Full wwPDB X-ray Structure Validation Report ⓘ

Jun 27, 2024 – 12:43 AM JST

PDB ID	:	8JXK
Title	:	Crystal Structure of Rv0047c from Mycobacterium tuberculosis
Authors	:	Ansari, M.S.; Yadav, V.; Zohib, M.; Pal, R.K.; Biswal, B.K.; Arora, A.
Deposited on	:	2023-06-30
Resolution	:	3.15 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 ⓘ 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 ⓘ) were used in the production of this report:

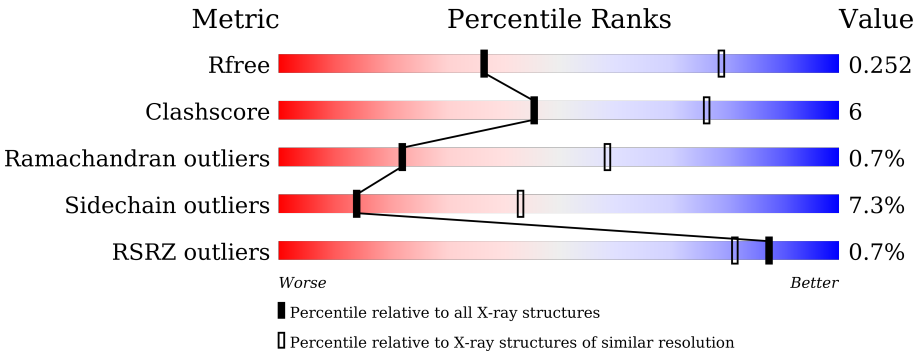
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 3.15 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	1665 (3.20-3.12)
Clashscore	141614	1804 (3.20-3.12)
Ramachandran outliers	138981	1770 (3.20-3.12)
Sidechain outliers	138945	1769 (3.20-3.12)
RSRZ outliers	127900	1616 (3.20-3.12)

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	182	71% 13% • 15%
1	B	182	% 69% 12% • 17%
1	C	182	74% 15% • 10%
1	D	182	80% 9% • 9%
1	E	182	% 76% 8% • 14%
1	F	182	2% 70% 9% • 20%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 6561 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 Conserved protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	154	Total	C	N	O	S	0	0	0
			1144	717	213	210	4			
1	B	151	Total	C	N	O	S	0	0	0
			981	604	185	191	1			
1	C	163	Total	C	N	O	S	0	0	0
			1169	729	223	214	3			
1	D	166	Total	C	N	O	S	0	0	0
			1202	756	227	216	3			
1	E	156	Total	C	N	O	S	0	0	0
			1088	679	203	204	2			
1	F	145	Total	C	N	O	S	0	0	0
			965	591	179	193	2			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	GLU	-	expression tag	UNP P71704
A	0	PHE	-	expression tag	UNP P71704
B	-1	GLU	-	expression tag	UNP P71704
B	0	PHE	-	expression tag	UNP P71704
C	-1	GLU	-	expression tag	UNP P71704
C	0	PHE	-	expression tag	UNP P71704
D	-1	GLU	-	expression tag	UNP P71704
D	0	PHE	-	expression tag	UNP P71704
E	-1	GLU	-	expression tag	UNP P71704
E	0	PHE	-	expression tag	UNP P71704
F	-1	GLU	-	expression tag	UNP P71704
F	0	PHE	-	expression tag	UNP P71704

- Molecule 2 is water.

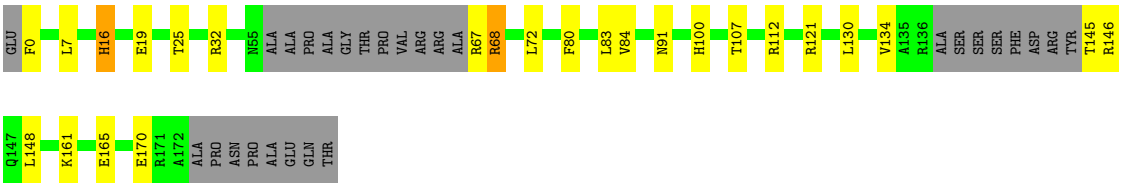
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	2	Total O 2 2	0	0
2	B	2	Total O 2 2	0	0
2	C	1	Total O 1 1	0	0
2	D	1	Total O 1 1	0	0
2	E	3	Total O 3 3	0	0
2	F	3	Total O 3 3	0	0

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.

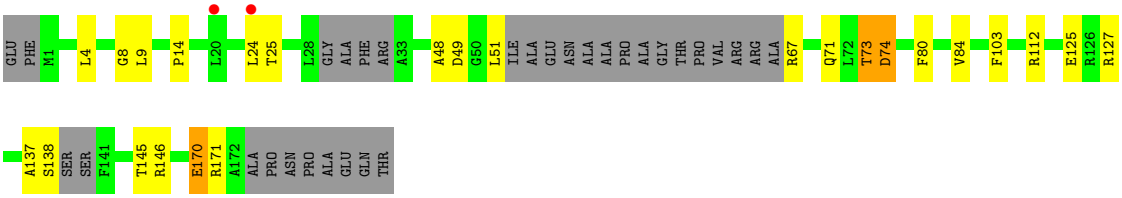
• Molecule 1: Conserved protein

Chain A:



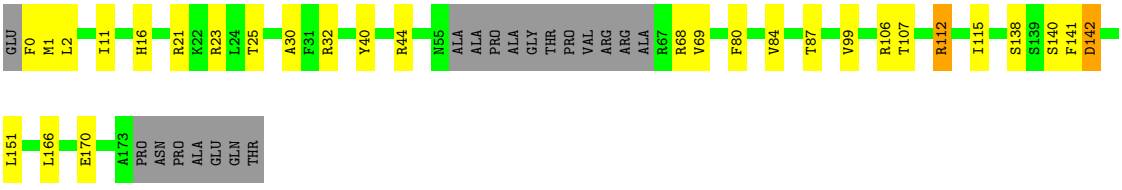
• Molecule 1: Conserved protein

Chain B:



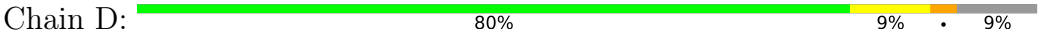
• Molecule 1: Conserved protein

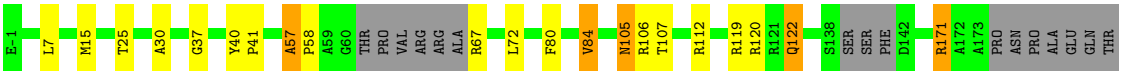
Chain C:



• Molecule 1: Conserved protein

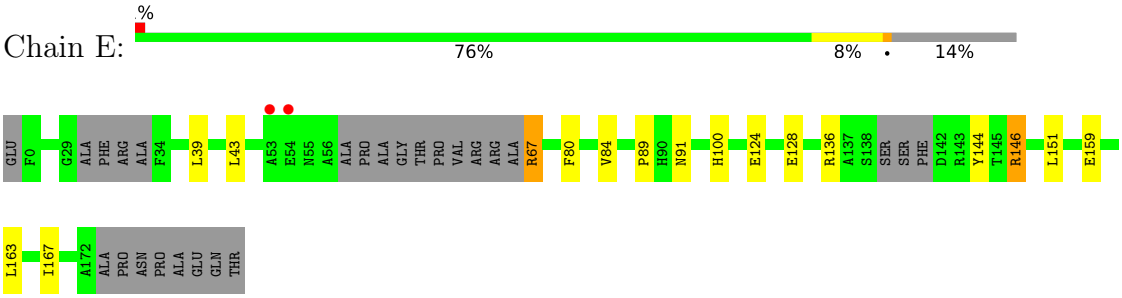
Chain D:





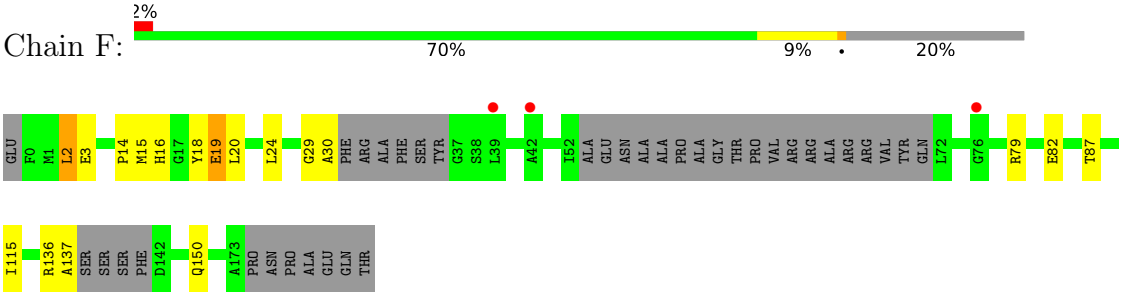
• Molecule 1: Conserved protein

Chain E:



• Molecule 1: Conserved protein

Chain F:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 61	Depositor
Cell constants a, b, c, α, β, γ	102.17Å 102.17Å 191.88Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	30.08 – 3.15 30.08 – 3.15	Depositor EDS
% Data completeness (in resolution range)	90.5 (30.08-3.15) 90.6 (30.08-3.15)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.92 (at 3.18Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.204 , 0.253 0.206 , 0.252	Depositor DCC
R_{free} test set	856 reflections (4.82%)	wwPDB-VP
Wilson B-factor (Å^2)	74.6	Xtriage
Anisotropy	0.597	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 73.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.43$, $< L^2 > = 0.25$	Xtriage
Estimated twinning fraction	0.286 for h,-h-k,-l	Xtriage
Reported twinning fraction	0.693 for H, K, L 0.307 for -K, -H, -L	Depositor
Outliers	0 of 17741 reflections	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	6561	wwPDB-VP
Average B, all atoms (Å^2)	79.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.53% 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	0/1158	0.84	2/1558 (0.1%)
1	B	0.49	1/991 (0.1%)	0.79	1/1346 (0.1%)
1	C	0.44	1/1187 (0.1%)	0.90	6/1607 (0.4%)
1	D	0.48	0/1219	0.90	2/1646 (0.1%)
1	E	0.43	0/1102	0.74	0/1495
1	F	0.47	0/974	0.80	0/1323
All	All	0.46	2/6631 (0.0%)	0.83	11/8975 (0.1%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	125	GLU	CD-OE1	6.34	1.32	1.25
1	C	32	ARG	CD-NE	6.11	1.56	1.46

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	32	ARG	NE-CZ-NH1	-8.99	115.81	120.30
1	C	32	ARG	CG-CD-NE	8.74	130.15	111.80
1	C	32	ARG	NE-CZ-NH2	8.16	124.38	120.30
1	D	119	ARG	NE-CZ-NH1	-6.43	117.09	120.30
1	A	121	ARG	CG-CD-NE	-6.39	98.38	111.80
1	A	146	ARG	CB-CA-C	6.02	122.44	110.40
1	C	106	ARG	CG-CD-NE	-5.91	99.39	111.80
1	D	171	ARG	CG-CD-NE	-5.88	99.44	111.80
1	C	112	ARG	CB-CG-CD	5.82	126.74	111.60
1	B	49	ASP	CB-CA-C	5.19	120.78	110.40
1	C	32	ARG	CD-NE-CZ	5.04	130.65	123.60

There are no chirality outliers.

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	1144	0	1068	17	0
1	B	981	0	771	10	0
1	C	1169	0	1051	16	0
1	D	1202	0	1126	13	0
1	E	1088	0	945	16	0
1	F	965	0	795	10	0
2	A	2	0	0	0	0
2	B	2	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
2	E	3	0	0	0	0
2	F	3	0	0	0	0
All	All	6561	0	5756	79	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 6.

All (79) 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:57:ALA:HB1	1:D:58:PRO:HD3	1.43	0.98
1:E:146:ARG:HH11	1:E:146:ARG:HB2	1.46	0.79
1:D:105:ASN:HD22	1:D:106:ARG:HG3	1.47	0.79
1:E:146:ARG:HH11	1:E:146:ARG:CB	1.96	0.79
1:E:163:LEU:O	1:E:167:ILE:HG13	1.85	0.77
1:D:57:ALA:HB1	1:D:58:PRO:CD	2.16	0.74
1:F:20:LEU:O	1:F:24:LEU:HB2	1.90	0.72
1:C:25:THR:HG23	1:C:30:ALA:HB3	1.74	0.70
1:A:16:HIS:H	1:A:16:HIS:CD2	2.11	0.67
1:D:25:THR:HG23	1:D:30:ALA:HB3	1.76	0.67
1:F:29:GLY:O	1:F:30:ALA:HB3	1.96	0.66
1:B:112:ARG:NE	1:B:170:GLU:OE2	2.30	0.64
1:B:51:LEU:C	1:B:73:THR:HG22	2.19	0.63
1:F:29:GLY:O	1:F:30:ALA:CB	2.49	0.61
1:E:67:ARG:NE	1:E:67:ARG:HA	2.16	0.61

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:F:15:MET:CG	1:F:16:HIS:H	2.14	0.60
1:C:112:ARG:NH1	1:C:170:GLU:OE1	2.35	0.59
1:E:124:GLU:O	1:E:128:GLU:HG2	2.02	0.59
1:B:51:LEU:C	1:B:73:THR:CG2	2.72	0.59
1:A:91:ASN:ND2	1:A:100:HIS:HD2	2.00	0.58
1:E:91:ASN:ND2	1:E:100:HIS:HD2	2.01	0.57
1:C:16:HIS:CD2	1:C:69:VAL:HG22	2.39	0.57
1:D:57:ALA:CB	1:D:58:PRO:CD	2.83	0.57
1:C:16:HIS:CD2	1:C:69:VAL:CG2	2.89	0.56
1:C:1:MET:HG2	1:C:99:VAL:HG21	1.87	0.55
1:B:4:LEU:O	1:B:8:GLY:N	2.37	0.55
1:C:107:THR:HB	1:C:112:ARG:HD3	1.89	0.53
1:A:7:LEU:HD23	1:A:72:LEU:HD11	1.90	0.53
1:A:16:HIS:CD2	1:A:16:HIS:N	2.77	0.52
1:A:0:PHE:CE2	1:A:83:LEU:HD22	2.44	0.52
1:E:159:GLU:O	1:E:163:LEU:HD12	2.10	0.52
1:A:161:LYS:O	1:A:165:GLU:HG3	2.09	0.52
1:F:15:MET:CG	1:F:16:HIS:N	2.74	0.51
1:A:68:ARG:HD3	1:A:68:ARG:H	1.74	0.51
1:B:137:ALA:CB	1:B:146:ARG:HG3	2.42	0.50
1:E:67:ARG:HA	1:E:67:ARG:HE	1.76	0.50
1:D:40:TYR:N	1:D:41:PRO:HD2	2.27	0.50
1:E:39:LEU:HD11	1:E:43:LEU:HD11	1.94	0.49
1:A:134:VAL:O	1:A:134:VAL:HG12	2.13	0.49
1:A:112:ARG:NE	1:A:170:GLU:OE1	2.35	0.49
1:C:166:LEU:HD11	1:E:144:TYR:CZ	2.48	0.49
1:B:74:ASP:OD1	1:B:74:ASP:N	2.34	0.49
1:D:80:PHE:O	1:D:84:VAL:HG22	2.12	0.48
1:C:80:PHE:CZ	1:C:84:VAL:HG11	2.49	0.48
1:C:107:THR:O	1:C:112:ARG:NE	2.44	0.48
1:A:80:PHE:CZ	1:A:84:VAL:HG11	2.49	0.48
1:C:138:SER:O	1:C:141:PHE:CB	2.62	0.47
1:F:136:ARG:O	1:F:137:ALA:HB2	2.15	0.47
1:D:7:LEU:HD23	1:D:72:LEU:HD11	1.96	0.47
1:C:166:LEU:HD11	1:E:144:TYR:CE1	2.49	0.47
1:E:146:ARG:HH11	1:E:146:ARG:CG	2.27	0.47
1:F:79:ARG:HA	1:F:82:GLU:HG3	1.98	0.46
1:D:105:ASN:ND2	1:D:106:ARG:HG3	2.25	0.45
1:B:9:LEU:CB	1:B:103:PHE:CZ	2.99	0.45
1:D:37:GLY:O	1:D:41:PRO:HG3	2.17	0.45
1:E:80:PHE:CZ	1:E:84:VAL:HG11	2.52	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:F:18:TYR:O	1:F:19:GLU:CB	2.65	0.44
1:A:68:ARG:H	1:A:68:ARG:CD	2.30	0.44
1:B:80:PHE:O	1:B:84:VAL:HG22	2.18	0.44
1:A:16:HIS:CD2	1:A:19:GLU:OE1	2.71	0.43
1:B:14:PRO:HB3	1:B:71:GLN:HA	1.98	0.43
1:B:9:LEU:CB	1:B:103:PHE:CE1	3.00	0.43
1:A:130:LEU:O	1:A:134:VAL:HG23	2.19	0.43
1:E:163:LEU:O	1:E:167:ILE:N	2.51	0.43
1:A:107:THR:O	1:A:112:ARG:NH1	2.53	0.42
1:C:166:LEU:CD1	1:E:144:TYR:CE1	3.02	0.42
1:C:16:HIS:HD2	1:C:69:VAL:HG22	1.83	0.42
1:A:80:PHE:O	1:A:84:VAL:HG22	2.20	0.42
1:A:7:LEU:CD2	1:A:72:LEU:HD11	2.49	0.42
1:A:25:THR:CG2	1:A:32:ARG:O	2.68	0.42
1:C:87:THR:HG22	1:C:115:ILE:HG13	2.02	0.42
1:D:107:THR:O	1:D:112:ARG:NH1	2.53	0.42
1:D:122:GLN:HE21	1:D:122:GLN:HA	1.85	0.41
1:E:80:PHE:O	1:E:84:VAL:HG22	2.21	0.41
1:F:87:THR:HG22	1:F:115:ILE:HG13	2.03	0.41
1:C:16:HIS:HA	1:C:68:ARG:O	2.21	0.41
1:F:2:LEU:C	1:F:2:LEU:HD12	2.42	0.40
1:D:7:LEU:CD2	1:D:72:LEU:HD11	2.51	0.40
1:C:40:TYR:O	1:C:44:ARG:HG2	2.21	0.40

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	148/182 (81%)	146 (99%)	2 (1%)	0	100 100
1	B	143/182 (79%)	135 (94%)	7 (5%)	1 (1%)	22 59

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	C	159/182 (87%)	157 (99%)	1 (1%)	1 (1%)	25 62
1	D	160/182 (88%)	156 (98%)	3 (2%)	1 (1%)	25 62
1	E	148/182 (81%)	147 (99%)	1 (1%)	0	100 100
1	F	137/182 (75%)	129 (94%)	5 (4%)	3 (2%)	6 32
All	All	895/1092 (82%)	870 (97%)	19 (2%)	6 (1%)	22 59

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	48	ALA
1	C	142	ASP
1	F	14	PRO
1	D	57	ALA
1	F	3	GLU
1	F	19	GLU

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	99/145 (68%)	94 (95%)	5 (5%)	24 56
1	B	62/145 (43%)	52 (84%)	10 (16%)	2 10
1	C	95/145 (66%)	87 (92%)	8 (8%)	11 37
1	D	100/145 (69%)	93 (93%)	7 (7%)	15 45
1	E	83/145 (57%)	78 (94%)	5 (6%)	19 51
1	F	70/145 (48%)	68 (97%)	2 (3%)	42 72
All	All	509/870 (58%)	472 (93%)	37 (7%)	14 43

All (37) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	16	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	67	ARG
1	A	68	ARG
1	A	145	THR
1	A	148	LEU
1	B	24	LEU
1	B	25	THR
1	B	67	ARG
1	B	73	THR
1	B	74	ASP
1	B	127	ARG
1	B	138	SER
1	B	145	THR
1	B	170	GLU
1	B	171	ARG
1	C	0	PHE
1	C	2	LEU
1	C	11	ILE
1	C	21	ARG
1	C	23	ARG
1	C	140	SER
1	C	142	ASP
1	C	151	LEU
1	D	15	MET
1	D	67	ARG
1	D	84	VAL
1	D	105	ASN
1	D	120	ARG
1	D	122	GLN
1	D	171	ARG
1	E	67	ARG
1	E	89	PRO
1	E	136	ARG
1	E	146	ARG
1	E	151	LEU
1	F	2	LEU
1	F	150	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (17) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	16	HIS
1	A	91	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	100	HIS
1	A	147	GLN
1	B	91	ASN
1	C	16	HIS
1	C	91	ASN
1	C	100	HIS
1	D	91	ASN
1	D	100	HIS
1	D	105	ASN
1	D	122	GLN
1	E	71	GLN
1	E	91	ASN
1	E	100	HIS
1	F	91	ASN
1	F	100	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	154/182 (84%)	-0.13	0 100 100	47, 68, 97, 109	0
1	B	151/182 (82%)	-0.14	2 (1%) 77 66	55, 86, 129, 140	0
1	C	163/182 (89%)	-0.03	0 100 100	55, 77, 112, 128	0
1	D	166/182 (91%)	-0.16	0 100 100	52, 72, 95, 107	0
1	E	156/182 (85%)	-0.08	2 (1%) 77 66	58, 83, 107, 119	0
1	F	145/182 (79%)	0.20	3 (2%) 63 49	55, 83, 125, 145	0
All	All	935/1092 (85%)	-0.06	7 (0%) 87 81	47, 77, 116, 145	0

All (7) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	54	GLU	4.4
1	E	53	ALA	3.3
1	F	42	ALA	2.8
1	B	20	LEU	2.7
1	F	76	GLY	2.7
1	B	24	LEU	2.2
1	F	39	LEU	2.1

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.