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

Aug 7, 2019 – 04:52 PM EDT

PDB ID	:	6RAP
EMDB ID:	:	EMD-4784
Title	:	Cryo-EM structure of the anti-feeding prophage cap (AFP tube terminating cap)
Authors	:	Desfosses, A.
Deposited on	:	2019-04-07
Resolution	:	3.30 Å(reported)

This is a Full wwPDB/EMDataBank EM Map/Model Validation 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.

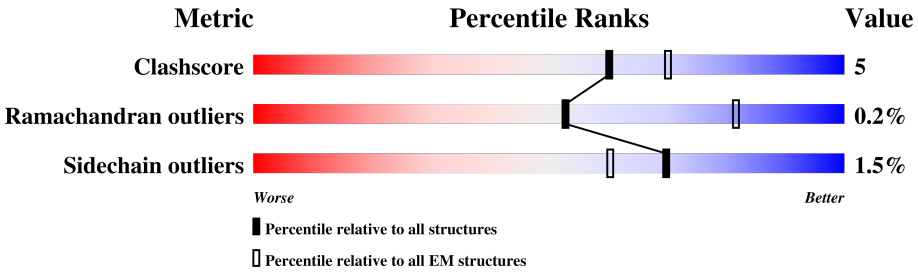
MolProbity	:	4.02b-467
Percentile statistics	:	20171227.v01 (using entries in the PDB archive December 27th 2017)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.4

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 3.30 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	83% 16% .
1	B	149	89% 10% .
2	C	354	91% 7% ..
3	D	451	72% 7% 21%
4	E	295	73% 20% 7%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 9879 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	B	148	Total	C	N	O	S	0	0
			1148	729	187	228	4		
1	A	149	Total	C	N	O	S	0	0
			1156	734	188	229	5		

- Molecule 2 is a protein called Afp2.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	C	349	Total	C	N	O	S	0	0
			2704	1732	447	519	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 Afp16.

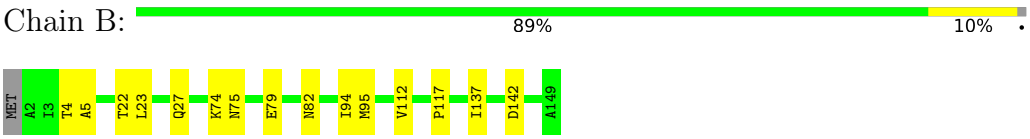
Mol	Chain	Residues	Atoms					AltConf	Trace
4	E	275	Total	C	N	O	S	0	0
			2124	1334	362	417	11		

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.

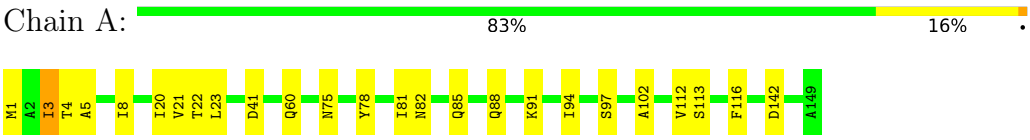
- Molecule 1: Afp1

Chain B:



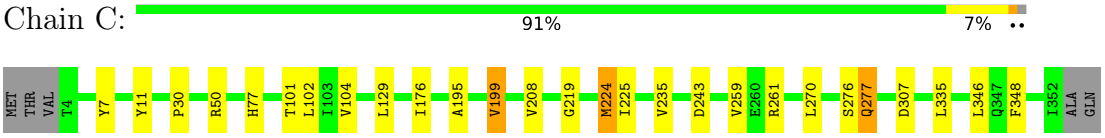
- Molecule 1: Afp1

Chain A:



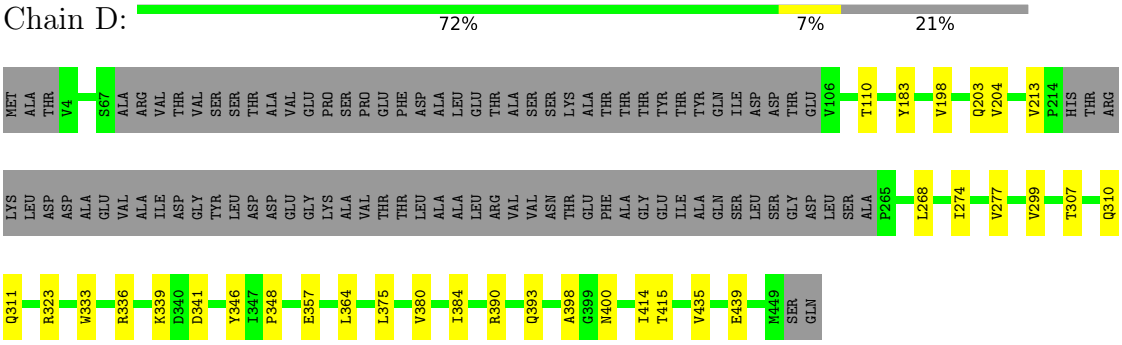
- Molecule 2: Afp2

Chain C:



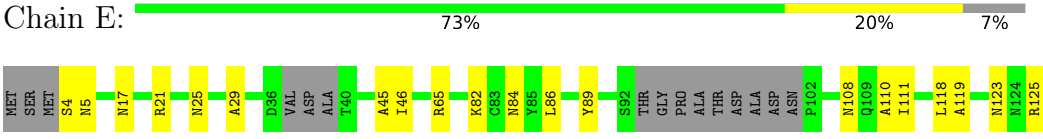
- Molecule 3: Afp3

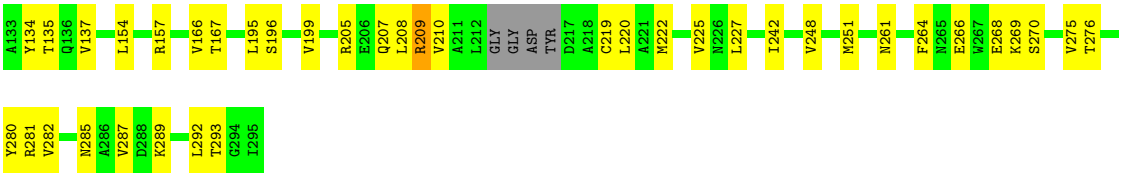
Chain D:



- Molecule 4: Afp16

Chain E:




A133
Y134
T135
Q136
V137
L154
R157
V166
T167
L195
S196
V199
R205
E206
Q207
L208
R209
V210
A211
L212
GLY
GLY
ASP
TYR
D217
A218
C219
L220
A221
M222
V225
W226
L227
I242
V248
M251
M261
F264
W265
E266
W267
E268
K269
S270
V275
T276
Y280
R281
V282
N285
A286
V287
D288
K289
L292
T293
G294
I295

4 Experimental information (i)

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C6	Depositor
Number of particles used	23797	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	0.46	0/1180	0.66	0/1611
1	B	0.34	0/1172	0.60	1/1601 (0.1%)
2	C	0.46	0/2775	0.66	1/3793 (0.0%)
3	D	0.32	0/2809	0.56	0/3833
4	E	0.38	0/2162	0.72	3/2948 (0.1%)
All	All	0.39	0/10098	0.64	5/13786 (0.0%)

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	E	227	LEU	CA-CB-CG	8.22	134.20	115.30
2	C	270	LEU	CA-CB-CG	6.29	129.76	115.30
4	E	86	LEU	CA-CB-CG	6.05	129.22	115.30
1	B	23	LEU	CA-CB-CG	5.19	127.23	115.30
4	E	220	LEU	CA-CB-CG	5.13	127.10	115.30

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	1156	0	1135	18	0
1	B	1148	0	1123	6	0
2	C	2704	0	2617	17	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	D	2747	0	2721	21	0
4	E	2124	0	2095	46	0
All	All	9879	0	9691	99	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 5.

All (99) 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)
4:E:266:GLU:O	4:E:270:SER:OG	1.68	1.12
4:E:261:ASN:OD1	4:E:289:LYS:CE	2.11	0.98
4:E:261:ASN:OD1	4:E:289:LYS:HE3	1.65	0.94
4:E:207:GLN:HG2	4:E:275:VAL:HB	1.55	0.86
4:E:276:THR:OG1	4:E:281:ARG:N	2.13	0.81
4:E:264:PHE:HB3	4:E:287:VAL:HG11	1.64	0.80
4:E:264:PHE:CG	4:E:287:VAL:HG11	2.16	0.80
4:E:207:GLN:CG	4:E:275:VAL:HB	2.15	0.76
4:E:261:ASN:OD1	4:E:289:LYS:NZ	2.22	0.71
4:E:276:THR:HB	4:E:281:ARG:HA	1.74	0.69
2:C:225:ILE:HG12	2:C:235:VAL:HG12	1.76	0.68
4:E:248:VAL:HG12	4:E:292:LEU:HD21	1.76	0.67
4:E:264:PHE:CB	4:E:287:VAL:HG11	2.24	0.67
1:B:74:LYS:HE2	1:B:79:GLU:HG3	1.77	0.67
4:E:208:LEU:HD23	4:E:208:LEU:O	1.98	0.63
2:C:208:VAL:HG11	2:C:224:MET:HG2	1.80	0.61
1:B:27:GLN:NE2	3:D:375:LEU:HD11	2.15	0.61
3:D:414:ILE:HG22	3:D:415:THR:HG23	1.81	0.61
1:A:94:ILE:HB	1:A:112:VAL:HB	1.82	0.61
4:E:293:THR:O	4:E:293:THR:HG22	2.01	0.61
4:E:248:VAL:CG1	4:E:292:LEU:HD21	2.32	0.59
2:C:348:PHE:CE2	3:D:357:GLU:HG2	2.38	0.59
1:B:94:ILE:HB	1:B:112:VAL:HB	1.83	0.59
2:C:346:LEU:HD21	3:D:364:LEU:HD12	1.84	0.58
1:A:85:GLN:HG2	1:A:88:GLN:HE21	1.69	0.58
1:A:1:MET:HB3	1:A:3:ILE:CG1	2.34	0.58
4:E:264:PHE:O	4:E:287:VAL:HG21	2.04	0.58
1:A:5:ALA:HB2	4:E:251:MET:CE	2.34	0.57
1:A:1:MET:HB3	1:A:3:ILE:HG12	1.87	0.57
1:B:22:THR:HB	1:B:95:MET:HB3	1.86	0.57
1:A:23:LEU:HA	1:A:94:ILE:HD13	1.89	0.55

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:142:ASP:N	1:A:142:ASP:OD1	2.40	0.55
4:E:82:LYS:NZ	4:E:84:ASN:OD1	2.40	0.54
3:D:213:VAL:HG22	3:D:299:VAL:HG12	1.90	0.54
1:A:5:ALA:HB2	4:E:251:MET:HE2	1.90	0.53
4:E:46:ILE:HG22	4:E:89:TYR:HD1	1.74	0.53
4:E:154:LEU:O	4:E:157:ARG:NH1	2.42	0.52
2:C:348:PHE:CE2	3:D:357:GLU:CG	2.93	0.52
2:C:259:VAL:HG11	2:C:335:LEU:HD21	1.91	0.51
4:E:205:ARG:NH2	4:E:222:MET:O	2.43	0.51
4:E:29:ALA:HB3	4:E:45:ALA:HA	1.92	0.51
2:C:104:VAL:HG12	2:C:129:LEU:HB2	1.92	0.51
1:A:91:LYS:HD3	1:A:116:PHE:HB3	1.93	0.51
3:D:390:ARG:HA	3:D:393:GLN:HG2	1.93	0.51
1:A:113:SER:OG	1:A:142:ASP:OD1	2.29	0.50
4:E:276:THR:CB	4:E:281:ARG:HA	2.40	0.50
1:B:117:PRO:HB3	1:B:137:ILE:HD11	1.94	0.49
1:A:102:ALA:O	4:E:65:ARG:NH1	2.44	0.49
4:E:268:GLU:HG3	4:E:285:ASN:O	2.13	0.49
3:D:311:GLN:HE22	3:D:333:TRP:HD1	1.59	0.49
3:D:307:THR:H	3:D:310:GLN:HE21	1.60	0.49
1:B:142:ASP:N	1:B:142:ASP:OD1	2.46	0.48
2:C:30:PRO:HA	2:C:102:LEU:HB2	1.95	0.48
3:D:183:TYR:O	3:D:203:GLN:NE2	2.42	0.48
3:D:198:VAL:HB	3:D:204:VAL:HG21	1.94	0.48
2:C:346:LEU:HD21	3:D:364:LEU:CD1	2.44	0.48
4:E:134:TYR:HB3	4:E:167:THR:HG22	1.96	0.48
4:E:275:VAL:HG13	4:E:275:VAL:O	2.13	0.47
4:E:4:SER:OG	4:E:5:ASN:N	2.48	0.47
1:A:20:ILE:HG23	1:A:97:SER:HB2	1.97	0.47
4:E:196:SER:HB2	4:E:199:VAL:HG23	1.98	0.46
4:E:275:VAL:O	4:E:282:VAL:HG12	2.15	0.46
3:D:311:GLN:OE1	3:D:323:ARG:NH1	2.47	0.46
2:C:195:ALA:HA	2:C:235:VAL:HG23	1.98	0.46
3:D:339:LYS:HG3	3:D:341:ASP:H	1.80	0.45
3:D:110:THR:HB	3:D:268:LEU:HD21	1.98	0.45
1:A:41:ASP:OD1	1:A:41:ASP:N	2.49	0.45
4:E:207:GLN:HG2	4:E:275:VAL:CB	2.37	0.45
1:A:3:ILE:HG13	1:A:3:ILE:H	1.62	0.45
4:E:207:GLN:O	4:E:210:VAL:HG12	2.15	0.45
3:D:364:LEU:HD22	3:D:380:VAL:HG13	1.99	0.45
2:C:276:SER:OG	2:C:277:GLN:N	2.49	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:E:119:ALA:O	4:E:123:ASN:ND2	2.50	0.45
4:E:289:LYS:O	4:E:289:LYS:HG3	2.17	0.45
1:A:41:ASP:HB3	1:A:60:GLN:HE21	1.83	0.44
4:E:108:ASN:HB2	4:E:111:ILE:HG22	1.99	0.43
3:D:274:ILE:HA	3:D:277:VAL:HG22	2.01	0.43
3:D:398:ALA:HB2	3:D:435:VAL:HB	2.00	0.43
3:D:346:TYR:HB3	3:D:348:PRO:HD2	2.00	0.43
4:E:21:ARG:NH1	4:E:25:ASN:O	2.53	0.42
3:D:364:LEU:HD11	3:D:384:ILE:HD11	2.02	0.42
1:A:23:LEU:HA	1:A:94:ILE:CD1	2.49	0.42
4:E:225:VAL:HG12	4:E:292:LEU:CD1	2.49	0.42
2:C:307:ASP:OD1	4:E:196:SER:OG	2.33	0.42
4:E:195:LEU:HD21	4:E:280:TYR:HD1	1.84	0.42
2:C:77:HIS:CE1	2:C:199:VAL:HG23	2.55	0.42
2:C:7:TYR:N	2:C:11:TYR:OH	2.54	0.41
4:E:269:LYS:HG3	4:E:269:LYS:O	2.19	0.41
4:E:108:ASN:HB3	4:E:110:ALA:H	1.86	0.41
4:E:209:ARG:NE	4:E:219:CYS:SG	2.94	0.41
2:C:219:GLY:N	2:C:243:ASP:OD2	2.40	0.41
1:A:78:TYR:HA	1:A:81:ILE:HG22	2.03	0.41
3:D:336:ARG:NH2	3:D:439:GLU:OE2	2.54	0.41
1:A:8:ILE:HA	1:A:8:ILE:HD13	1.94	0.41
4:E:135:THR:HG22	4:E:166:VAL:HG22	2.03	0.40
2:C:176:ILE:HG23	2:C:199:VAL:HG11	2.02	0.40
4:E:242:ILE:HG23	4:E:282:VAL:HG23	2.03	0.40
4:E:125:ARG:HG2	4:E:137:VAL:HG21	2.03	0.40
2:C:101:THR:HG22	2:C:261:ARG:HH12	1.86	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 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	147/149 (99%)	137 (93%)	10 (7%)	0	100	100
1	B	146/149 (98%)	136 (93%)	8 (6%)	2 (1%)	12	43
2	C	347/354 (98%)	308 (89%)	39 (11%)	0	100	100
3	D	352/451 (78%)	328 (93%)	24 (7%)	0	100	100
4	E	267/295 (90%)	225 (84%)	42 (16%)	0	100	100
All	All	1259/1398 (90%)	1134 (90%)	123 (10%)	2 (0%)	53	80

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	5	ALA
1	B	4	THR

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	128/128 (100%)	122 (95%)	6 (5%)	29	62
1	B	127/128 (99%)	125 (98%)	2 (2%)	65	83
2	C	283/287 (99%)	279 (99%)	4 (1%)	69	85
3	D	293/368 (80%)	292 (100%)	1 (0%)	93	97
4	E	235/249 (94%)	232 (99%)	3 (1%)	71	85
All	All	1066/1160 (92%)	1050 (98%)	16 (2%)	70	84

All (16) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	75	ASN
1	B	82	ASN
1	A	3	ILE
1	A	4	THR
1	A	21	VAL
1	A	22	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	75	ASN
1	A	82	ASN
2	C	50	ARG
2	C	199	VAL
2	C	224	MET
2	C	277	GLN
3	D	400	ASN
4	E	17	ASN
4	E	118	LEU
4	E	209	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (14) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	27	GLN
1	A	60	GLN
1	A	63	ASN
1	A	82	ASN
1	A	88	GLN
2	C	277	GLN
2	C	325	ASN
3	D	310	GLN
3	D	400	ASN
4	E	17	ASN
4	E	54	HIS
4	E	116	GLN
4	E	124	ASN
4	E	207	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 carbohydrates 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.