wwPDB X-ray Structure Validation Summary Report (i)

May 21, 2024 – 02:04 PM EDT

PDB ID	:	8GCJ
Title	:	PCNA
Authors	:	Vandborg, B.C.; Bruning, J.B.
Deposited on	:	2023-03-01
Resolution	:	2.85 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:

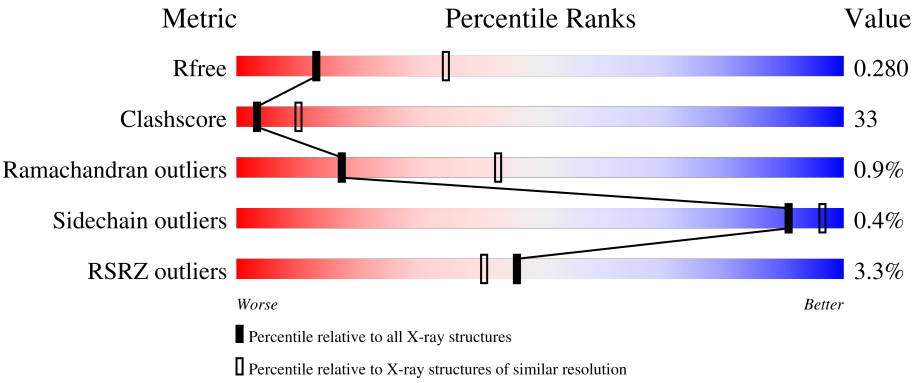
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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 2.85 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	3168 (2.90-2.82)
Clashscore	141614	3438 (2.90-2.82)
Ramachandran outliers	138981	3348 (2.90-2.82)
Sidechain outliers	138945	3351 (2.90-2.82)
RSRZ outliers	127900	3103 (2.90-2.82)

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	261	2% 50% 44%
1	B	261	3% 51% 44%
1	C	261	5% 59% 35% 5%
1	D	261	6% 55% 40% 5%
1	E	261	% 52% 45%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	261	

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 11280 atoms, of which 29 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 Proliferating cell nuclear antigen.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	250	Total	C	H	N	O	S	0	1	0
			1851	1167	1	307	361	15			
1	C	249	Total	C	H	N	O	S	0	0	0
			1829	1151	10	299	354	15			
1	E	256	Total	C	H	N	O	S	0	0	0
			1880	1188	3	312	361	16			
1	F	257	Total	C	H	N	O	S	0	0	0
			1920	1205	4	315	380	16			
1	B	253	Total	C	H	N	O	S	0	0	0
			1868	1175	7	302	368	16			
1	D	249	Total	C	H	N	O	S	0	0	0
			1809	1140	4	302	348	15			

- Molecule 2 is water.

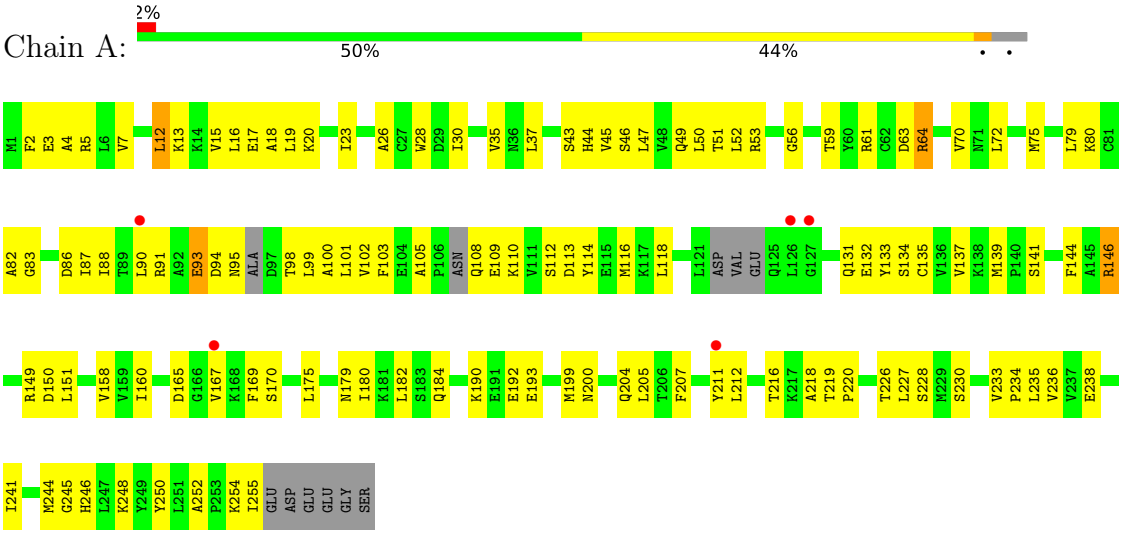
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	21	Total	O	0	0
			21	21		
2	C	36	Total	O	0	0
			36	36		
2	E	62	Total	O	0	0
			62	62		
2	F	1	Total	O	0	0
			1	1		
2	D	3	Total	O	0	0
			3	3		

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.

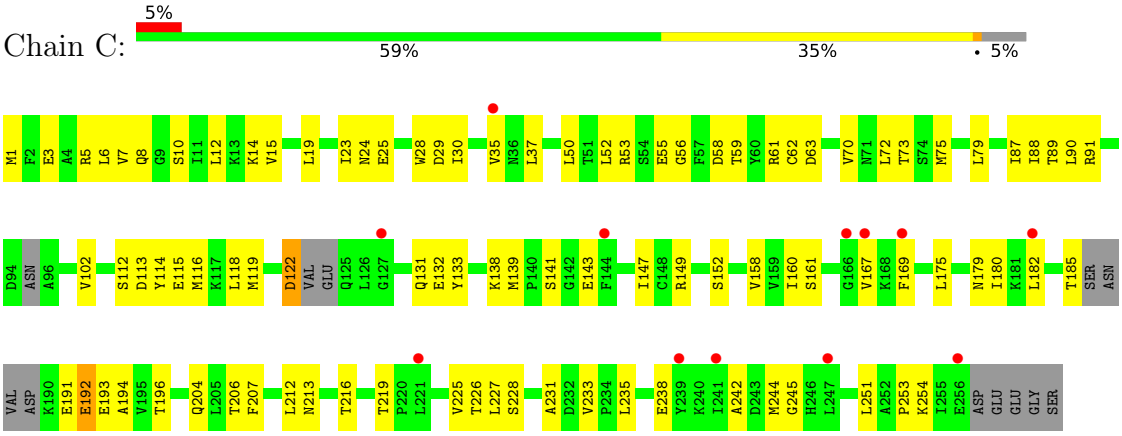
• Molecule 1: Proliferating cell nuclear antigen

Chain A:



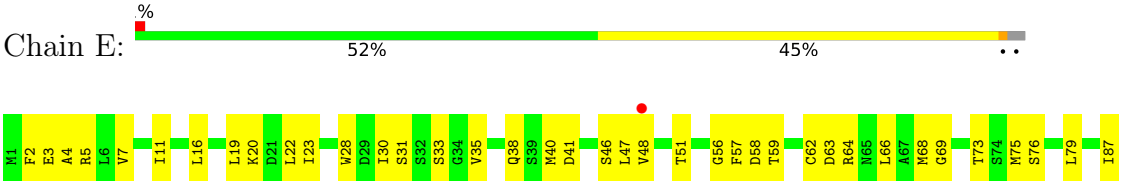
• Molecule 1: Proliferating cell nuclear antigen

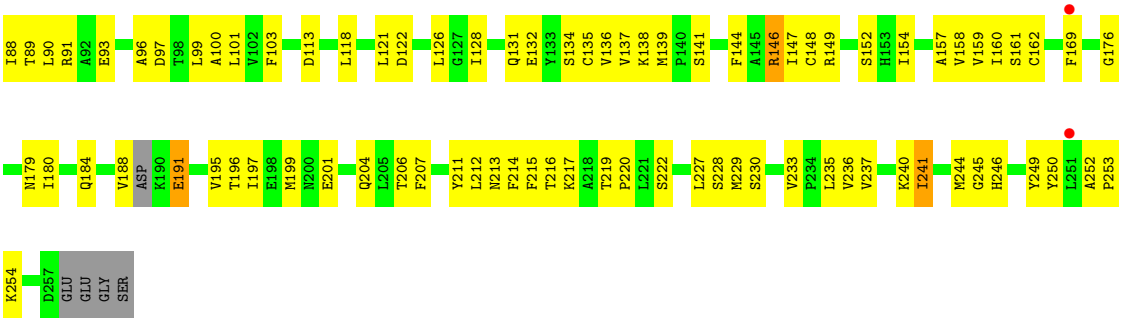
Chain C:



• Molecule 1: Proliferating cell nuclear antigen

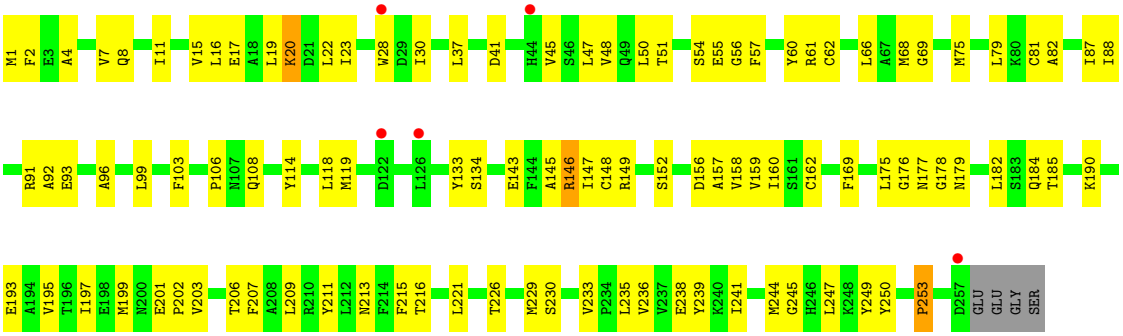
Chain E:





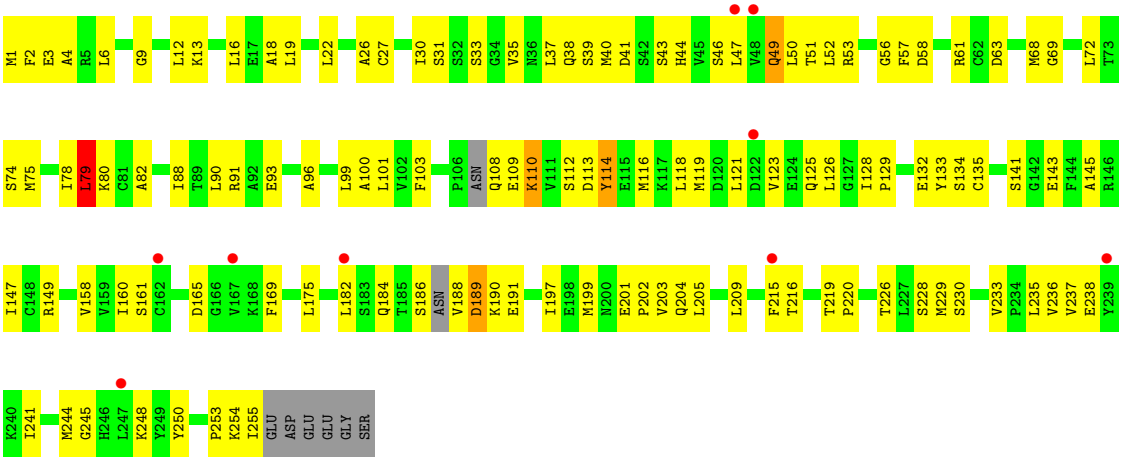
• Molecule 1: Proliferating cell nuclear antigen

Chain F:



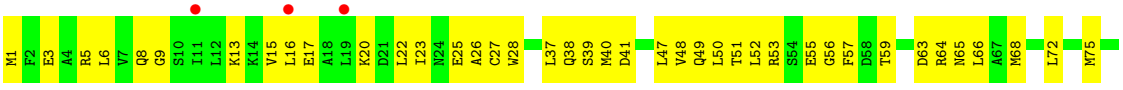
• Molecule 1: Proliferating cell nuclear antigen

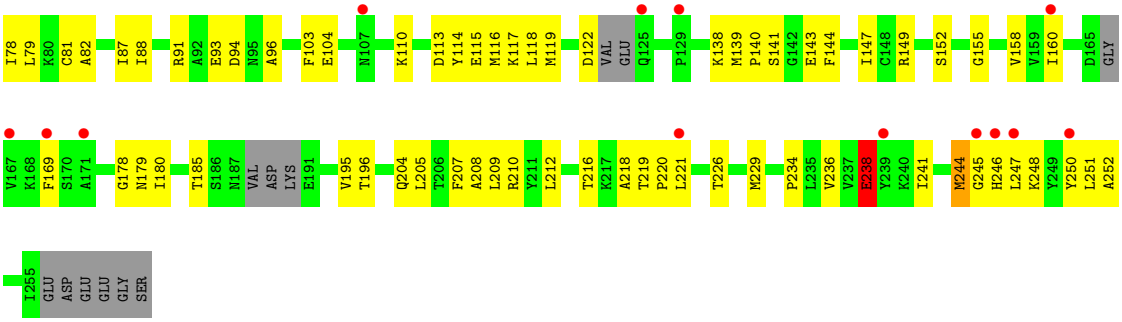
Chain B:



• Molecule 1: Proliferating cell nuclear antigen

Chain D:




I78 L79 K80 C81 A82 I87 I88 R91 A92 E93 D94 N95 A96 F103 E104 N107 K110 D113 Y114 E115 M116 K117 L118 M119 D122 VAL GLU Q125 P129 K138 M139 P140 S141 G142 E143 F144 I147 C148 R149 S152 G155 V158 V159 I160 D165 GLY
V167 K168 F169 S170 A171 G178 N179 I180 T185 S186 N187 VAL ASP LYS E191 V195 T196 Q204 L205 T206 F207 A208 L209 R210 Y211 L212 T216 K217 A218 T219 P220 L221 T226 M229 P234 L235 V236 V237 E238 Y239 K240 I241 M244 G245 H246 L247 K248 Y249 Y250 L251 A252
I255 GLU ASP GLU GLU GLY SER

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	147.10Å 85.19Å 149.81Å 90.00° 116.87° 90.00°	Depositor
Resolution (Å)	48.39 – 2.85 48.39 – 2.85	Depositor EDS
% Data completeness (in resolution range)	97.8 (48.39-2.85) 97.8 (48.39-2.85)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	5.25 (at 2.86Å)	Xtriage
Refinement program	PHENIX 1.18.2_3874	Depositor
R, R_{free}	0.231 , 0.278 0.232 , 0.280	Depositor DCC
R_{free} test set	1989 reflections (5.24%)	wwPDB-VP
Wilson B-factor (Å^2)	60.3	Xtriage
Anisotropy	0.945	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 64.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.000 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	11280	wwPDB-VP
Average B, all atoms (Å^2)	73.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 46.70 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.1023e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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	0.44	2/1872 (0.1%)	0.70	3/2535 (0.1%)
1	B	0.35	1/1884 (0.1%)	0.65	2/2554 (0.1%)
1	C	0.37	0/1842	0.74	2/2494 (0.1%)
1	D	0.38	2/1827 (0.1%)	0.66	0/2476
1	E	0.29	0/1902	0.61	1/2580 (0.0%)
1	F	0.31	0/1942	0.64	2/2630 (0.1%)
All	All	0.36	5/11269 (0.0%)	0.67	10/15269 (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

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	64	ARG	CZ-NH2	7.17	1.42	1.33
1	D	238	GLU	CD-OE2	6.23	1.32	1.25
1	A	192	GLU	CD-OE2	6.07	1.32	1.25
1	B	110	LYS	CE-NZ	5.97	1.64	1.49
1	D	238	GLU	CD-OE1	5.36	1.31	1.25

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	122	ASP	CB-CG-OD1	-11.72	107.75	118.30
1	F	146	ARG	NE-CZ-NH2	-8.09	116.25	120.30
1	E	146	ARG	NE-CZ-NH1	-6.40	117.10	120.30
1	A	64	ARG	CA-CB-CG	6.05	126.72	113.40
1	A	192	GLU	OE1-CD-OE2	5.99	130.49	123.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	95	ASN	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	1850	1	1790	136	1
1	B	1861	7	1782	133	0
1	C	1819	10	1747	112	1
1	D	1805	4	1719	114	0
1	E	1877	3	1829	146	2
1	F	1916	4	1857	101	0
2	A	21	0	0	12	0
2	C	36	0	0	21	0
2	D	3	0	0	3	0
2	E	62	0	0	71	1
2	F	1	0	0	1	0
All	All	11251	29	10724	713	3

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 33.

The worst 5 of 713 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:C:75:MET:HG3	1:C:116:MET:HE1	1.25	1.15
1:E:149:ARG:N	2:E:302:HOH:O	1.82	1.13
1:D:1:MET:N	1:D:94:ASP:OD1	1.83	1.12
1:D:238:GLU:OE1	1:D:248:LYS:HE2	1.55	1.07
1:B:114:TYR:HD2	1:D:178:GLY:HA3	1.00	1.06

All (3) 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:64:ARG:NH2	2:E:310:HOH:O[3_455]	2.10	0.10
1:C:62:CYS:O	1:E:64:ARG:NH2[1_565]	2.18	0.02
1:E:5:ARG:NH2	1:E:58:ASP:OD2[2_656]	2.18	0.02

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	245/261 (94%)	224 (91%)	20 (8%)	1 (0%)	34	62
1	B	247/261 (95%)	217 (88%)	27 (11%)	3 (1%)	13	35
1	C	241/261 (92%)	220 (91%)	18 (8%)	3 (1%)	13	35
1	D	241/261 (92%)	224 (93%)	16 (7%)	1 (0%)	34	62
1	E	252/261 (97%)	235 (93%)	15 (6%)	2 (1%)	19	46
1	F	255/261 (98%)	237 (93%)	15 (6%)	3 (1%)	13	35
All	All	1481/1566 (95%)	1357 (92%)	111 (8%)	13 (1%)	17	43

5 of 13 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	192	GLU
1	C	242	ALA
1	E	241	ILE
1	F	106	PRO
1	B	190	LYS

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	193/228 (85%)	191 (99%)	2 (1%)	76 91
1	B	193/228 (85%)	192 (100%)	1 (0%)	88 96
1	C	186/228 (82%)	185 (100%)	1 (0%)	88 96
1	D	182/228 (80%)	180 (99%)	2 (1%)	73 90
1	E	196/228 (86%)	196 (100%)	0	100 100
1	F	204/228 (90%)	204 (100%)	0	100 100
All	All	1154/1368 (84%)	1148 (100%)	6 (0%)	91 96

5 of 6 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	114	TYR
1	D	210	ARG
1	D	238	GLU
1	A	146[B]	ARG
1	A	146[A]	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	184	GLN
1	D	246	HIS
1	E	131	GLN
1	F	179	ASN
1	F	184	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)

There are no ligands in this entry.

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

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	250/261 (95%)	0.02	5 (2%)	65	62	45, 74, 98, 122	0
1	B	253/261 (96%)	0.08	9 (3%)	42	37	54, 80, 106, 127	0
1	C	249/261 (95%)	0.05	12 (4%)	30	26	46, 70, 102, 113	0
1	D	249/261 (95%)	0.06	16 (6%)	19	15	58, 79, 107, 117	0
1	E	256/261 (98%)	-0.10	3 (1%)	79	78	30, 57, 96, 114	0
1	F	257/261 (98%)	0.01	5 (1%)	66	64	46, 65, 101, 120	0
All	All	1514/1566 (96%)	0.02	50 (3%)	46	41	30, 72, 104, 127	0

The worst 5 of 50 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	166	GLY	6.0
1	C	221	LEU	5.9
1	F	122	ASP	4.8
1	B	162	CYS	4.5
1	D	221	LEU	3.9

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.