



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 10, 2023 – 04:33 AM EDT

PDB ID : 6XRI
Title : MSMEG_2027 domain-swapped dimer
Authors : Jackson, C.J.; Antony, J.P.; Ahmed, F.H.
Deposited on : 2020-07-13
Resolution : 2.37 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 ⓘ 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 : 2.35.1
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) : 2.35.1

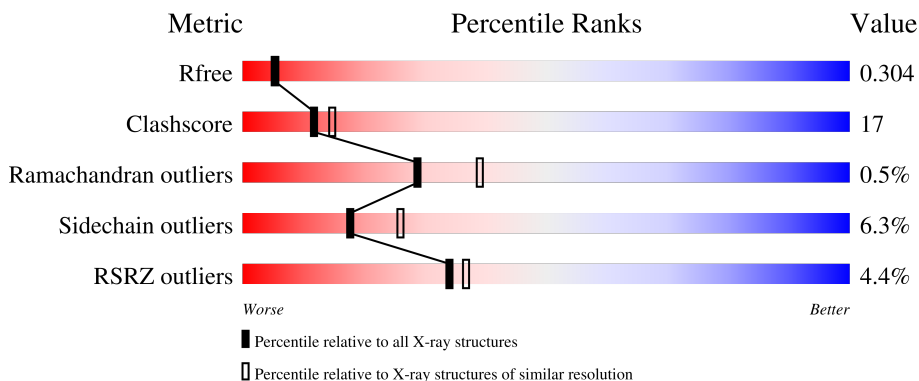
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.37 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	5509 (2.40-2.36)
Clashscore	141614	6082 (2.40-2.36)
Ramachandran outliers	138981	5973 (2.40-2.36)
Sidechain outliers	138945	5975 (2.40-2.36)
RSRZ outliers	127900	5397 (2.40-2.36)

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 $\leq 5\%$. 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	140	 76% 20% . .
1	B	140	 5% 72% 24% . .
1	C	140	 4% 69% 26% . .
1	D	140	 4% 66% 29% . .
1	E	140	 10% 64% 29% . .

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	140	 76% 19% . .
1	G	140	 72% 22% . .
1	H	140	 2% 73% 23% .
1	I	140	 58% 38% . .
1	J	140	 64% 30% . .
1	K	140	 10% 72% 21% . .
1	L	140	 6% 64% 25% . 8%
1	M	140	 4% 66% 26% . .
1	N	140	 14% 55% 34% 5% 6%

2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 15752 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 Uncharacterized protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	135	1093	695	189	207	2	0	0	0
1	B	135	1092	695	189	206	2	0	0	0
1	C	134	1093	694	189	208	2	0	1	0
1	D	135	1092	695	189	206	2	0	0	0
1	E	134	1084	689	188	205	2	0	0	0
1	F	134	1085	689	188	206	2	0	0	0
1	G	135	1092	695	189	206	2	0	0	0
1	H	134	1084	689	188	205	2	0	0	0
1	I	135	1107	706	191	208	2	0	1	0
1	J	134	1084	689	188	205	2	0	0	0
1	K	135	1092	695	189	206	2	0	0	0
1	L	129	1042	662	182	196	2	0	0	0
1	M	135	1092	695	189	206	2	0	0	0
1	N	132	1071	681	186	202	2	0	0	0


- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	57	Total 57	O 57	0	0
2	B	31	Total 31	O 31	0	0
2	C	32	Total 32	O 32	0	0
2	D	24	Total 24	O 24	0	0
2	E	35	Total 35	O 35	0	0
2	F	56	Total 56	O 56	0	0
2	G	54	Total 54	O 54	0	0
2	H	57	Total 57	O 57	0	0
2	I	47	Total 47	O 47	0	0
2	J	51	Total 51	O 51	0	0
2	K	25	Total 25	O 25	0	0
2	L	20	Total 20	O 20	0	0
2	M	30	Total 30	O 30	0	0
2	N	30	Total 30	O 30	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.

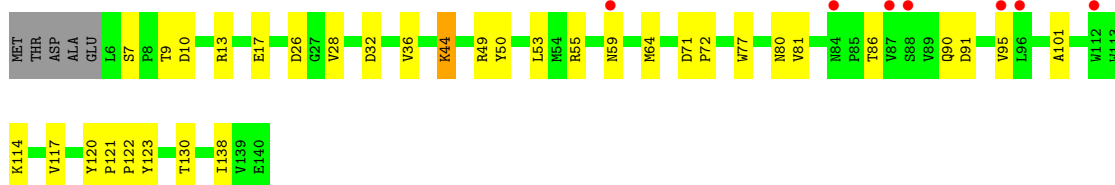
- Molecule 1: Uncharacterized protein

Chain A: 



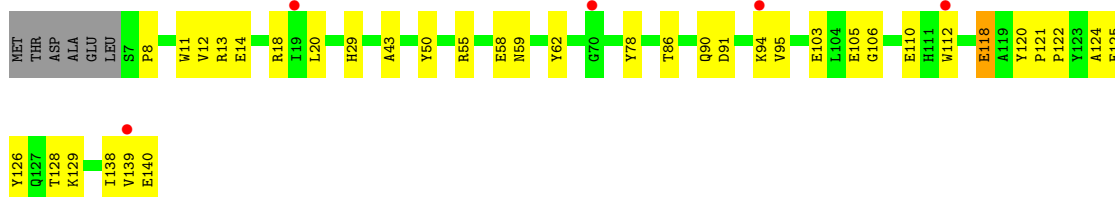
- Molecule 1: Uncharacterized protein

Chain B: 



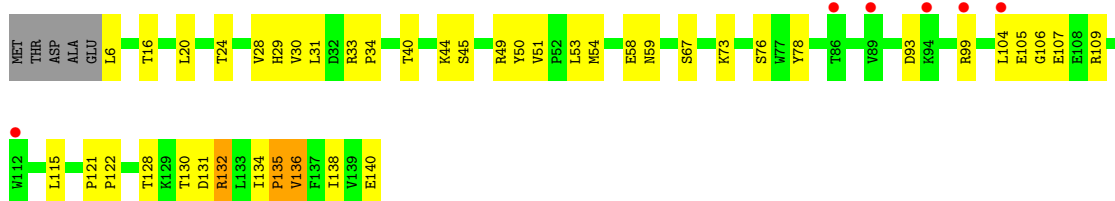
- Molecule 1: Uncharacterized protein

Chain C: 

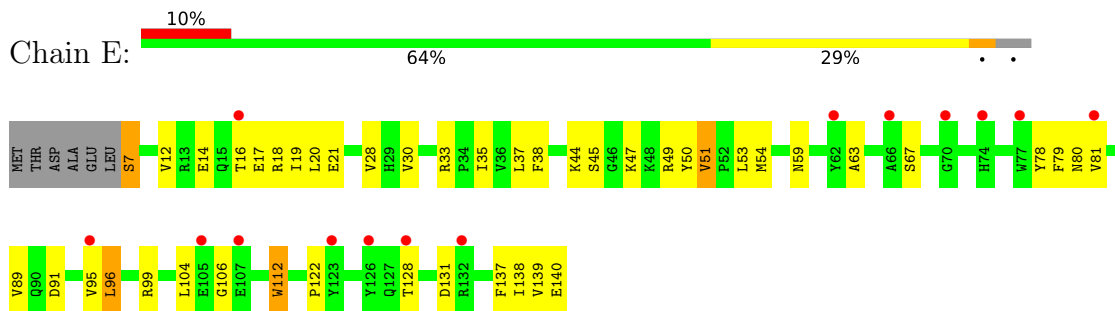


- Molecule 1: Uncharacterized protein

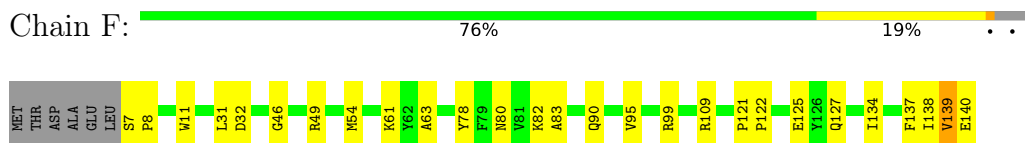
Chain D: 



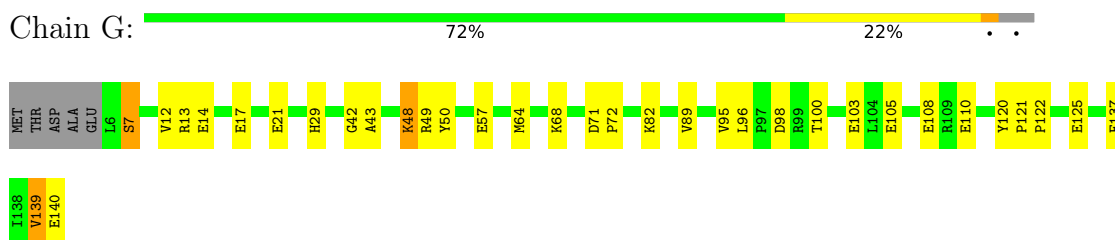
- Molecule 1: Uncharacterized protein



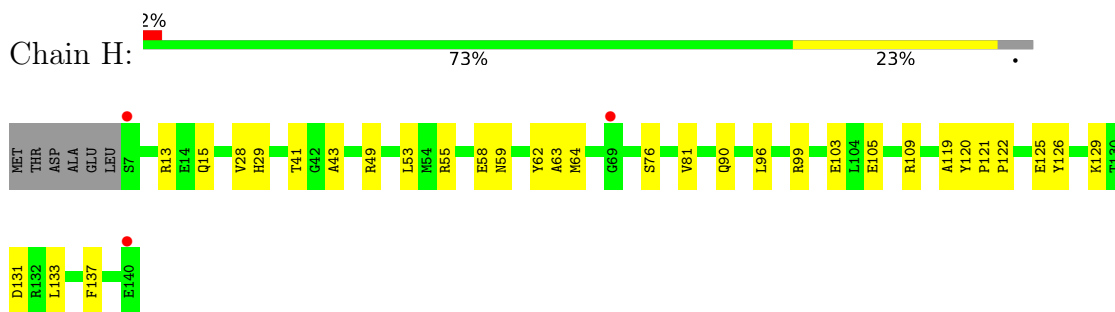
- Molecule 1: Uncharacterized protein



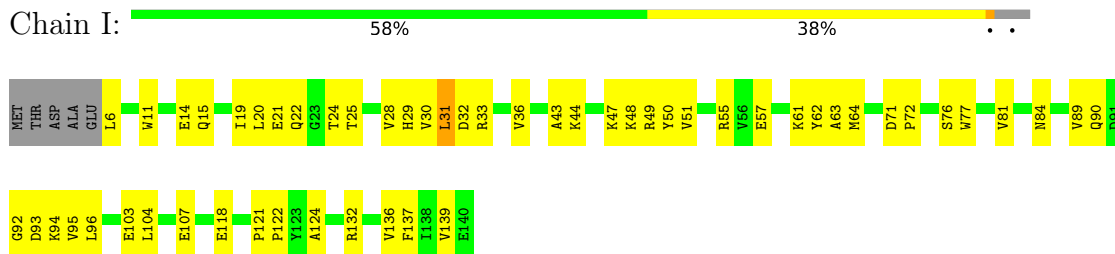
- Molecule 1: Uncharacterized protein



- Molecule 1: Uncharacterized protein

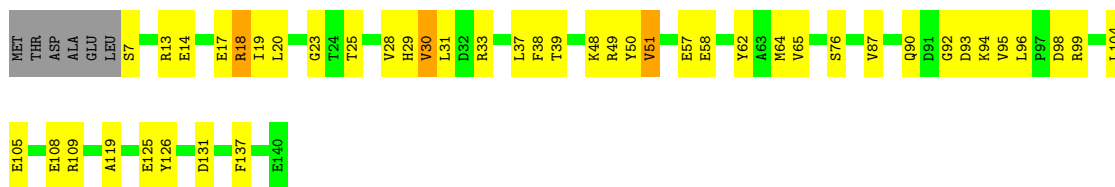


- Molecule 1: Uncharacterized protein



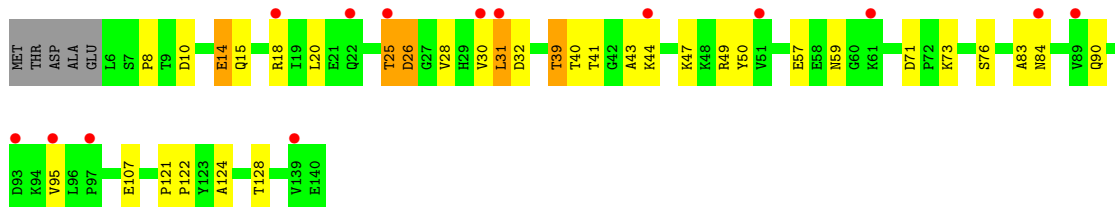
- Molecule 1: Uncharacterized protein

Chain J:  64% 30%



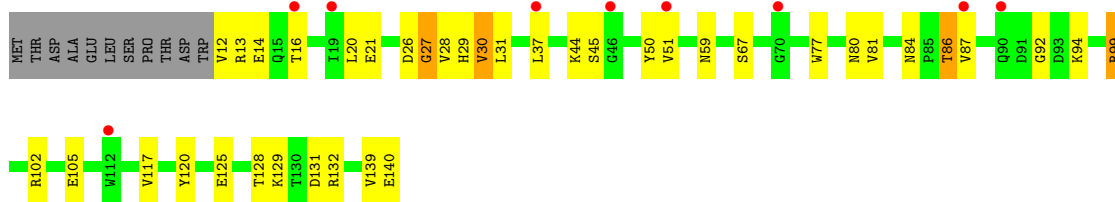
• Molecule 1: Uncharacterized protein

Chain K:  10% 72% 21%



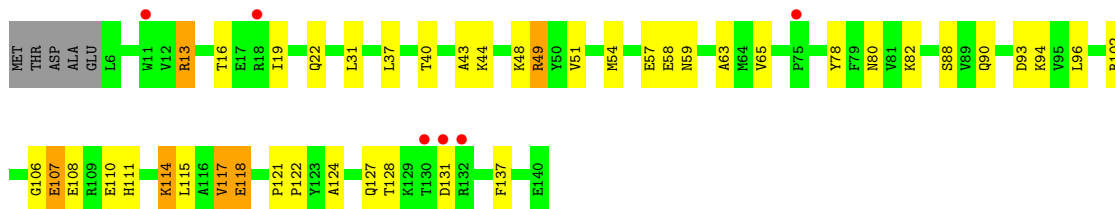
• Molecule 1: Uncharacterized protein

Chain L:  6% 64% 25% 8%



• Molecule 1: Uncharacterized protein

Chain M:  4% 66% 26%



• Molecule 1: Uncharacterized protein

Chain N:  14% 55% 34% 5% 6%





4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	73.55Å 110.23Å 140.15Å 90.00° 105.21° 90.00°	Depositor
Resolution (Å)	34.18 – 2.37 34.18 – 2.37	Depositor EDS
% Data completeness (in resolution range)	92.8 (34.18-2.37) 92.8 (34.18-2.37)	Depositor EDS
R_{merge}	0.20	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.54 (at 2.36Å)	Xtrriage
Refinement program	PHENIX 1.18_3855	Depositor
R, R_{free}	0.242 , 0.301 0.246 , 0.304	Depositor DCC
R_{free} test set	4026 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å ²)	28.5	Xtrriage
Anisotropy	0.637	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.28 , 22.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.52$, $\langle L^2 \rangle = 0.35$	Xtrriage
Estimated twinning fraction	0.448 for h,-k,-h-l	Xtrriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	15752	wwPDB-VP
Average B, all atoms (Å ²)	56.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 62.53 % 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 1.1114e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality [i](#)

5.1 Standard geometry [i](#)

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	0.51	0/1122	0.64	1/1525 (0.1%)
1	B	0.37	0/1121	0.54	0/1525
1	C	0.41	0/1122	0.52	0/1526
1	D	0.36	0/1121	0.52	0/1525
1	E	0.37	0/1113	0.53	0/1514
1	F	0.50	0/1114	0.58	0/1514
1	G	0.49	0/1121	0.59	0/1525
1	H	0.45	0/1113	0.57	0/1514
1	I	0.46	0/1138	0.59	0/1548
1	J	0.44	0/1113	0.57	0/1514
1	K	0.38	0/1121	0.52	0/1525
1	L	0.39	0/1068	0.52	0/1450
1	M	0.45	0/1121	0.58	0/1525
1	N	0.35	0/1099	0.50	0/1494
All	All	0.43	0/15607	0.56	1/21224 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	30	VAL	O-C-N	-5.90	113.27	122.70

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	1093	0	1074	32	0
1	B	1092	0	1074	33	1
1	C	1093	0	1068	35	0
1	D	1092	0	1074	44	0
1	E	1084	0	1063	48	1
1	F	1085	0	1063	36	0
1	G	1092	0	1074	31	0
1	H	1084	0	1063	29	0
1	I	1107	0	1083	60	1
1	J	1084	0	1063	48	0
1	K	1092	0	1074	37	0
1	L	1042	0	1030	36	1
1	M	1092	0	1074	45	0
1	N	1071	0	1051	54	0
2	A	57	0	0	26	0
2	B	31	0	0	20	0
2	C	32	0	0	20	0
2	D	24	0	0	27	0
2	E	35	0	0	21	0
2	F	56	0	0	23	0
2	G	54	0	0	19	0
2	H	57	0	0	25	0
2	I	47	0	0	32	0
2	J	51	0	0	32	0
2	K	25	0	0	21	0
2	L	20	0	0	14	0
2	M	30	0	0	19	0
2	N	30	0	0	23	0
All	All	15752	0	14928	525	2

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 17.

The worst 5 of 525 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:I:104:LEU:HG	2:I:205:HOH:O	1.35	1.26
1:D:6:LEU:N	2:D:202:HOH:O	1.87	1.07
2:A:256:HOH:O	1:F:54:MET:SD	2.11	1.07
1:K:49:ARG:NH2	2:K:202:HOH:O	1.88	1.05
1:F:109:ARG:NH2	2:F:202:HOH:O	1.89	1.05

All (2) 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:I:25:THR:O	1:L:59:ASN:ND2[1_655]	2.07	0.13
1:B:59:ASN:ND2	1:E:95:VAL:O[1_455]	2.17	0.03

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	133/140 (95%)	129 (97%)	4 (3%)	0	100	100
1	B	133/140 (95%)	131 (98%)	2 (2%)	0	100	100
1	C	133/140 (95%)	131 (98%)	2 (2%)	0	100	100
1	D	133/140 (95%)	126 (95%)	4 (3%)	3 (2%)	6	6
1	E	132/140 (94%)	126 (96%)	5 (4%)	1 (1%)	19	27
1	F	132/140 (94%)	131 (99%)	1 (1%)	0	100	100
1	G	133/140 (95%)	132 (99%)	1 (1%)	0	100	100
1	H	132/140 (94%)	128 (97%)	4 (3%)	0	100	100
1	I	134/140 (96%)	126 (94%)	8 (6%)	0	100	100
1	J	132/140 (94%)	126 (96%)	6 (4%)	0	100	100
1	K	133/140 (95%)	127 (96%)	6 (4%)	0	100	100
1	L	127/140 (91%)	123 (97%)	1 (1%)	3 (2%)	6	5
1	M	133/140 (95%)	129 (97%)	3 (2%)	1 (1%)	19	27
1	N	130/140 (93%)	123 (95%)	6 (5%)	1 (1%)	19	27
All	All	1850/1960 (94%)	1788 (97%)	53 (3%)	9 (0%)	29	39

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	M	107	GLU
1	D	106	GLY
1	E	106	GLY
1	L	26	ASP
1	D	135	PRO

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	118/122 (97%)	114 (97%)	4 (3%)	37 53
1	B	118/122 (97%)	112 (95%)	6 (5%)	24 36
1	C	118/122 (97%)	112 (95%)	6 (5%)	24 36
1	D	118/122 (97%)	112 (95%)	6 (5%)	24 36
1	E	117/122 (96%)	108 (92%)	9 (8%)	13 18
1	F	117/122 (96%)	115 (98%)	2 (2%)	60 76
1	G	118/122 (97%)	109 (92%)	9 (8%)	13 18
1	H	117/122 (96%)	111 (95%)	6 (5%)	24 36
1	I	119/122 (98%)	112 (94%)	7 (6%)	19 29
1	J	117/122 (96%)	110 (94%)	7 (6%)	19 28
1	K	118/122 (97%)	108 (92%)	10 (8%)	10 14
1	L	112/122 (92%)	100 (89%)	12 (11%)	6 8
1	M	118/122 (97%)	110 (93%)	8 (7%)	16 23
1	N	115/122 (94%)	102 (89%)	13 (11%)	6 7
All	All	1640/1708 (96%)	1535 (94%)	105 (6%)	18 25

5 of 105 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	J	51	VAL
1	K	57	GLU
1	N	105	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	J	94	LYS
1	K	25	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	F	127	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 [i](#)

6.1 Protein, DNA and RNA chains [i](#)

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	135/140 (96%)	-0.22	0 100 100	9, 35, 60, 74	0
1	B	135/140 (96%)	0.24	7 (5%) 27 30	40, 59, 85, 99	0
1	C	134/140 (95%)	0.18	5 (3%) 41 45	37, 58, 76, 93	0
1	D	135/140 (96%)	0.43	6 (4%) 34 37	59, 74, 96, 116	0
1	E	134/140 (95%)	0.54	14 (10%) 6 7	56, 75, 100, 112	0
1	F	134/140 (95%)	-0.28	0 100 100	13, 30, 56, 80	0
1	G	135/140 (96%)	-0.18	0 100 100	12, 36, 65, 79	0
1	H	134/140 (95%)	-0.06	3 (2%) 62 63	26, 40, 75, 96	0
1	I	135/140 (96%)	-0.10	0 100 100	24, 42, 65, 78	0
1	J	134/140 (95%)	-0.02	0 100 100	26, 46, 77, 105	0
1	K	135/140 (96%)	0.57	14 (10%) 6 7	46, 74, 104, 119	0
1	L	129/140 (92%)	0.51	9 (6%) 16 17	54, 74, 105, 109	0
1	M	135/140 (96%)	0.24	6 (4%) 34 37	32, 54, 85, 104	0
1	N	132/140 (94%)	0.89	19 (14%) 2 2	51, 71, 108, 122	0
All	All	1876/1960 (95%)	0.19	83 (4%) 34 37	9, 57, 93, 122	0

The worst 5 of 83 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	N	139	VAL	6.2
1	N	112	TRP	5.8
1	M	132	ARG	4.5
1	N	102	ARG	4.3
1	N	135	PRO	4.2

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.