Full wwPDB X-ray Structure Validation Report (i)

Oct 17, 2022 – 10:13 pm BST

PDB ID	:	7ZGW
Title	:	Serratia NucC apo form
Authors	:	Garcia-Doval, C.; Mayo-Munoz, D.; Smith, L.M.; Fineran, P.C.
Deposited on	:	2022-04-04
Resolution	:	1.83 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 (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:

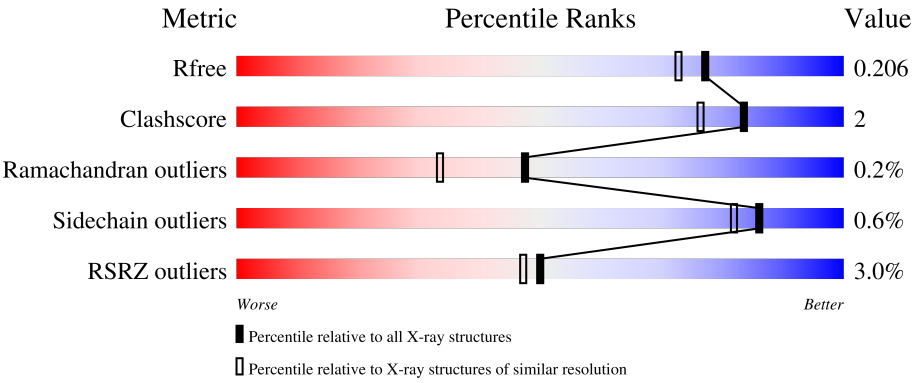
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.31.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.2

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 1.83 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.



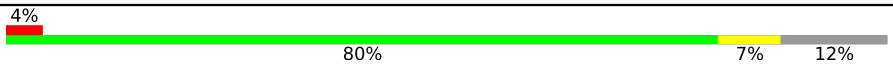
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	4003 (1.86-1.82)
Clashscore	141614	4233 (1.86-1.82)
Ramachandran outliers	138981	4185 (1.86-1.82)
Sidechain outliers	138945	4186 (1.86-1.82)
RSRZ outliers	127900	3957 (1.86-1.82)

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	256	88% · 7%
1	B	256	86% 7% 7%
1	C	256	5% 79% 7% 14%
1	D	256	4% 81% 6% 13%
1	E	256	3% 86% 7% 7%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	256	

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 22746 atoms, of which 10920 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 Serratia NucC.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	238	Total	C	H	N	O	S	0	0	0
			3759	1202	1871	320	363	3			
1	B	239	Total	C	H	N	O	S	0	0	0
			3775	1207	1877	321	367	3			
1	D	223	Total	C	H	N	O	S	0	0	0
			3529	1130	1752	300	345	2			
1	C	221	Total	C	H	N	O	S	0	0	0
			3508	1123	1744	298	341	2			
1	E	239	Total	C	H	N	O	S	0	0	0
			3775	1207	1877	321	367	3			
1	F	226	Total	C	H	N	O	S	0	0	0
			3603	1150	1799	308	344	2			

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-5	LYS	-	expression tag	UNP A0A2I5TBB8
A	-4	GLU	-	expression tag	UNP A0A2I5TBB8
A	-3	GLU	-	expression tag	UNP A0A2I5TBB8
A	-2	LYS	-	expression tag	UNP A0A2I5TBB8
A	-1	LEU	-	expression tag	UNP A0A2I5TBB8
A	0	THR	-	expression tag	UNP A0A2I5TBB8
B	-5	LYS	-	expression tag	UNP A0A2I5TBB8
B	-4	GLU	-	expression tag	UNP A0A2I5TBB8
B	-3	GLU	-	expression tag	UNP A0A2I5TBB8
B	-2	LYS	-	expression tag	UNP A0A2I5TBB8
B	-1	LEU	-	expression tag	UNP A0A2I5TBB8
B	0	THR	-	expression tag	UNP A0A2I5TBB8
D	-5	LYS	-	expression tag	UNP A0A2I5TBB8
D	-4	GLU	-	expression tag	UNP A0A2I5TBB8
D	-3	GLU	-	expression tag	UNP A0A2I5TBB8
D	-2	LYS	-	expression tag	UNP A0A2I5TBB8
D	-1	LEU	-	expression tag	UNP A0A2I5TBB8

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	0	THR	-	expression tag	UNP A0A2I5TBB8
C	-5	LYS	-	expression tag	UNP A0A2I5TBB8
C	-4	GLU	-	expression tag	UNP A0A2I5TBB8
C	-3	GLU	-	expression tag	UNP A0A2I5TBB8
C	-2	LYS	-	expression tag	UNP A0A2I5TBB8
C	-1	LEU	-	expression tag	UNP A0A2I5TBB8
C	0	THR	-	expression tag	UNP A0A2I5TBB8
E	-5	LYS	-	expression tag	UNP A0A2I5TBB8
E	-4	GLU	-	expression tag	UNP A0A2I5TBB8
E	-3	GLU	-	expression tag	UNP A0A2I5TBB8
E	-2	LYS	-	expression tag	UNP A0A2I5TBB8
E	-1	LEU	-	expression tag	UNP A0A2I5TBB8
E	0	THR	-	expression tag	UNP A0A2I5TBB8
F	-5	LYS	-	expression tag	UNP A0A2I5TBB8
F	-4	GLU	-	expression tag	UNP A0A2I5TBB8
F	-3	GLU	-	expression tag	UNP A0A2I5TBB8
F	-2	LYS	-	expression tag	UNP A0A2I5TBB8
F	-1	LEU	-	expression tag	UNP A0A2I5TBB8
F	0	THR	-	expression tag	UNP A0A2I5TBB8

- Molecule 2 is water.

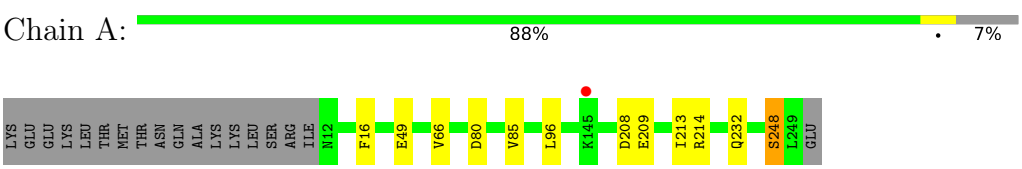
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	197	Total O 197 197	0	0
2	B	162	Total O 162 162	0	0
2	D	114	Total O 114 114	0	0
2	C	102	Total O 102 102	0	0
2	E	149	Total O 149 149	0	0
2	F	73	Total O 73 73	0	0

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.

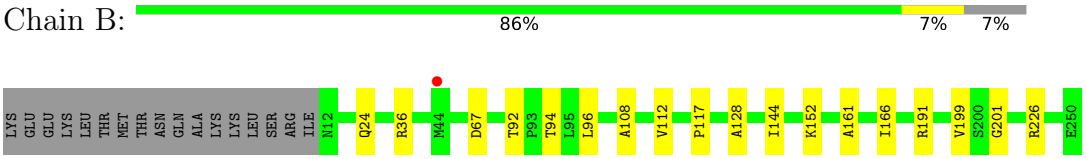
• Molecule 1: Serratia NucC

Chain A:



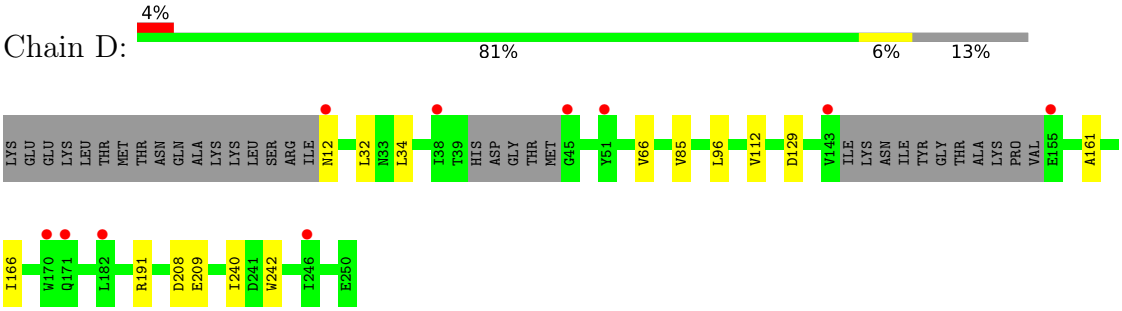
• Molecule 1: Serratia NucC

Chain B:



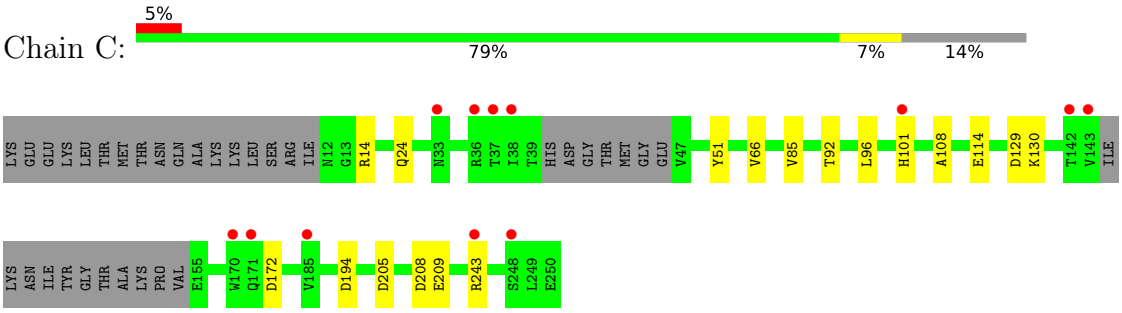
• Molecule 1: Serratia NucC

Chain D:



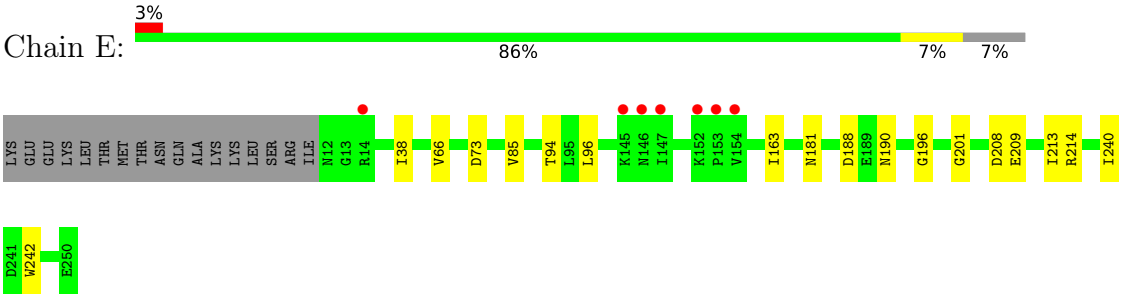
• Molecule 1: Serratia NucC

Chain C:



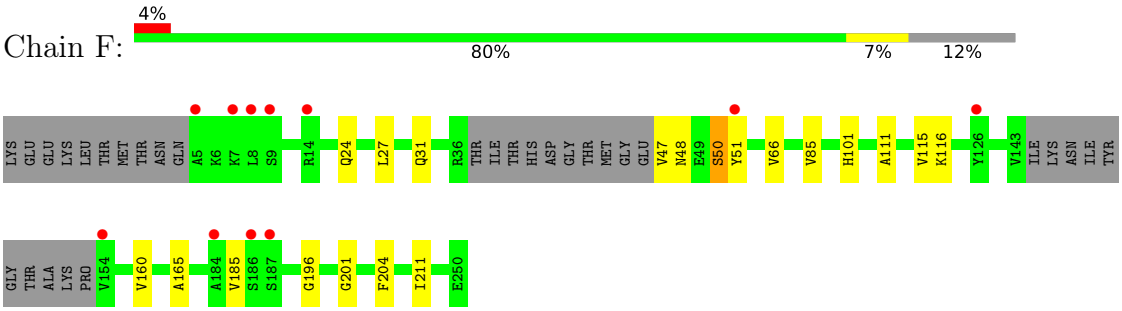
• Molecule 1: Serratia NucC

Chain E:



• Molecule 1: Serratia NucC

Chain F:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	79.47Å 95.49Å 94.54Å 90.00° 92.05° 90.00°	Depositor
Resolution (Å)	42.61 – 1.83 42.61 – 1.83	Depositor EDS
% Data completeness (in resolution range)	100.0 (42.61-1.83) 95.6 (42.61-1.83)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.17 (at 1.83Å)	Xtriage
Refinement program	PHENIX 1.19.2_4158	Depositor
R, R_{free}	0.167 , 0.208 0.165 , 0.206	Depositor DCC
R_{free} test set	6326 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å^2)	26.1	Xtriage
Anisotropy	0.147	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.013 for -h,-l,-k 0.000 for -h,l,k 0.020 for h,-k,-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	22746	wwPDB-VP
Average B, all atoms (Å^2)	38.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.69% 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	0.56	0/1924	0.70	0/2609
1	B	0.55	0/1934	0.70	1/2621 (0.0%)
1	C	0.48	0/1795	0.65	0/2430
1	D	0.53	0/1808	0.65	0/2447
1	E	0.53	0/1934	0.68	1/2621 (0.0%)
1	F	0.43	0/1835	0.62	0/2482
All	All	0.52	0/11230	0.67	2/15210 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	67	ASP	CB-CG-OD1	6.17	123.85	118.30
1	E	73	ASP	CB-CG-OD1	5.04	122.84	118.30

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.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1888	1871	1870	9	0
1	B	1898	1877	1876	11	0
1	C	1764	1744	1741	11	0
1	D	1777	1752	1750	11	0
1	E	1898	1877	1876	8	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	F	1804	1799	1796	14	0
2	A	197	0	0	0	0
2	B	162	0	0	2	0
2	C	102	0	0	0	0
2	D	114	0	0	0	0
2	E	149	0	0	0	0
2	F	73	0	0	0	0
All	All	11826	10920	10909	53	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 2.

All (53) 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:A:80:ASP:OD1	1:E:38:ILE:HD12	1.88	0.74
1:D:129:ASP:OD1	1:D:191:ARG:NH1	2.28	0.66
1:B:152:LYS:NZ	2:B:301:HOH:O	2.32	0.62
1:B:166:ILE:HA	1:B:199:VAL:HG22	1.86	0.57
1:B:94:THR:HG22	1:B:96:LEU:H	1.72	0.55
1:B:24:GLN:HG3	1:C:96:LEU:HD13	1.88	0.54
1:F:185:VAL:O	1:F:185:VAL:HG12	2.08	0.53
1:D:240:ILE:HD11	1:D:242:TRP:CE3	2.46	0.50
1:C:114:GLU:OE2	1:C:130:LYS:NZ	2.39	0.50
1:E:66:VAL:HA	1:E:85:VAL:O	2.12	0.49
1:E:240:ILE:HD11	1:E:242:TRP:CE3	2.47	0.49
1:C:101:HIS:CG	1:F:51:TYR:OH	2.66	0.49
1:E:163:ILE:O	1:E:196:GLY:HA2	2.13	0.48
1:C:66:VAL:HA	1:C:85:VAL:O	2.13	0.48
1:B:128:ALA:HB1	1:B:191:ARG:HB3	1.96	0.48
1:F:196:GLY:HA3	1:F:204:PHE:CE1	2.49	0.48
1:F:27:LEU:O	1:F:31:GLN:HG3	2.15	0.47
1:F:31:GLN:HB3	1:F:51:TYR:CD2	2.50	0.47
1:A:213:ILE:O	1:A:214:ARG:HD2	2.16	0.46
1:C:51:TYR:CZ	1:F:101:HIS:HB2	2.51	0.46
1:A:16:PHE:CD2	1:D:34:LEU:HD22	2.50	0.46
1:D:66:VAL:HA	1:D:85:VAL:O	2.15	0.46
1:A:96:LEU:HD13	1:C:24:GLN:HG3	1.98	0.46
1:F:211:ILE:HD12	1:F:211:ILE:N	2.31	0.45
1:F:47:VAL:CG2	1:F:115:VAL:HG12	2.46	0.45
1:B:92:THR:HG21	1:B:108:ALA:HB2	1.98	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:92:THR:HG21	1:C:108:ALA:HB2	1.99	0.44
1:E:188:ASP:OD2	1:E:190:ASN:HB2	2.18	0.44
1:F:111:ALA:HA	1:F:160:VAL:O	2.18	0.43
1:D:32:LEU:HB3	1:D:166:ILE:HG12	2.00	0.43
1:E:208:ASP:O	1:E:209:GLU:HB2	2.18	0.43
1:A:66:VAL:HA	1:A:85:VAL:O	2.18	0.43
1:A:232:GLN:OE1	1:B:144:ILE:HD12	2.19	0.43
1:C:208:ASP:O	1:C:209:GLU:HB2	2.18	0.43
1:B:226:ARG:NH1	1:C:243:ARG:NH1	2.67	0.43
1:E:213:ILE:O	1:E:214:ARG:HD2	2.18	0.42
1:F:31:GLN:HB3	1:F:51:TYR:CE2	2.54	0.42
1:D:240:ILE:C	1:D:240:ILE:HD12	2.39	0.42
1:D:208:ASP:O	1:D:209:GLU:HB2	2.19	0.42
1:C:194:ASP:O	1:C:205:ASP:HA	2.20	0.42
1:E:94:THR:HG22	1:E:96:LEU:H	1.84	0.42
1:F:66:VAL:HA	1:F:85:VAL:O	2.20	0.41
1:F:48:ASN:OD1	1:F:50:SER:HB2	2.20	0.41
1:F:116:LYS:O	1:F:165:ALA:HA	2.20	0.41
1:D:32:LEU:O	1:D:166:ILE:HD11	2.20	0.41
1:A:208:ASP:O	1:A:209:GLU:HB2	2.21	0.41
1:A:16:PHE:CG	1:D:34:LEU:HD22	2.55	0.41
1:B:117:PRO:O	1:B:166:ILE:HG22	2.21	0.41
1:A:248:SER:HB3	1:C:14:ARG:NH1	2.36	0.40
1:B:112:VAL:O	1:B:161:ALA:HA	2.21	0.40
1:D:96:LEU:HD13	1:F:24:GLN:HG3	2.02	0.40
1:D:112:VAL:O	1:D:161:ALA:HA	2.22	0.40
1:B:36:ARG:HD3	2:B:426:HOH:O	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	236/256 (92%)	230 (98%)	6 (2%)	0	100	100
1	B	237/256 (93%)	231 (98%)	5 (2%)	1 (0%)	34	20
1	C	215/256 (84%)	210 (98%)	5 (2%)	0	100	100
1	D	217/256 (85%)	212 (98%)	5 (2%)	0	100	100
1	E	237/256 (93%)	232 (98%)	4 (2%)	1 (0%)	34	20
1	F	220/256 (86%)	216 (98%)	3 (1%)	1 (0%)	29	15
All	All	1362/1536 (89%)	1331 (98%)	28 (2%)	3 (0%)	47	33

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	201	GLY
1	E	201	GLY
1	F	201	GLY

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	205/222 (92%)	203 (99%)	2 (1%)	76	68
1	B	206/222 (93%)	206 (100%)	0	100	100
1	C	192/222 (86%)	190 (99%)	2 (1%)	76	68
1	D	193/222 (87%)	192 (100%)	1 (0%)	88	85
1	E	206/222 (93%)	205 (100%)	1 (0%)	88	85
1	F	196/222 (88%)	195 (100%)	1 (0%)	88	85
All	All	1198/1332 (90%)	1191 (99%)	7 (1%)	86	82

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	49	GLU
1	A	248	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	12	ASN
1	C	129	ASP
1	C	172	ASP
1	E	181	ASN
1	F	50	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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	238/256 (92%)	0.05	1 (0%)	92	92	18, 25, 41, 67	0
1	B	239/256 (93%)	-0.01	1 (0%)	92	92	18, 29, 51, 80	0
1	C	221/256 (86%)	0.33	12 (5%)	25	23	22, 34, 69, 100	0
1	D	223/256 (87%)	0.24	10 (4%)	33	30	21, 33, 62, 94	0
1	E	239/256 (93%)	0.15	7 (2%)	51	49	21, 29, 53, 84	0
1	F	226/256 (88%)	0.38	11 (4%)	29	27	27, 38, 74, 94	0
All	All	1386/1536 (90%)	0.19	42 (3%)	50	47	18, 32, 62, 100	0

All (42) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	38	ILE	6.6
1	F	184	ALA	6.2
1	A	145	LYS	5.7
1	D	171	GLN	5.1
1	C	171	GLN	4.1
1	C	37	THR	4.1
1	D	246	ILE	3.9
1	C	142	THR	3.8
1	C	143	VAL	3.6
1	E	152	LYS	3.5
1	F	7	LYS	3.4
1	D	51	TYR	3.2
1	F	51	TYR	3.2
1	C	185	VAL	3.2
1	E	154	VAL	3.2
1	F	154	VAL	3.1
1	C	101	HIS	3.1
1	B	44	MET	3.0
1	F	9	SER	3.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	248	SER	2.9
1	D	45	GLY	2.9
1	C	170	TRP	2.8
1	C	243	ARG	2.8
1	F	186	SER	2.5
1	E	146	ASN	2.5
1	F	126	TYR	2.5
1	F	187	SER	2.4
1	D	170	TRP	2.4
1	F	5	ALA	2.4
1	C	33	ASN	2.3
1	E	145	LYS	2.3
1	D	182	LEU	2.2
1	F	8	LEU	2.2
1	D	155	GLU	2.2
1	E	153	PRO	2.1
1	E	147	ILE	2.1
1	E	14	ARG	2.1
1	F	14	ARG	2.1
1	C	36	ARG	2.1
1	D	12	ASN	2.0
1	D	143	VAL	2.0
1	D	38	ILE	2.0

6.2 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.