wwPDB X-ray Structure Validation Summary Report (i)

Nov 14, 2023 – 12:44 PM JST

PDB ID	:	5ZAD
Title	:	Human topoisomerase II beta in complex with DNA
Authors	:	Sun, L.Y.; Zhu, L.W.; Tang, Y.J.
Deposited on	:	2018-02-07
Resolution	:	2.54 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

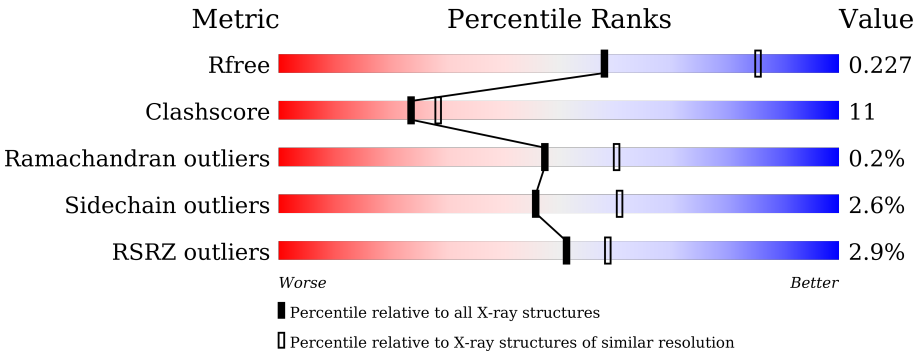
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.54 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	1284 (2.56-2.52)
Clashscore	141614	1332 (2.56-2.52)
Ramachandran outliers	138981	1315 (2.56-2.52)
Sidechain outliers	138945	1315 (2.56-2.52)
RSRZ outliers	127900	1272 (2.56-2.52)

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 >=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	803	2% 70% 19% • 9%
1	B	803	3% 70% 19% • 9%
2	C	8	100%
2	D	8	88% 12%
3	E	12	83% 17%
3	F	12	67% 33%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 12706 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 DNA topoisomerase 2-beta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	729	Total	C	N	O	S	0	3	0
			5903	3762	1012	1105	24			
1	B	729	Total	C	N	O	S	0	1	0
			5893	3756	1012	1101	24			

There are 92 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	419	MET	-	expression tag	UNP Q02880
A	420	ALA	-	expression tag	UNP Q02880
A	421	SER	-	expression tag	UNP Q02880
A	422	TRP	-	expression tag	UNP Q02880
A	423	SER	-	expression tag	UNP Q02880
A	424	HIS	-	expression tag	UNP Q02880
A	425	PRO	-	expression tag	UNP Q02880
A	426	GLN	-	expression tag	UNP Q02880
A	427	PHE	-	expression tag	UNP Q02880
A	428	GLU	-	expression tag	UNP Q02880
A	429	LYS	-	expression tag	UNP Q02880
A	430	GLY	-	expression tag	UNP Q02880
A	431	ALA	-	expression tag	UNP Q02880
A	432	ASP	-	expression tag	UNP Q02880
A	433	ASP	-	expression tag	UNP Q02880
A	434	ASP	-	expression tag	UNP Q02880
A	435	ASP	-	expression tag	UNP Q02880
A	436	LYS	-	expression tag	UNP Q02880
A	437	VAL	-	expression tag	UNP Q02880
A	438	PRO	-	expression tag	UNP Q02880
A	439	ASP	-	expression tag	UNP Q02880
A	440	PRO	-	expression tag	UNP Q02880
A	441	THR	-	expression tag	UNP Q02880
A	442	SER	-	expression tag	UNP Q02880
A	443	VAL	-	expression tag	UNP Q02880

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	444	ASP	-	expression tag	UNP Q02880
A	1202	GLY	-	expression tag	UNP Q02880
A	1203	ALA	-	expression tag	UNP Q02880
A	1204	PRO	-	expression tag	UNP Q02880
A	1205	GLY	-	expression tag	UNP Q02880
A	1206	PHE	-	expression tag	UNP Q02880
A	1207	SER	-	expression tag	UNP Q02880
A	1208	SER	-	expression tag	UNP Q02880
A	1209	ILE	-	expression tag	UNP Q02880
A	1210	SER	-	expression tag	UNP Q02880
A	1211	ALA	-	expression tag	UNP Q02880
A	1212	HIS	-	expression tag	UNP Q02880
A	1213	HIS	-	expression tag	UNP Q02880
A	1214	HIS	-	expression tag	UNP Q02880
A	1215	HIS	-	expression tag	UNP Q02880
A	1216	HIS	-	expression tag	UNP Q02880
A	1217	HIS	-	expression tag	UNP Q02880
A	1218	HIS	-	expression tag	UNP Q02880
A	1219	HIS	-	expression tag	UNP Q02880
A	1220	HIS	-	expression tag	UNP Q02880
A	1221	HIS	-	expression tag	UNP Q02880
B	419	MET	-	expression tag	UNP Q02880
B	420	ALA	-	expression tag	UNP Q02880
B	421	SER	-	expression tag	UNP Q02880
B	422	TRP	-	expression tag	UNP Q02880
B	423	SER	-	expression tag	UNP Q02880
B	424	HIS	-	expression tag	UNP Q02880
B	425	PRO	-	expression tag	UNP Q02880
B	426	GLN	-	expression tag	UNP Q02880
B	427	PHE	-	expression tag	UNP Q02880
B	428	GLU	-	expression tag	UNP Q02880
B	429	LYS	-	expression tag	UNP Q02880
B	430	GLY	-	expression tag	UNP Q02880
B	431	ALA	-	expression tag	UNP Q02880
B	432	ASP	-	expression tag	UNP Q02880
B	433	ASP	-	expression tag	UNP Q02880
B	434	ASP	-	expression tag	UNP Q02880
B	435	ASP	-	expression tag	UNP Q02880
B	436	LYS	-	expression tag	UNP Q02880
B	437	VAL	-	expression tag	UNP Q02880
B	438	PRO	-	expression tag	UNP Q02880
B	439	ASP	-	expression tag	UNP Q02880

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	440	PRO	-	expression tag	UNP Q02880
B	441	THR	-	expression tag	UNP Q02880
B	442	SER	-	expression tag	UNP Q02880
B	443	VAL	-	expression tag	UNP Q02880
B	444	ASP	-	expression tag	UNP Q02880
B	1202	GLY	-	expression tag	UNP Q02880
B	1203	ALA	-	expression tag	UNP Q02880
B	1204	PRO	-	expression tag	UNP Q02880
B	1205	GLY	-	expression tag	UNP Q02880
B	1206	PHE	-	expression tag	UNP Q02880
B	1207	SER	-	expression tag	UNP Q02880
B	1208	SER	-	expression tag	UNP Q02880
B	1209	ILE	-	expression tag	UNP Q02880
B	1210	SER	-	expression tag	UNP Q02880
B	1211	ALA	-	expression tag	UNP Q02880
B	1212	HIS	-	expression tag	UNP Q02880
B	1213	HIS	-	expression tag	UNP Q02880
B	1214	HIS	-	expression tag	UNP Q02880
B	1215	HIS	-	expression tag	UNP Q02880
B	1216	HIS	-	expression tag	UNP Q02880
B	1217	HIS	-	expression tag	UNP Q02880
B	1218	HIS	-	expression tag	UNP Q02880
B	1219	HIS	-	expression tag	UNP Q02880
B	1220	HIS	-	expression tag	UNP Q02880
B	1221	HIS	-	expression tag	UNP Q02880

- Molecule 2 is a DNA chain called DNA (5'-D(P*AP*GP*CP*CP*GP*AP*GP*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	8	Total	C	N	O	P	0	0	0
			165	77	34	46	8			
2	D	8	Total	C	N	O	P	0	0	0
			165	77	34	46	8			

- Molecule 3 is a DNA chain called DNA (5'-D(P*TP*GP*CP*AP*GP*CP*TP*CP*GP*GP*CP*T)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	F	12	Total	C	N	O	P	0	0	0
			245	116	43	74	12			
3	E	12	Total	C	N	O	P	0	0	0
			245	116	43	74	12			

- Molecule 4 is water.

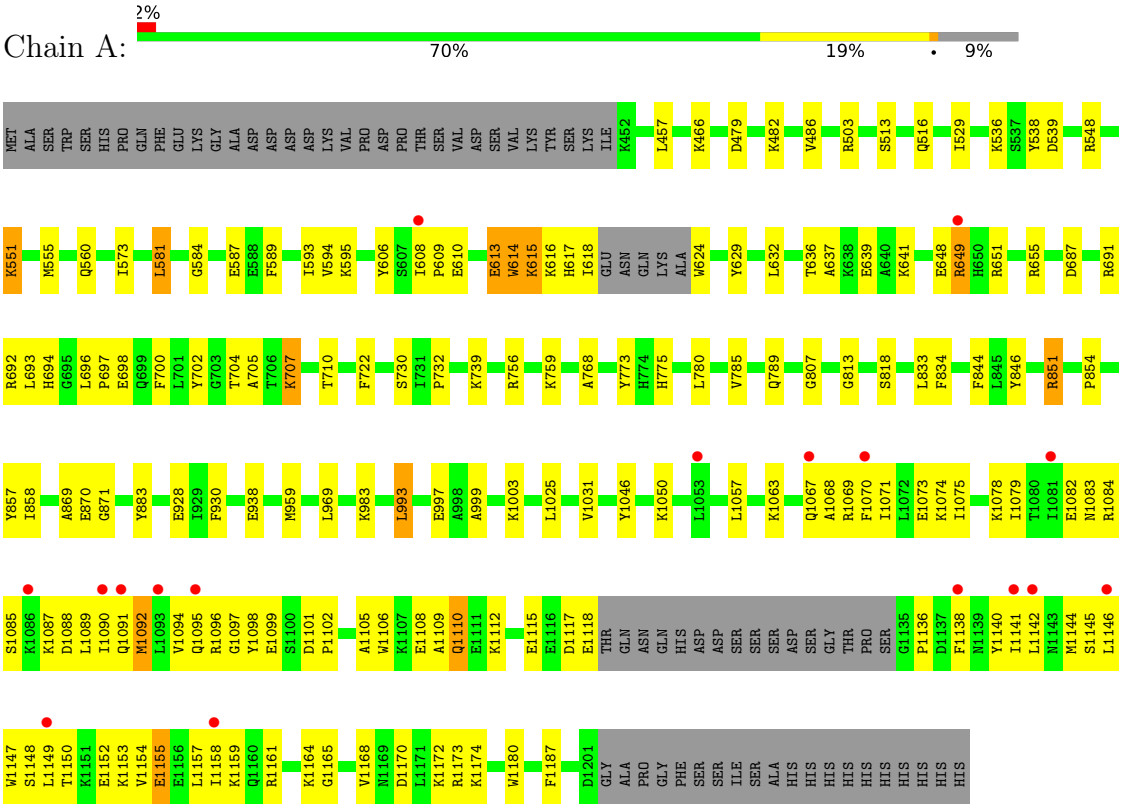
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	39	Total O 39 39	0	0
4	C	5	Total O 5 5	0	0
4	F	5	Total O 5 5	0	0
4	B	31	Total O 31 31	0	0
4	D	5	Total O 5 5	0	0
4	E	5	Total O 5 5	0	0

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.

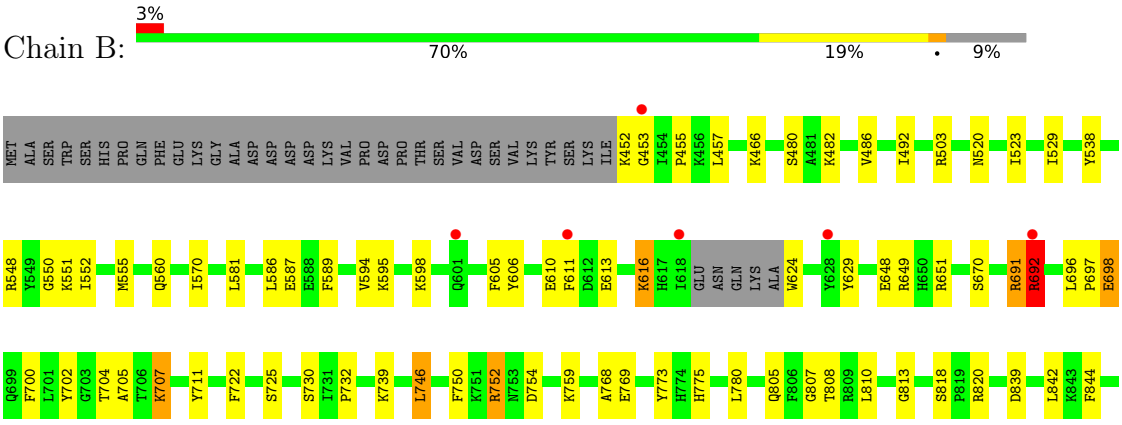
• Molecule 1: DNA topoisomerase 2-beta

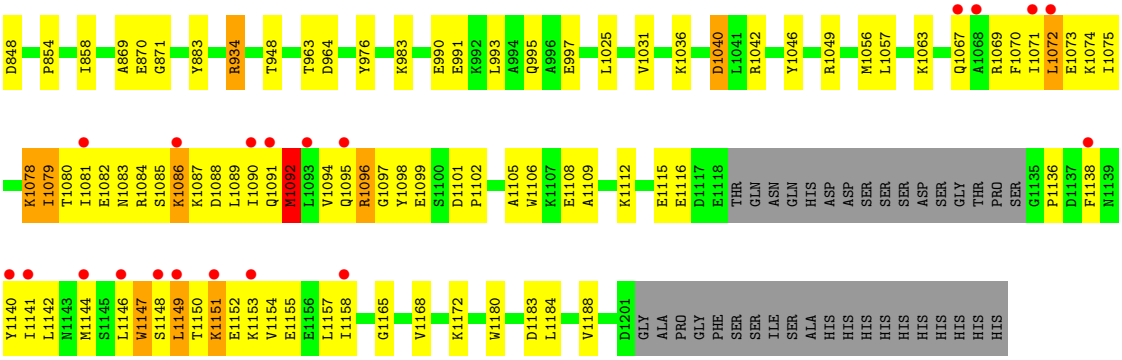
Chain A:



• Molecule 1: DNA topoisomerase 2-beta

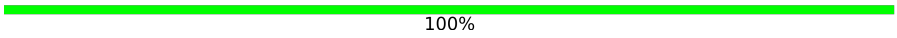
Chain B:





• Molecule 2: DNA (5'-D(P*AP*GP*CP*CP*GP*AP*GP*C)-3')

Chain C:



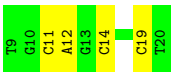
There are no outlier residues recorded for this chain.

• Molecule 2: DNA (5'-D(P*AP*GP*CP*CP*GP*AP*GP*C)-3')

Chain D:

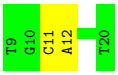
• Molecule 3: DNA (5'-D(P*TP*GP*CP*AP*GP*CP*TP*CP*GP*GP*CP*T)-3')

Chain F:



• Molecule 3: DNA (5'-D(P*TP*GP*CP*AP*GP*CP*TP*CP*GP*GP*CP*T)-3')

Chain E:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, α, β, γ	95.01Å 95.01Å 230.29Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	28.06 – 2.54 47.50 – 2.20	Depositor EDS
% Data completeness (in resolution range)	99.8 (28.06-2.54) 99.5 (47.50-2.20)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	0.08	Depositor
$< I/\sigma(I) >$ [1]	1.13 (at 2.20Å)	Xtriage
Refinement program	PHENIX (1.13_2998: ???)	Depositor
R, R_{free}	0.180 , 0.227 0.180 , 0.227	Depositor DCC
R_{free} test set	6114 reflections (5.18%)	wwPDB-VP
Wilson B-factor (Å^2)	55.5	Xtriage
Anisotropy	0.236	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 51.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.021 for -h,-k,l 0.487 for h,-h-k,-l 0.022 for -k,-h,-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	12706	wwPDB-VP
Average B, all atoms (Å^2)	74.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.81% 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	0.47	1/6030 (0.0%)	0.66	6/8124 (0.1%)
1	B	0.48	1/6014 (0.0%)	0.76	15/8103 (0.2%)
2	C	1.06	0/185	1.07	0/283
2	D	1.02	0/185	1.01	0/283
3	E	1.11	0/273	1.09	0/419
3	F	1.12	0/273	1.22	3/419 (0.7%)
All	All	0.54	2/12960 (0.0%)	0.75	24/17631 (0.1%)

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	B	1	0

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	698	GLU	CB-CG	5.46	1.62	1.52
1	A	614	TRP	CB-CG	-5.05	1.41	1.50

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	1149	LEU	CB-CG-CD2	-20.67	75.86	111.00
1	B	1149	LEU	CB-CG-CD1	13.49	133.94	111.00
1	B	692	ARG	NE-CZ-NH1	10.41	125.51	120.30
1	B	1086	LYS	CD-CE-NZ	-9.95	88.81	111.70
1	A	993	LEU	CB-CG-CD2	8.93	126.17	111.00

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	B	1079	ILE	CB

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	5903	0	5923	150	0
1	B	5893	0	5909	138	0
2	C	165	0	89	0	0
2	D	165	0	89	2	0
3	E	245	0	136	2	0
3	F	245	0	136	3	0
4	A	39	0	0	1	0
4	B	31	0	0	1	0
4	C	5	0	0	0	0
4	D	5	0	0	0	0
4	E	5	0	0	0	0
4	F	5	0	0	0	0
All	All	12706	0	12282	278	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 11.

The worst 5 of 278 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:691:ARG:NH2	1:B:698:GLU:OE1	1.85	1.08
1:A:1089:LEU:HA	1:A:1092:MET:HB3	1.43	0.96
1:A:617:HIS:HE1	1:A:624:TRP:CH2	1.84	0.94
1:B:732:PRO:HG2	1:B:869:ALA:HB1	1.50	0.94
1:A:617:HIS:CE1	1:A:624:TRP:CZ3	2.58	0.91

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	726/803 (90%)	687 (95%)	38 (5%)	1 (0%)	51 65
1	B	724/803 (90%)	684 (94%)	38 (5%)	2 (0%)	41 51
All	All	1450/1606 (90%)	1371 (95%)	76 (5%)	3 (0%)	47 60

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	616	LYS
1	B	616	LYS
1	B	455	PRO

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	642/704 (91%)	627 (98%)	15 (2%)	50 65
1	B	639/704 (91%)	621 (97%)	18 (3%)	43 58
All	All	1281/1408 (91%)	1248 (97%)	33 (3%)	46 61

5 of 33 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	1092	MET
1	B	1096	ARG
1	B	1172	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	1092	MET
1	A	1078	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	617	HIS
1	A	1110	GLN
1	B	617	HIS
1	B	1083	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

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	729/803 (90%)	-0.23	17 (2%) 60 67	41, 65, 131, 173	6 (0%)
1	B	729/803 (90%)	-0.23	26 (3%) 42 49	41, 66, 131, 172	4 (0%)
2	C	8/8 (100%)	-0.69	0 100 100	47, 49, 74, 107	0
2	D	8/8 (100%)	-0.65	0 100 100	47, 50, 73, 106	0
3	E	12/12 (100%)	-0.58	0 100 100	49, 59, 106, 146	0
3	F	12/12 (100%)	-0.67	0 100 100	48, 59, 111, 136	0
All	All	1498/1646 (91%)	-0.24	43 (2%) 51 59	41, 65, 132, 173	10 (0%)

The worst 5 of 43 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1090	ILE	7.1
1	A	1141	ILE	6.5
1	A	1146	LEU	5.2
1	B	1081	ILE	5.2
1	B	1086	LYS	5.0

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.