WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 11, 2024 – 03:45 PM EST

PDB ID	:	3BQ4
Title	:	Crystal Structure of Ad35 fiber knob
Authors	:	Pache, L.; Venkataraman, S.; Nemerow, G.R.; Reddy, V.S.
Deposited on	:	2007-12-19
Resolution	:	2.70 Å(reported)

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

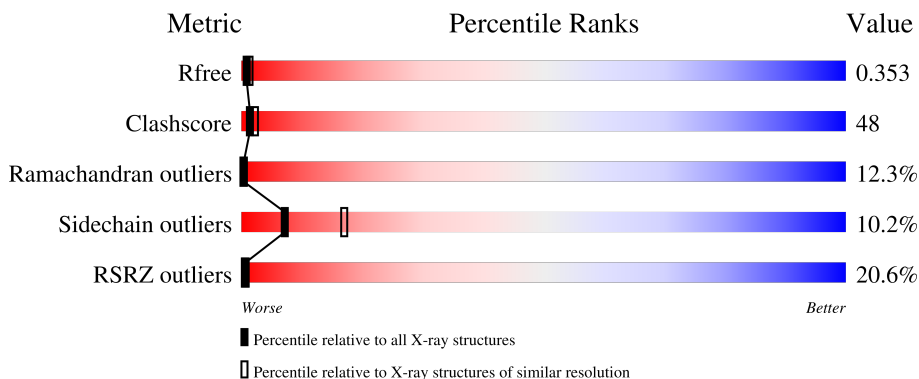|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation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

1 Overall quality at a glance ⓘ

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

The reported resolution of this entry is 2.70 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.
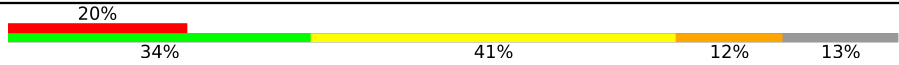
Metric	Percentile Ranks	Value
Rfree		0.353
Clashscore		48
Ramachandran outliers		12.3%
Sidechain outliers		10.2%
RSRZ outliers		20.6%

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	130704	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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	220	18% / 35% · 39% · 13% · 13%
1	B	220	19% / 27% · 45% · 13% · 13%
1	D	220	18% / 29% · 44% · 14% · 13%
1	E	220	17% / 32% · 40% · 14% · 13%
1	F	220	17% / 34% · 39% · 12% · 13%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	220	

2 Entry composition ⓘ

There is only 1 type of molecule in this entry. The entry contains 8976 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 Fiber.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	192	Total	C	N	O	S	0	0	0
			1496	944	240	305	7			
1	B	192	Total	C	N	O	S	0	0	0
			1496	944	240	305	7			
1	D	192	Total	C	N	O	S	0	0	0
			1496	944	240	305	7			
1	E	192	Total	C	N	O	S	0	0	0
			1496	944	240	305	7			
1	F	192	Total	C	N	O	S	0	0	0
			1496	944	240	305	7			
1	G	192	Total	C	N	O	S	0	0	0
			1496	944	240	305	7			

There are 102 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	104	GLY	-	expression tag	UNP Q7T925
A	105	SER	-	expression tag	UNP Q7T925
A	106	HIS	-	expression tag	UNP Q7T925
A	107	MET	-	expression tag	UNP Q7T925
A	108	ALA	-	expression tag	UNP Q7T925
A	109	SER	-	expression tag	UNP Q7T925
A	110	MET	-	expression tag	UNP Q7T925
A	111	THR	-	expression tag	UNP Q7T925
A	112	GLY	-	expression tag	UNP Q7T925
A	113	GLY	-	expression tag	UNP Q7T925
A	114	GLN	-	expression tag	UNP Q7T925
A	115	GLN	-	expression tag	UNP Q7T925
A	116	MET	-	expression tag	UNP Q7T925
A	117	GLY	-	expression tag	UNP Q7T925
A	118	ARG	-	expression tag	UNP Q7T925
A	119	GLY	-	expression tag	UNP Q7T925
A	120	SER	-	expression tag	UNP Q7T925

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	104	GLY	-	expression tag	UNP Q7T925
B	105	SER	-	expression tag	UNP Q7T925
B	106	HIS	-	expression tag	UNP Q7T925
B	107	MET	-	expression tag	UNP Q7T925
B	108	ALA	-	expression tag	UNP Q7T925
B	109	SER	-	expression tag	UNP Q7T925
B	110	MET	-	expression tag	UNP Q7T925
B	111	THR	-	expression tag	UNP Q7T925
B	112	GLY	-	expression tag	UNP Q7T925
B	113	GLY	-	expression tag	UNP Q7T925
B	114	GLN	-	expression tag	UNP Q7T925
B	115	GLN	-	expression tag	UNP Q7T925
B	116	MET	-	expression tag	UNP Q7T925
B	117	GLY	-	expression tag	UNP Q7T925
B	118	ARG	-	expression tag	UNP Q7T925
B	119	GLY	-	expression tag	UNP Q7T925
B	120	SER	-	expression tag	UNP Q7T925
D	104	GLY	-	expression tag	UNP Q7T925
D	105	SER	-	expression tag	UNP Q7T925
D	106	HIS	-	expression tag	UNP Q7T925
D	107	MET	-	expression tag	UNP Q7T925
D	108	ALA	-	expression tag	UNP Q7T925
D	109	SER	-	expression tag	UNP Q7T925
D	110	MET	-	expression tag	UNP Q7T925
D	111	THR	-	expression tag	UNP Q7T925
D	112	GLY	-	expression tag	UNP Q7T925
D	113	GLY	-	expression tag	UNP Q7T925
D	114	GLN	-	expression tag	UNP Q7T925
D	115	GLN	-	expression tag	UNP Q7T925
D	116	MET	-	expression tag	UNP Q7T925
D	117	GLY	-	expression tag	UNP Q7T925
D	118	ARG	-	expression tag	UNP Q7T925
D	119	GLY	-	expression tag	UNP Q7T925
D	120	SER	-	expression tag	UNP Q7T925
E	104	GLY	-	expression tag	UNP Q7T925
E	105	SER	-	expression tag	UNP Q7T925
E	106	HIS	-	expression tag	UNP Q7T925
E	107	MET	-	expression tag	UNP Q7T925
E	108	ALA	-	expression tag	UNP Q7T925
E	109	SER	-	expression tag	UNP Q7T925
E	110	MET	-	expression tag	UNP Q7T925
E	111	THR	-	expression tag	UNP Q7T925

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	112	GLY	-	expression tag	UNP Q7T925
E	113	GLY	-	expression tag	UNP Q7T925
E	114	GLN	-	expression tag	UNP Q7T925
E	115	GLN	-	expression tag	UNP Q7T925
E	116	MET	-	expression tag	UNP Q7T925
E	117	GLY	-	expression tag	UNP Q7T925
E	118	ARG	-	expression tag	UNP Q7T925
E	119	GLY	-	expression tag	UNP Q7T925
E	120	SER	-	expression tag	UNP Q7T925
F	104	GLY	-	expression tag	UNP Q7T925
F	105	SER	-	expression tag	UNP Q7T925
F	106	HIS	-	expression tag	UNP Q7T925
F	107	MET	-	expression tag	UNP Q7T925
F	108	ALA	-	expression tag	UNP Q7T925
F	109	SER	-	expression tag	UNP Q7T925
F	110	MET	-	expression tag	UNP Q7T925
F	111	THR	-	expression tag	UNP Q7T925
F	112	GLY	-	expression tag	UNP Q7T925
F	113	GLY	-	expression tag	UNP Q7T925
F	114	GLN	-	expression tag	UNP Q7T925
F	115	GLN	-	expression tag	UNP Q7T925
F	116	MET	-	expression tag	UNP Q7T925
F	117	GLY	-	expression tag	UNP Q7T925
F	118	ARG	-	expression tag	UNP Q7T925
F	119	GLY	-	expression tag	UNP Q7T925
F	120	SER	-	expression tag	UNP Q7T925
G	104	GLY	-	expression tag	UNP Q7T925
G	105	SER	-	expression tag	UNP Q7T925
G	106	HIS	-	expression tag	UNP Q7T925
G	107	MET	-	expression tag	UNP Q7T925
G	108	ALA	-	expression tag	UNP Q7T925
G	109	SER	-	expression tag	UNP Q7T925
G	110	MET	-	expression tag	UNP Q7T925
G	111	THR	-	expression tag	UNP Q7T925
G	112	GLY	-	expression tag	UNP Q7T925
G	113	GLY	-	expression tag	UNP Q7T925
G	114	GLN	-	expression tag	UNP Q7T925
G	115	GLN	-	expression tag	UNP Q7T925
G	116	MET	-	expression tag	UNP Q7T925
G	117	GLY	-	expression tag	UNP Q7T925
G	118	ARG	-	expression tag	UNP Q7T925
G	119	GLY	-	expression tag	UNP Q7T925

Continued from previous page...

| Chain	Residue	Modelled	Actual	Comment	Reference
G	120	SER	-	expression tag	UNP Q7T925

3 Residue-property plots ⓘ

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.
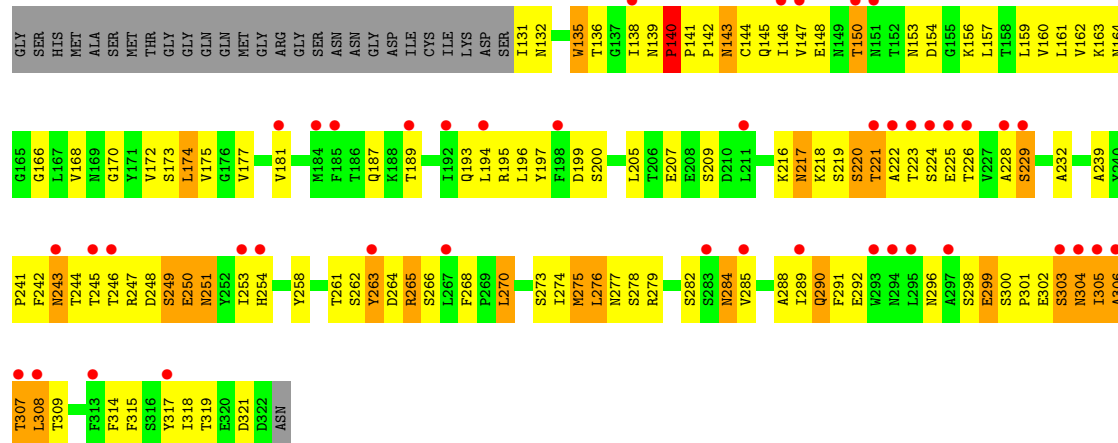
• Molecule 1: Fiber

Chain A:

• Molecule 1: Fiber

Chain B:

• Molecule 1: Fiber

Chain D:

- Molecule 1: Fiber
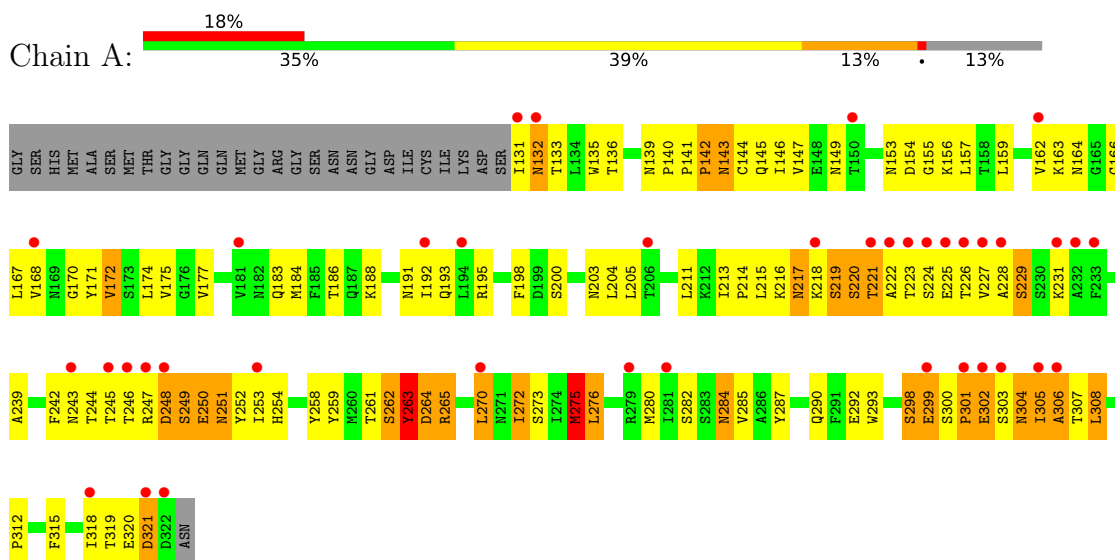
Chain E:

- Molecule 1: Fiber
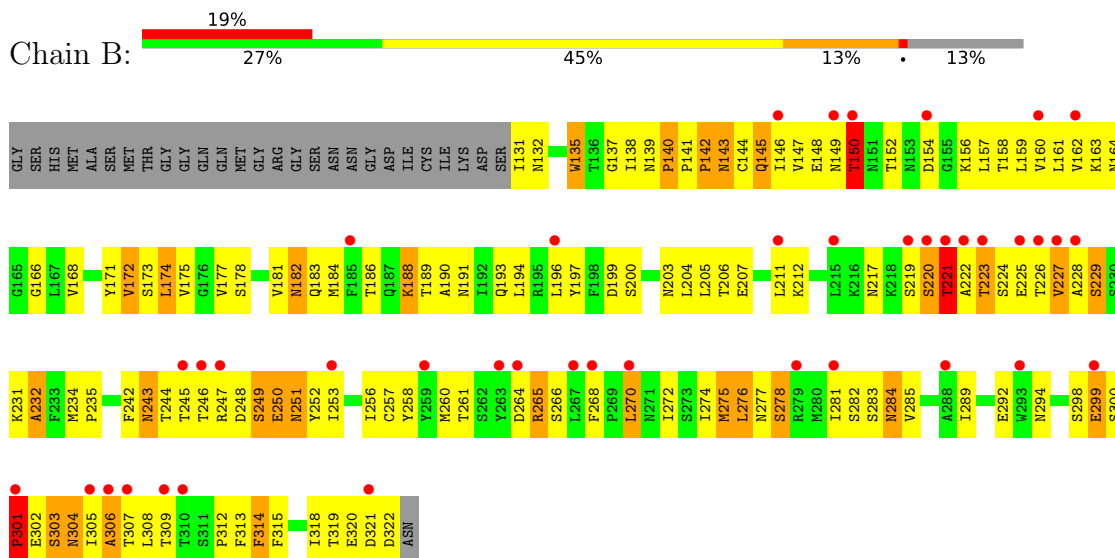
Chain F:

- Molecule 1: Fiber

Chain G:

20%

34%　　　41%　　12%　　13%

4   Data and refinement statistics ⓘ

Property	Value	Source
Space group	I 4	Depositor
Cell constants a, b, c, α, β, γ	173.91Å 173.91Å 154.15Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.03 – 2.70 46.03 – 2.70	Depositor EDS
% Data completeness (in resolution range)	91.5 (46.03-2.70) 91.5 (46.03-2.70)	Depositor EDS
R_{merge}	0.15	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.33 (at 2.69Å)	Xtriage
Refinement program	CNS 1.2	Depositor
R, R_{free}	0.332 , 0.352 0.332 , 0.353	Depositor DCC
R_{free} test set	2911 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å^2)	62.8	Xtriage
Anisotropy	0.829	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 66.8	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.468 for -k,-h,-l	Xtriage
F_o,F_c correlation	0.87	EDS
Total number of atoms	8976	wwPDB-VP
Average B, all atoms (Å^2)	85.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.34% 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	0.50	0/1527	0.80	1/2084 (0.0%)
1	B	0.48	0/1527	0.82	1/2084 (0.0%)
1	D	0.47	0/1527	0.75	1/2084 (0.0%)
1	E	0.50	1/1527 (0.1%)	0.80	1/2084 (0.0%)
1	F	0.49	0/1527	0.82	1/2084 (0.0%)
1	G	0.48	0/1527	0.75	1/2084 (0.0%)
All	All	0.49	1/9162 (0.0%)	0.79	6/12504 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	144	CYS	CB-SG	-6.08	1.72	1.82

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	221	THR	N-CA-C	-7.55	90.62	111.00
1	A	221	THR	N-CA-C	-7.17	91.64	111.00
1	G	221	THR	N-CA-C	-6.30	94.00	111.00
1	D	221	THR	N-CA-C	-6.18	94.31	111.00
1	B	221	THR	N-CA-C	-5.64	95.76	111.00

There are no chirality outliers.

There are no planarity outliers.

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.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1496	0	1453	153	1
1	B	1496	0	1453	154	0
1	D	1496	0	1453	160	0
1	E	1496	0	1453	149	1
1	F	1496	0	1453	133	0
1	G	1496	0	1453	150	0
All	All	8976	0	8718	840	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 48.

The worst 5 of 840 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:E:217:ASN:ND2	1:E:224:SER:H	1.57	1.00
1:D:261:THR:HG22	1:D:301:PRO:HG3	1.45	0.98
1:A:143:ASN:HD21	1:A:157:LEU:H	1.09	0.96
1:B:268:PHE:HE1	1:B:300:SER:HA	1.27	0.96
1:F:242:PHE:CE1	1:F:277:ASN:ND2	2.38	0.92

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)
1:A:186:THR:O	1:E:302:GLU:OE1[7_444]	2.08	0.12

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	190/220 (86%)	139 (73%)	29 (15%)	22 (12%)	0	0
1	B	190/220 (86%)	132 (70%)	31 (16%)	27 (14%)	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	D	190/220 (86%)	139 (73%)	30 (16%)	21 (11%)	0	0
1	E	190/220 (86%)	135 (71%)	32 (17%)	23 (12%)	0	0
1	F	190/220 (86%)	139 (73%)	28 (15%)	23 (12%)	0	0
1	G	190/220 (86%)	136 (72%)	30 (16%)	24 (13%)	0	0
All	All	1140/1320 (86%)	820 (72%)	180 (16%)	140 (12%)	0	0

5 of 140 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	220	SER
1	A	221	THR
1	A	248	ASP
1	A	250	GLU
1	A	304	ASN

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	175/196 (89%)	159 (91%)	16 (9%)	9	21
1	B	175/196 (89%)	159 (91%)	16 (9%)	9	21
1	D	175/196 (89%)	155 (89%)	20 (11%)	5	13
1	E	175/196 (89%)	156 (89%)	19 (11%)	6	14
1	F	175/196 (89%)	153 (87%)	22 (13%)	4	10
1	G	175/196 (89%)	161 (92%)	14 (8%)	12	27
All	All	1050/1176 (89%)	943 (90%)	107 (10%)	7	17

5 of 107 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	173	SER
1	F	135	TRP

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	G	200	SER
1	E	188	LYS
1	E	270	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 62 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	243	ASN
1	G	203	ASN
1	E	145	GLN
1	G	164	ASN
1	G	284	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, 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	192/220 (87%)	1.54	39 (20%)	1	0	34, 79, 108, 113	72 (37%)
1	B	192/220 (87%)	1.46	41 (21%)	0	0	41, 83, 108, 113	77 (40%)
1	D	192/220 (87%)	1.45	40 (20%)	1	0	51, 90, 108, 113	62 (32%)
1	E	192/220 (87%)	1.49	37 (19%)	1	0	42, 79, 108, 113	70 (36%)
1	F	192/220 (87%)	1.54	37 (19%)	1	0	42, 80, 108, 113	69 (35%)
1	G	192/220 (87%)	1.29	43 (22%)	0	0	55, 89, 108, 113	69 (35%)
All	All	1152/1320 (87%)	1.46	237 (20%)	1	0	34, 83, 108, 113	419 (36%)

The worst 5 of 237 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	246	THR	29.7
1	A	225	GLU	20.3
1	D	223	THR	17.1
1	E	225	GLU	16.4
1	A	224	SER	16.0

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 ⓘ

There are no such residues in this entry.