wwPDB X-ray Structure Validation Summary Report (i)

Oct 25, 2023 – 12:11 AM EDT

PDB ID	:	2Z5C
Title	:	Crystal Structure of a Novel Chaperone Complex for Yeast 20S Proteasome Assembly
Authors	:	Yashiroda, H.; Mizushima, T.; Okamoto, K.; Kameyama, T.; Hayashi, H.; Kishimoto, T.; Kasahara, M.; Kurimoto, E.; Sakata, E.; Suzuki, A.; Hirano, Y.; Murata, S.; Kato, K.; Yamane, T.; Tanaka, K.
Deposited on	:	2007-07-03
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 (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 (1)) were used in the production of this report:

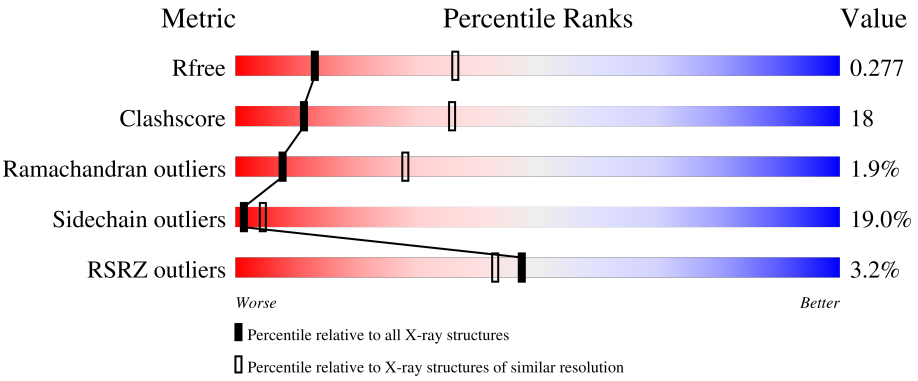
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation 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 2.90 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.



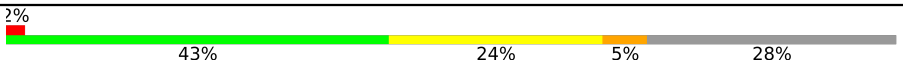
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (2.90-2.90)

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 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	151	3% 44% 28% 8% • 18%
1	D	151	2% 44% 27% 9% • 18%
2	B	149	3% 36% 30% 5% 29%
2	E	149	3% 40% 26% 5% 29%
3	C	262	2% 42% 25% 5% 28%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	F	262	

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 6598 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 Protein YPL144W.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	124	Total	C	N	O	S	0	0	0
			980	623	166	184	7			
1	D	124	Total	C	N	O	S	0	0	0
			980	623	166	184	7			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	GLY	-	expression tag	UNP Q12245
A	-1	SER	-	expression tag	UNP Q12245
A	0	HIS	-	expression tag	UNP Q12245
D	-2	GLY	-	expression tag	UNP Q12245
D	-1	SER	-	expression tag	UNP Q12245
D	0	HIS	-	expression tag	UNP Q12245

- Molecule 2 is a protein called Uncharacterized protein YLR021W.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	106	Total	C	N	O	S	0	0	0
			856	557	138	156	5			
2	E	106	Total	C	N	O	S	0	0	0
			856	557	138	156	5			

- Molecule 3 is a protein called Proteasome component PUP2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	189	Total	C	N	O	S	0	0	0
			1463	915	247	295	6			
3	F	189	Total	C	N	O	S	0	0	0
			1463	915	247	295	6			

There are 4 discrepancies between the modelled and reference sequences:

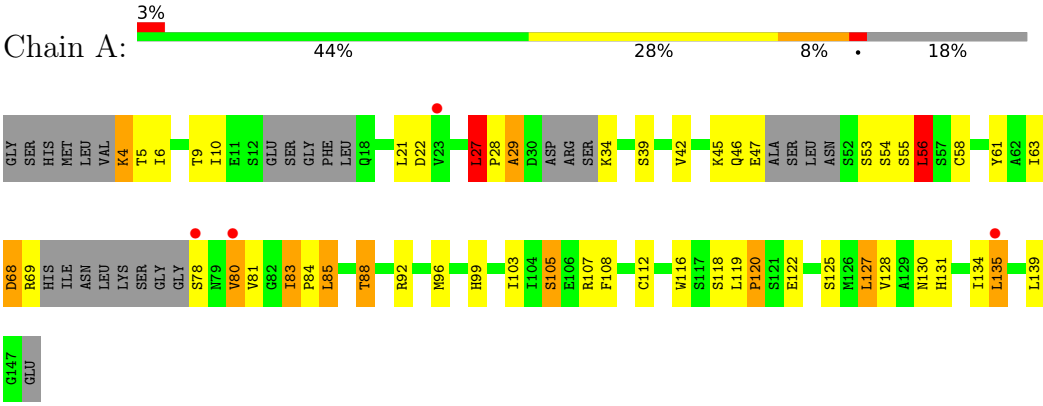
Chain	Residue	Modelled	Actual	Comment	Reference
C	-1	GLY	-	expression tag	UNP P32379
C	0	SER	-	expression tag	UNP P32379
F	-1	GLY	-	expression tag	UNP P32379
F	0	SER	-	expression tag	UNP P32379

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.

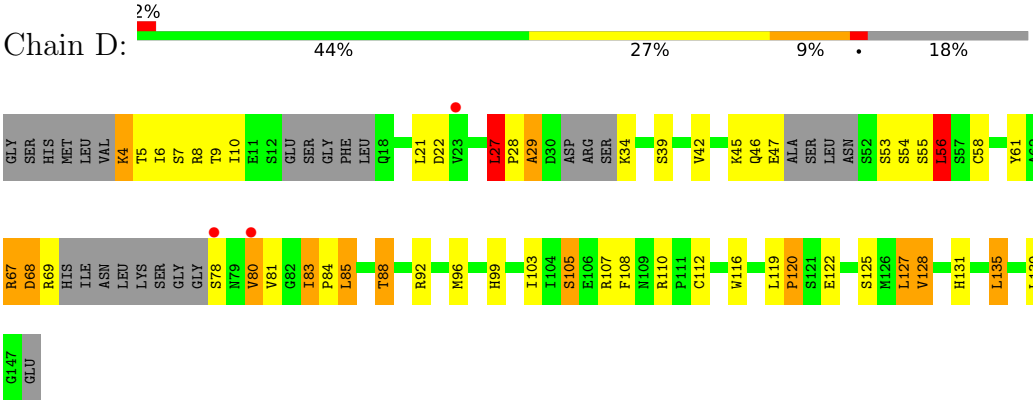
• Molecule 1: Protein YPL144W

Chain A:



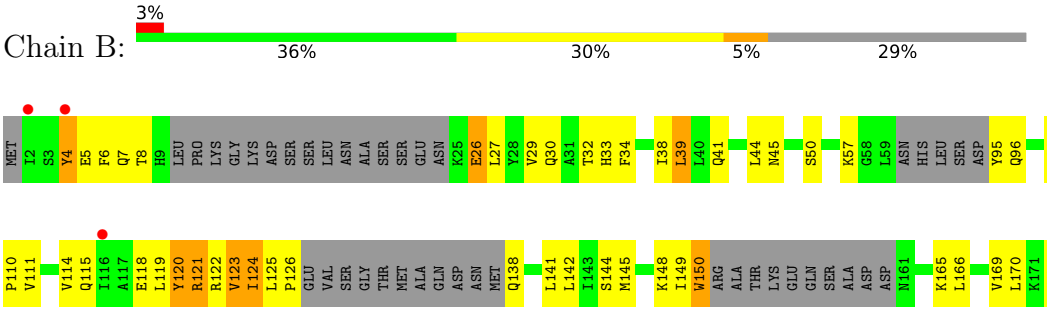
• Molecule 1: Protein YPL144W

Chain D:



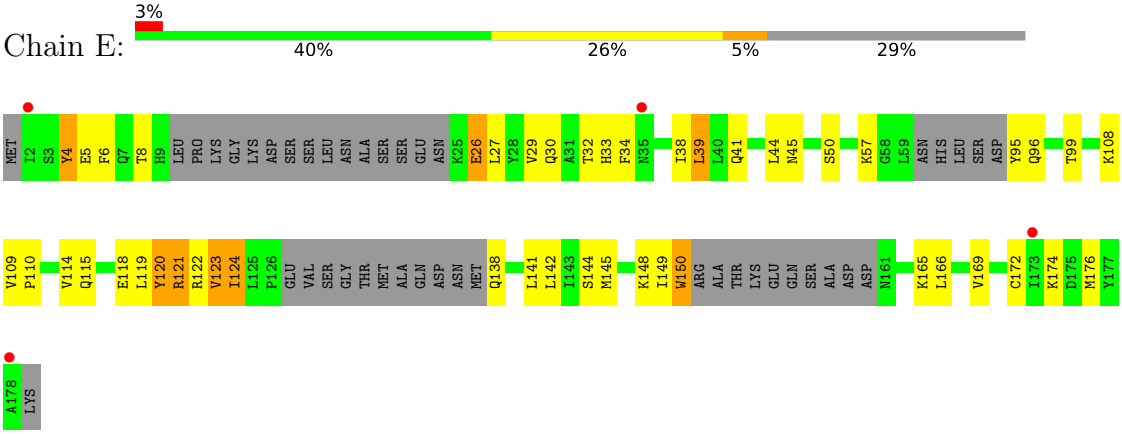
• Molecule 2: Uncharacterized protein YLR021W

Chain B:



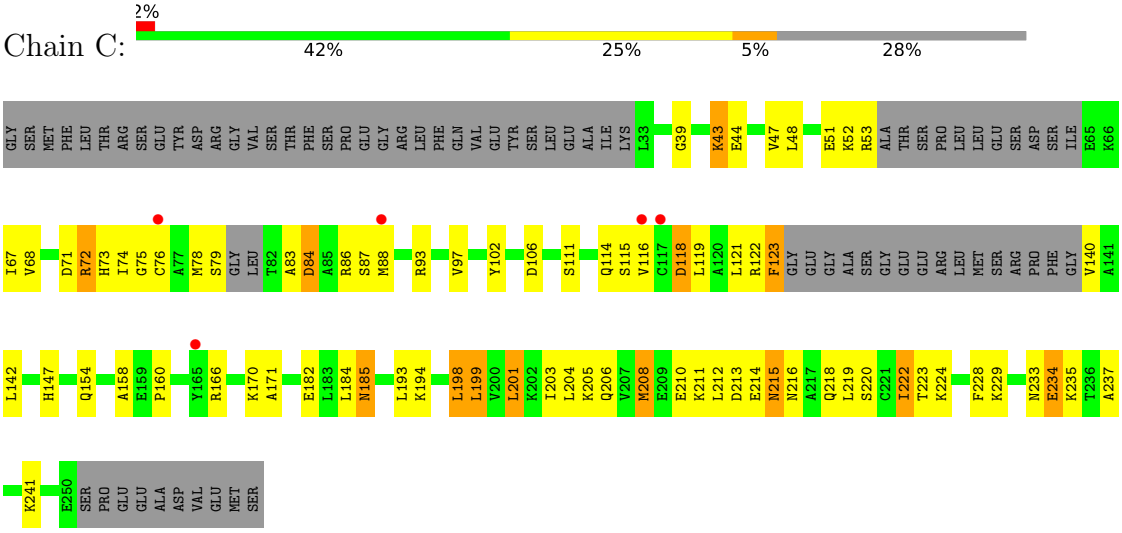
• Molecule 2: Uncharacterized protein YLR021W

Chain E:



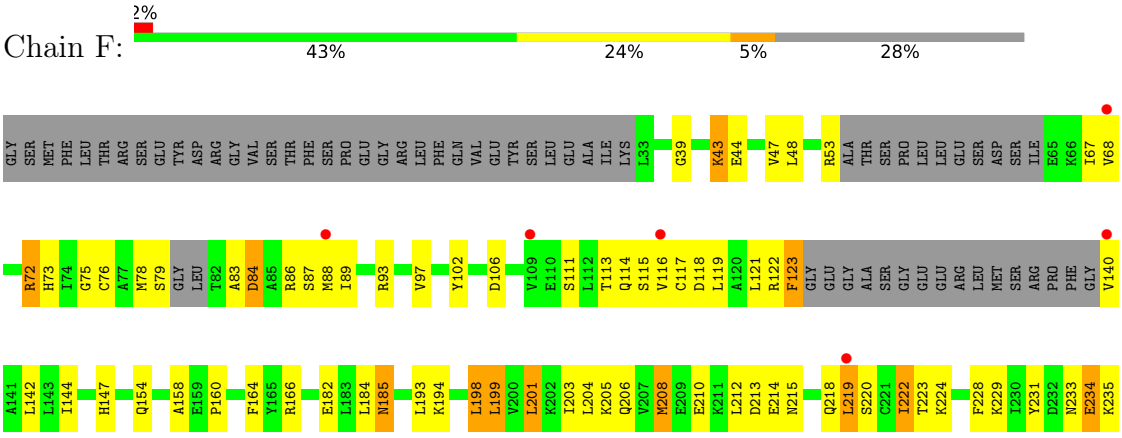
• Molecule 3: Proteasome component PUP2

Chain C:



• Molecule 3: Proteasome component PUP2

Chain F:




T236
A237
K241
E250
SER
PRO
GLU
GLU
ALA
ASP
VAL
GLU
MET
SER

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	158.69Å 159.18Å 65.03Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.38 – 2.90 50.36 – 2.90	Depositor EDS
% Data completeness (in resolution range)	99.4 (50.38-2.90) 98.8 (50.36-2.90)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	5.81 (at 2.91Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.250 , 0.283 0.245 , 0.277	Depositor DCC
R_{free} test set	1859 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	66.1	Xtriage
Anisotropy	0.673	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 41.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	0.458 for k,h,-l	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	6598	wwPDB-VP
Average B, all atoms (Å^2)	66.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.72% 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	0.80	0/994	1.06	6/1341 (0.4%)
1	D	0.79	0/994	1.07	6/1341 (0.4%)
2	B	0.59	0/869	0.70	0/1169
2	E	0.60	0/869	0.70	0/1169
3	C	0.70	0/1478	0.80	0/1987
3	F	0.70	0/1478	0.81	0/1987
All	All	0.71	0/6682	0.87	12/8994 (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	A	0	1
1	D	0	1
All	All	0	2

There are no bond length outliers.

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	67	ARG	NE-CZ-NH1	11.95	126.27	120.30
1	A	67	ARG	NE-CZ-NH2	9.56	125.08	120.30
1	A	67	ARG	NE-CZ-NH1	-8.47	116.06	120.30
1	D	67	ARG	NE-CZ-NH2	-7.27	116.66	120.30
1	A	67	ARG	CD-NE-CZ	6.27	132.38	123.60

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	53	SER	Peptide
1	D	53	SER	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	980	0	1011	35	0
1	D	980	0	1011	38	1
2	B	856	0	872	36	0
2	E	856	0	872	33	0
3	C	1463	0	1448	50	1
3	F	1463	0	1448	50	0
All	All	6598	0	6662	236	1

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 18.

The worst 5 of 236 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:D:68:ASP:OD1	1:D:68:ASP:O	1.72	1.08
1:A:68:ASP:OD1	1:A:68:ASP:O	1.74	1.05
3:F:154:GLN:HE21	3:F:166:ARG:HH22	0.99	0.99
3:C:154:GLN:HE21	3:C:166:ARG:HH22	1.05	0.98
3:C:182:GLU:HG2	3:C:206:GLN:HE22	1.31	0.94

All (1) 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)
3:C:106:ASP:OD2	1:D:67:ARG:NH2[1_554]	2.06	0.14

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	114/151 (76%)	100 (88%)	11 (10%)	3 (3%)	5 20
1	D	114/151 (76%)	100 (88%)	11 (10%)	3 (3%)	5 20
2	B	96/149 (64%)	83 (86%)	10 (10%)	3 (3%)	4 16
2	E	96/149 (64%)	80 (83%)	13 (14%)	3 (3%)	4 16
3	C	181/262 (69%)	167 (92%)	12 (7%)	2 (1%)	14 42
3	F	181/262 (69%)	166 (92%)	14 (8%)	1 (1%)	25 58
All	All	782/1124 (70%)	696 (89%)	71 (9%)	15 (2%)	8 28

5 of 15 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	121	ARG
2	B	123	VAL
2	E	121	ARG
2	E	123	VAL
1	A	29	ALA

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	115/137 (84%)	85 (74%)	30 (26%)	0 1
1	D	115/137 (84%)	86 (75%)	29 (25%)	0 2
2	B	98/135 (73%)	85 (87%)	13 (13%)	4 11

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
2	E	98/135 (73%)	86 (88%)	12 (12%)	5 15
3	C	155/216 (72%)	126 (81%)	29 (19%)	1 5
3	F	155/216 (72%)	128 (83%)	27 (17%)	2 6
All	All	736/976 (75%)	596 (81%)	140 (19%)	1 4

5 of 140 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	F	53	ARG
3	F	79	SER
3	F	198	LEU
3	C	76	CYS
3	C	72	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 15 such sidechains are listed below:

Mol	Chain	Res	Type
2	E	7	GLN
3	F	206	GLN
2	E	45	ASN
3	F	218	GLN
3	F	154	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

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	124/151 (82%)	0.41	4 (3%)	47	43	39, 59, 74, 84	0
1	D	124/151 (82%)	0.38	3 (2%)	59	56	39, 59, 74, 84	0
2	B	106/149 (71%)	0.60	5 (4%)	31	28	53, 73, 99, 102	0
2	E	106/149 (71%)	0.64	4 (3%)	40	36	53, 73, 99, 102	0
3	C	189/262 (72%)	0.41	5 (2%)	56	52	39, 61, 97, 109	0
3	F	189/262 (72%)	0.40	6 (3%)	47	43	39, 61, 97, 109	0
All	All	838/1124 (74%)	0.46	27 (3%)	47	43	39, 64, 97, 109	0

The worst 5 of 27 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	E	178	ALA	6.3
1	A	78	SER	5.7
2	B	178	ALA	5.1
1	D	78	SER	3.3
3	C	88	MET	2.7

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.