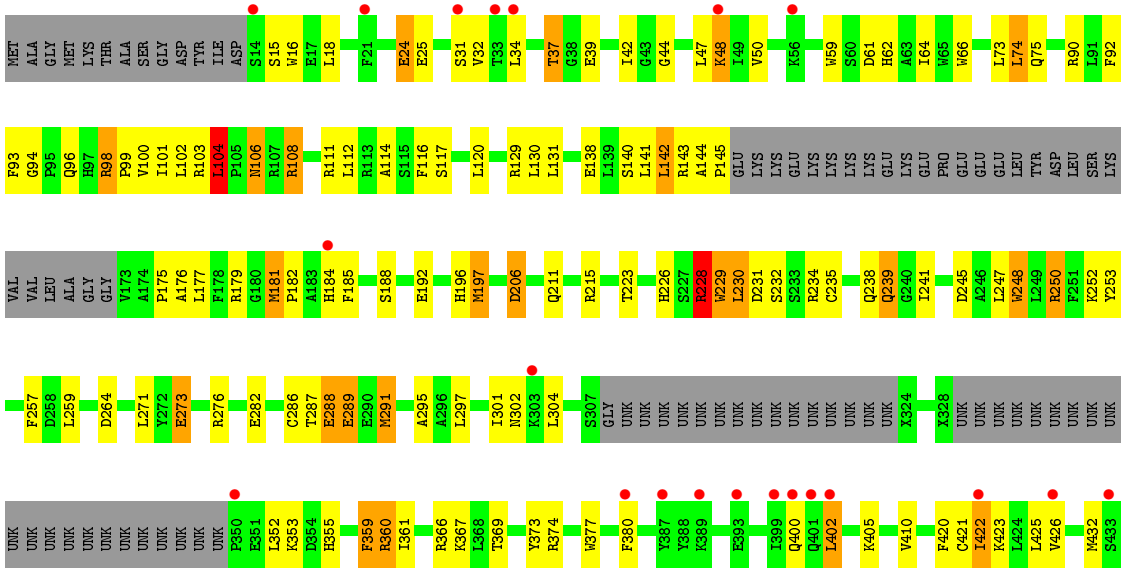
Chain B:

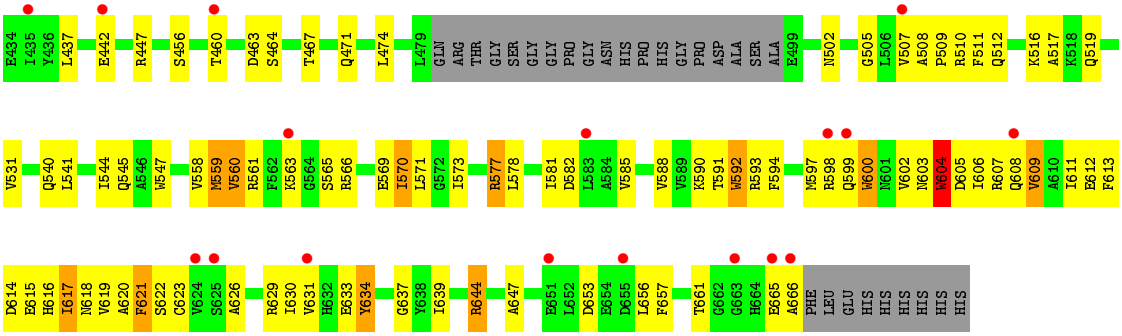


• Molecule 1: Fermitin family homolog 3,Fermitin family homolog 3,Fermitin family homolog 3

Chain C:






E434 I435 Y436 L437 E442 R447 S456 T460 D463 S464 T467 Q471 L474 L479 GLN ARG THR GLY SER GLY GLY PRO GLY ASN HIS PRO HIS GLY PRO ASP ALA SER ALA E499 N502 G505 L506 V507 A508 P509 R510 F511 Q512 K516 A517 K518 Q519
V531 Q540 L541 I544 Q545 A546 W547 V558 M559 V560 R561 F562 K563 G564 S565 R566 E569 I570 L571 G572 I573 R577 L578 I581 D582 L583 A584 V585 V588 V589 K590 T591 W592 R593 F594 M597 R598 Q599 W600 N601 V602 N603 W604 D605 I606 R607 Q608 V609 A610 I611 E612 F613
D614 E615 H616 I617 N618 V619 A620 F621 S622 C623 V624 S625 A626 R629 I630 V631 H632 E633 Y634 G637 Y638 I639 R644 A647 E651 L652 D653 E654 D655 L656 F657 T661 G662 G663 H664 E665 A666 PHE LEU GLU HIS HIS HIS HIS HIS HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	115.76Å 204.62Å 269.95Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.20 – 3.60 49.20 – 3.60	Depositor EDS
% Data completeness (in resolution range)	99.9 (49.20-3.60) 100.0 (49.20-3.60)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.10 (at 3.57Å)	Xtriage
Refinement program	PHENIX (1.13_2998)	Depositor
R, R_{free}	0.305 , 0.317 0.305 , 0.317	Depositor DCC
R_{free} test set	1771 reflections (4.72%)	wwPDB-VP
Wilson B-factor (Å^2)	170.3	Xtriage
Anisotropy	0.325	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 140.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.29$	Xtriage
Estimated twinning fraction	0.034 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.038 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	13538	wwPDB-VP
Average B, all atoms (Å^2)	173.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.76% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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

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	0.27	0/4648	0.54	1/6306 (0.0%)
1	B	0.30	0/4537	0.56	0/6157
1	C	0.28	0/4571	0.55	0/6202
All	All	0.28	0/13756	0.55	1/18665 (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	A	0	1
1	B	0	5
1	C	0	4
All	All	0	10

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	501	LEU	C-N-CA	5.53	135.52	121.70

There are no chirality outliers.

5 of 10 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	603	ASN	Peptide
1	B	324	UNK	Peptide
1	B	363	ARG	Peptide
1	B	564	GLY	Peptide
1	B	595	SER	Peptide

5.2 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	4575	0	4424	132	0
1	B	4473	0	4275	159	0
1	C	4490	0	4317	172	0
All	All	13538	0	13016	463	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 17.

The worst 5 of 463 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:B:181:MET:HB3	1:B:182:PRO:HD2	1.50	0.92
1:C:104:LEU:HB2	1:C:106:ASN:H	1.33	0.92
1:B:600:TRP:HB3	1:B:611:ILE:HG23	1.52	0.89
1:C:360:ARG:HD3	1:C:361:ILE:H	1.39	0.88
1:A:120:LEU:HB3	1:A:232:SER:HA	1.55	0.88

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	562/675 (83%)	476 (85%)	76 (14%)	10 (2%)	8 43
1	B	554/675 (82%)	455 (82%)	82 (15%)	17 (3%)	4 32
1	C	557/675 (82%)	469 (84%)	78 (14%)	10 (2%)	8 43

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	1673/2025 (83%)	1400 (84%)	236 (14%)	37 (2%)	6 39

5 of 37 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	502	ASN
1	B	536	LEU
1	B	565	SER
1	B	594	PHE
1	B	595	SER

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	471/551 (86%)	443 (94%)	28 (6%)	19 55
1	B	455/551 (83%)	416 (91%)	39 (9%)	10 41
1	C	461/551 (84%)	409 (89%)	52 (11%)	6 30
All	All	1387/1653 (84%)	1268 (91%)	119 (9%)	10 41

5 of 119 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	555	ILE
1	B	634	TYR
1	C	604	TRP
1	B	560	VAL
1	B	596	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	75	GLN
1	B	190	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	62	HIS
1	B	54	ASN
1	B	616	HIS

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

There are no ligands in this entry.

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^{th} 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	570/675 (84%)	0.30	18 (3%)	47	32	126, 154, 192, 224	0
1	B	562/675 (83%)	0.35	27 (4%)	30	19	143, 184, 225, 246	0
1	C	565/675 (83%)	0.46	38 (6%)	17	10	139, 179, 214, 227	0
All	All	1697/2025 (83%)	0.37	83 (4%)	29	18	126, 173, 215, 246	0

The worst 5 of 83 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	557	TYR	7.9
1	C	666	ALA	7.7
1	C	599	GLN	5.0
1	B	666	ALA	4.8
1	A	484	SER	4.8

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 monosaccharides in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.