wwPDB/EMDataBank EM Map/Model Validation Summary Report (i)

Aug 7, 2019 – 01:51 PM EDT

PDB ID	:	6RAO
EMDB ID:	:	EMD-4782
Title	:	Cryo-EM structure of the anti-feeding prophage (AFP) baseplate, 6-fold symmetrised
Authors	:	Desfosses, A.
Deposited on	:	2019-04-06
Resolution	:	3.10 Å(reported)

This is a wwPDB/EMDataBank EM Map/Model Validation Summary Report for a publicly released PDB/EMDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/EMValidationReportHelp
with specific help available everywhere you see the (i) symbol.

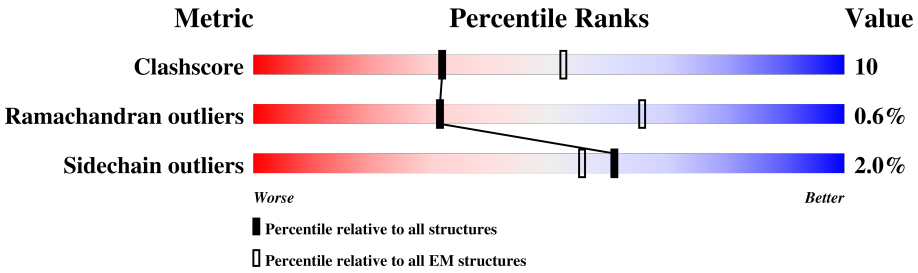
MolProbity	:	4.02b-467
Percentile statistics	:	20171227.v01 (using entries in the PDB archive December 27th 2017)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.4

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

The reported resolution of this entry is 3.10 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.



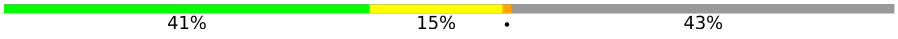
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	136327	1886
Ramachandran outliers	132723	1663
Sidechain outliers	132532	1531

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. 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%

Mol	Chain	Length	Quality of chain
1	A	149	89% 10% •
1	B	149	85% 14% •
2	C	354	85% 14% ••
3	D	451	71% 8% 21%
4	F	149	84% 15% •
5	H	140	64% 22% • 11%
6	E	417	72% 20% • 7%
7	G	229	79% 17% ••
8	I	607	59% 13% • 27%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	J	963	

2 Entry composition (i)

There are 9 unique types of molecules in this entry. The entry contains 22409 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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 Afp1.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	148	Total	C	N	O	S	0	0
			1148	729	187	228	4		
1	B	148	Total	C	N	O	S	0	0
			1148	729	187	228	4		

- Molecule 2 is a protein called Afp2.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	C	351	Total	C	N	O	S	0	0
			2718	1740	450	522	6		

- Molecule 3 is a protein called Afp3.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	D	358	Total	C	N	O	S	0	0
			2747	1753	462	525	7		

- Molecule 4 is a protein called Afp5.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	F	149	Total	C	N	O	S	0	0
			1196	757	218	215	6		

- Molecule 5 is a protein called Afp9.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	H	125	Total	C	N	O	S	0	0
			994	628	177	184	5		

- Molecule 6 is a protein called Afp4.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	E	386	Total	C	N	O	S	0	0
			3005	1920	512	561	12		

- Molecule 7 is a protein called Afp7.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	G	224	Total	C	N	O	S	0	0
			1737	1096	292	343	6		

- Molecule 8 is a protein called Afp11.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	I	443	Total	C	N	O	S	0	0
			3441	2189	592	648	12		

- Molecule 9 is a protein called Afp12.

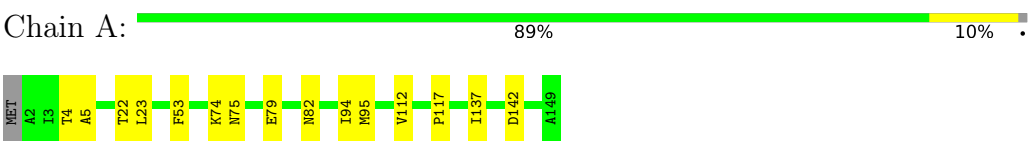
Mol	Chain	Residues	Atoms					AltConf	Trace
9	J	548	Total	C	N	O	S	0	0
			4275	2711	737	819	8		

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA 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. 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. 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.

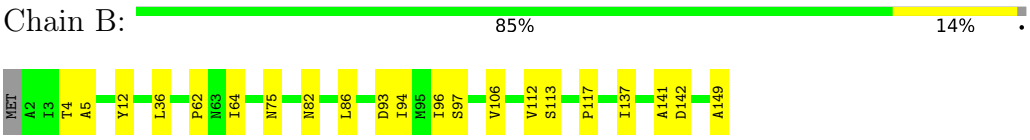
- Molecule 1: Afp1

Chain A:



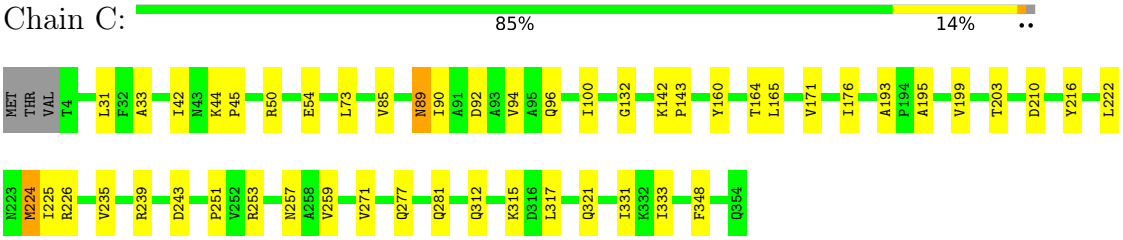
- Molecule 1: Afp1

Chain B:



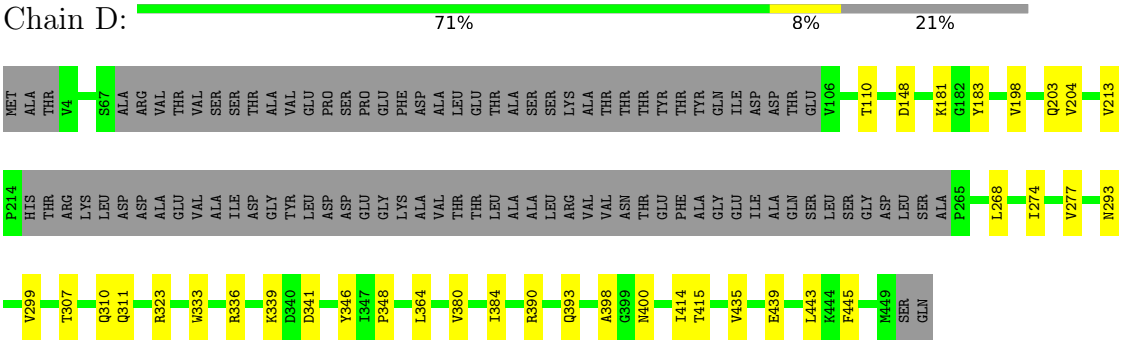
- Molecule 2: Afp2

Chain C:



- Molecule 3: Afp3

Chain D:

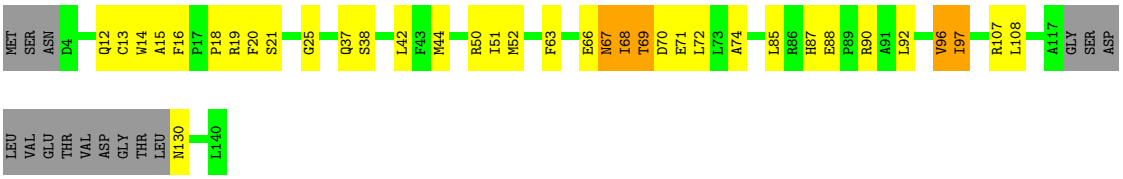


- Molecule 4: Afp5

Chain F:

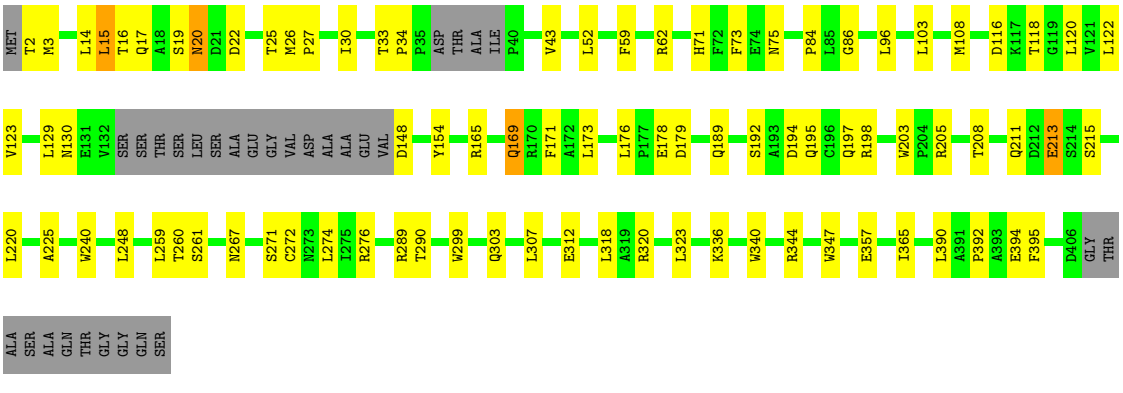
• Molecule 5: Afp9

Chain H:



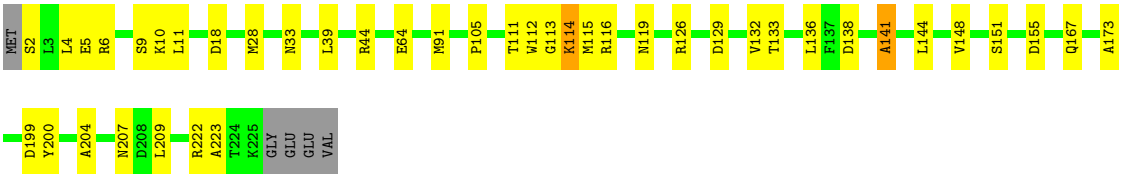
• Molecule 6: Afp4

Chain E:



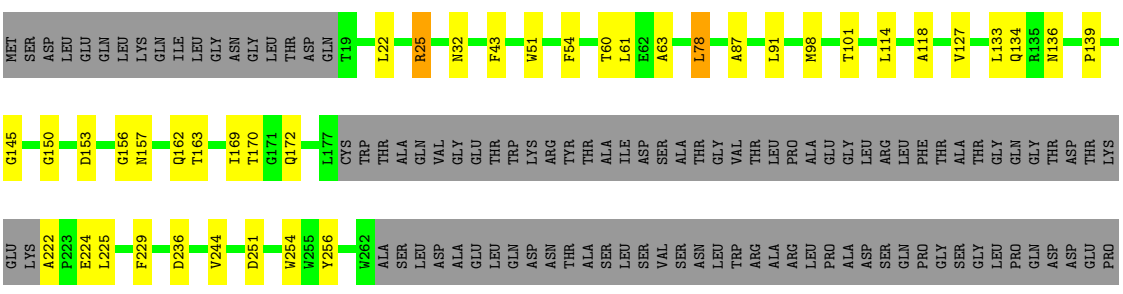
• Molecule 7: Afp7

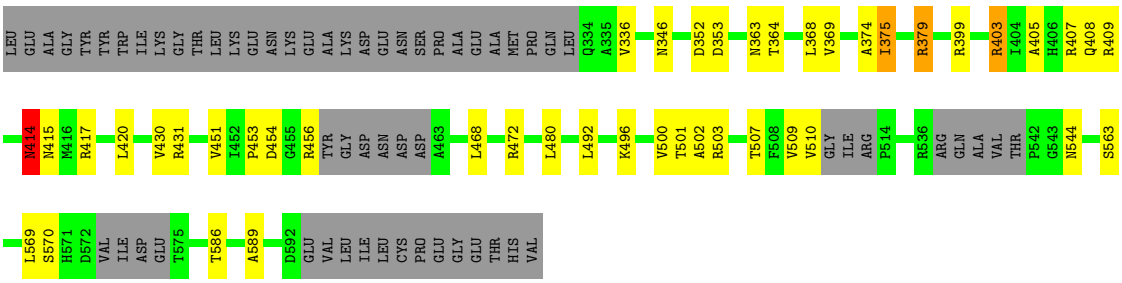
Chain G:



• Molecule 8: Afp11

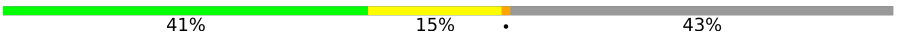
Chain I:

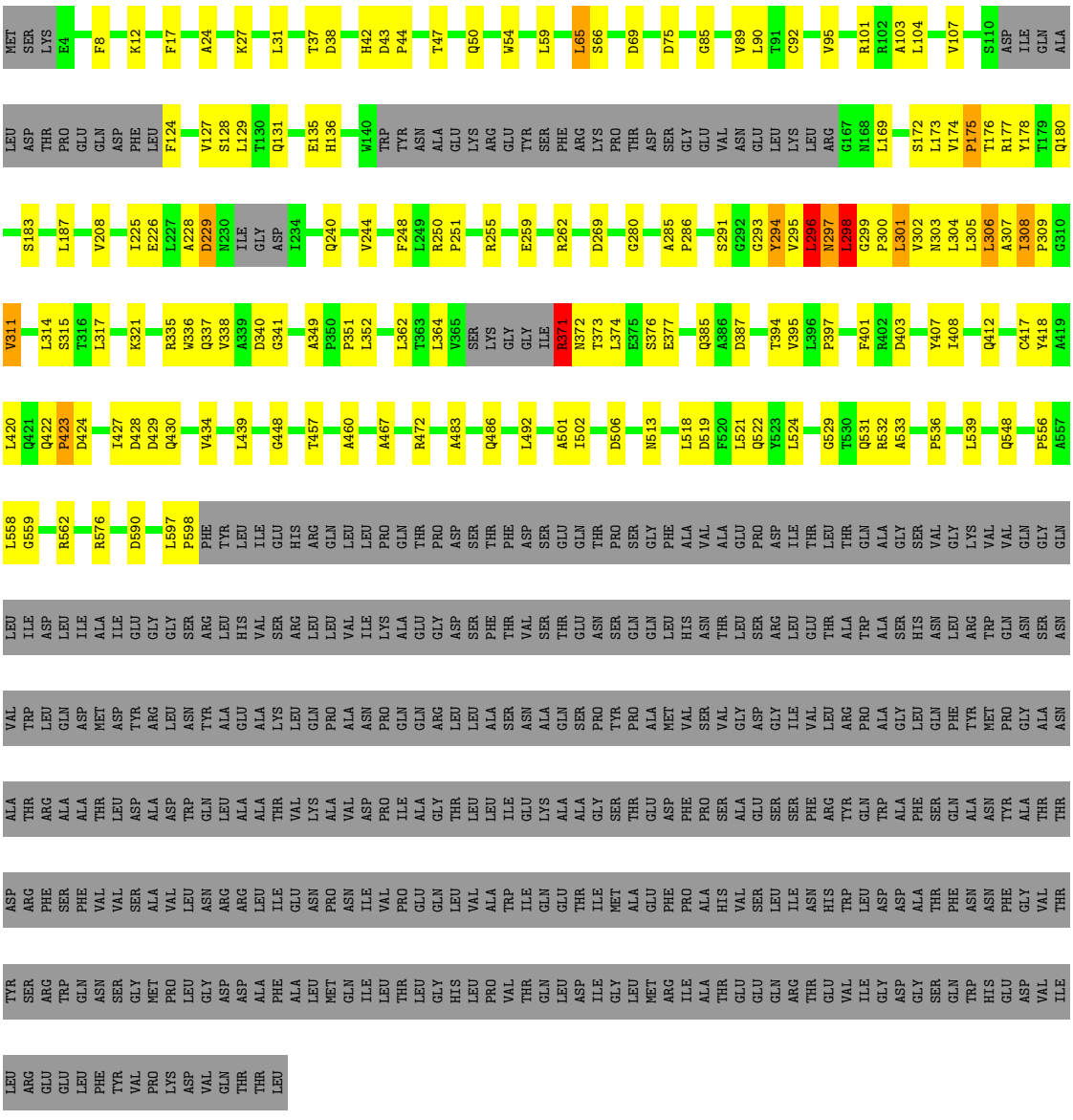




• Molecule 9: Afp12

Chain J:





4 Experimental information (i)

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C6	Depositor
Number of particles used	46991	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	20	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	Not provided	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor

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	#$\|Z\|>2$	RMSZ	#$\|Z\|>2$
1	A	0.34	0/1172	0.60	1/1601 (0.1%)
1	B	0.42	0/1172	0.65	2/1601 (0.1%)
2	C	0.40	0/2789	0.62	1/3812 (0.0%)
3	D	0.32	0/2809	0.56	0/3833
4	F	0.45	0/1220	0.66	1/1658 (0.1%)
5	H	0.36	0/1011	0.69	0/1366
6	E	0.38	0/3073	0.66	4/4195 (0.1%)
7	G	0.40	0/1770	0.75	4/2410 (0.2%)
8	I	0.41	0/3524	0.68	6/4815 (0.1%)
9	J	0.44	0/4370	0.71	5/5974 (0.1%)
All	All	0.40	0/22910	0.66	24/31265 (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
7	G	0	1
8	I	0	5
9	J	0	6
All	All	0	12

There are no bond length outliers.

The worst 5 of 24 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
9	J	423	PRO	CA-N-CD	-8.46	99.65	111.50
7	G	144	LEU	CA-CB-CG	8.25	134.27	115.30
1	B	86	LEU	CA-CB-CG	7.08	131.58	115.30
4	F	75	LEU	CA-CB-CG	6.62	130.51	115.30
8	I	133	LEU	CA-CB-CG	6.36	129.92	115.30

There are no chirality outliers.

5 of 12 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
7	G	141	ALA	Peptide
8	I	254	TRP	Peptide
8	I	353	ASP	Peptide
8	I	368	LEU	Peptide
8	I	375	ILE	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	1148	0	1123	7	0
1	B	1148	0	1123	12	0
2	C	2718	0	2630	35	0
3	D	2747	0	2721	21	0
4	F	1196	0	1206	12	0
5	H	994	0	998	76	0
6	E	3005	0	3004	83	0
7	G	1737	0	1720	26	0
8	I	3441	0	3352	59	0
9	J	4275	0	4189	231	0
All	All	22409	0	22066	459	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 10.

The worst 5 of 459 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)
5:H:21:SER:CA	8:I:43:PHE:HZ	1.00	1.58
5:H:50:ARG:HH22	9:J:47:THR:CG2	1.14	1.55
5:H:21:SER:CA	8:I:43:PHE:CZ	1.88	1.54
9:J:298:LEU:CB	9:J:304:LEU:HD12	1.04	1.50
5:H:21:SER:HA	8:I:43:PHE:CZ	1.44	1.48

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 PDB entries followed by that with respect to all EM entries.

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	146/149 (98%)	136 (93%)	8 (6%)	2 (1%)	12 44
1	B	146/149 (98%)	133 (91%)	11 (8%)	2 (1%)	12 44
2	C	349/354 (99%)	314 (90%)	34 (10%)	1 (0%)	43 77
3	D	352/451 (78%)	328 (93%)	24 (7%)	0	100 100
4	F	147/149 (99%)	132 (90%)	15 (10%)	0	100 100
5	H	121/140 (86%)	109 (90%)	8 (7%)	4 (3%)	4 24
6	E	380/417 (91%)	338 (89%)	41 (11%)	1 (0%)	43 77
7	G	222/229 (97%)	187 (84%)	33 (15%)	2 (1%)	19 56
8	I	429/607 (71%)	365 (85%)	63 (15%)	1 (0%)	49 81
9	J	538/963 (56%)	438 (81%)	97 (18%)	3 (1%)	27 64
All	All	2830/3608 (78%)	2480 (88%)	334 (12%)	16 (1%)	31 64

5 of 16 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
5	H	68	ILE
5	H	69	THR
9	J	296	LEU
9	J	395	VAL
1	A	5	ALA

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 PDB entries followed by that with respect to all EM entries.

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	127/128 (99%)	125 (98%)	2 (2%)	65	87
1	B	127/128 (99%)	125 (98%)	2 (2%)	65	87
2	C	284/287 (99%)	281 (99%)	3 (1%)	76	90
3	D	293/368 (80%)	292 (100%)	1 (0%)	93	97
4	F	134/134 (100%)	132 (98%)	2 (2%)	67	87
5	H	108/121 (89%)	104 (96%)	4 (4%)	37	71
6	E	324/345 (94%)	316 (98%)	8 (2%)	50	80
7	G	191/195 (98%)	187 (98%)	4 (2%)	56	82
8	I	364/509 (72%)	356 (98%)	8 (2%)	55	81
9	J	455/807 (56%)	442 (97%)	13 (3%)	45	76
All	All	2407/3022 (80%)	2360 (98%)	47 (2%)	61	83

5 of 47 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
6	E	267	ASN
8	I	25	ARG
9	J	308	ILE
7	G	114	LYS
8	I	32	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 43 such sidechains are listed below:

Mol	Chain	Res	Type
7	G	33	ASN
8	I	134	GLN
9	J	522	GLN
7	G	149	GLN
7	G	167	GLN

5.3.3 RNA

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 carbohydrates 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.