



wwPDB/EMDataBank EM Map/Model Validation Summary Report ⓘ

Aug 17, 2017 – 12:31 PM EDT

PDB ID : 2YEW
EMDB ID: : EMD-1886
Title : Modeling Barmah Forest virus structural proteins
Authors : Kostyuchenko, V.A.; Jakana, J.; Liu, X.; Haddow, A.D.; Aung, M.; Weaver, S.C.; Chiu, W.; Lok, S.M.
Deposited on : unknown
Resolution : 5.00 Å(reported)
Based on PDB ID : 2XFB, 2ALA, 1SVP

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
<http://wwpdb.org/validation/2016/EMValidationReportHelp>
with specific help available everywhere you see the ⓘ symbol.

MolProbity : 4.02b-467
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20029824

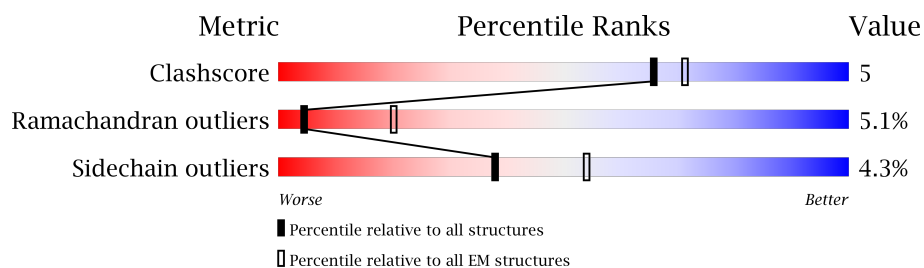
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 5.00 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	125131	1336
Ramachandran outliers	121729	1120
Sidechain outliers	121581	1026

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 $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	A	253	
1	D	253	
1	G	253	
1	J	253	
2	B	427	
2	E	427	
2	H	427	
2	K	427	
3	C	421	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	F	421	 74%22%••
3	I	421	 70%24%5%
3	L	421	 71%25%•

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 31120 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 CAPSID PROTEIN.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	171	Total	C	N	O	S	0	0
			1316	827	234	244	11		
1	D	171	Total	C	N	O	S	0	0
			1316	827	234	244	11		
1	G	171	Total	C	N	O	S	0	0
			1316	827	234	244	11		
1	J	171	Total	C	N	O	S	0	0
			1316	827	234	244	11		

- Molecule 2 is a protein called E1 ENVELOPE GLYCOPROTEIN.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	426	Total	C	N	O	S	0	1
			3225	2045	526	631	23		
2	E	426	Total	C	N	O	S	0	1
			3225	2045	526	631	23		
2	H	426	Total	C	N	O	S	0	1
			3225	2045	526	631	23		
2	K	426	Total	C	N	O	S	0	1
			3225	2045	526	631	23		

There are 52 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	?	-	HIS	deletion	UNP P89946
B	?	-	HIS	deletion	UNP P89946
B	?	-	HIS	deletion	UNP P89946
B	?	-	HIS	deletion	UNP P89946
B	?	-	HIS	deletion	UNP P89946
B	?	-	HIS	deletion	UNP P89946
B	?	-	HIS	deletion	UNP P89946
B	?	-	HIS	deletion	UNP P89946
B	?	-	HIS	deletion	UNP P89946

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	?	-	HIS	deletion	UNP P89946
B	?	-	HIS	deletion	UNP P89946
B	?	-	HIS	deletion	UNP P89946
B	?	-	HIS	deletion	UNP P89946
E	?	-	HIS	deletion	UNP P89946
E	?	-	HIS	deletion	UNP P89946
E	?	-	HIS	deletion	UNP P89946
E	?	-	HIS	deletion	UNP P89946
E	?	-	HIS	deletion	UNP P89946
E	?	-	HIS	deletion	UNP P89946
E	?	-	HIS	deletion	UNP P89946
E	?	-	HIS	deletion	UNP P89946
E	?	-	HIS	deletion	UNP P89946
E	?	-	HIS	deletion	UNP P89946
E	?	-	HIS	deletion	UNP P89946
E	?	-	HIS	deletion	UNP P89946
E	?	-	HIS	deletion	UNP P89946
E	?	-	HIS	deletion	UNP P89946
H	?	-	HIS	deletion	UNP P89946
H	?	-	HIS	deletion	UNP P89946
H	?	-	HIS	deletion	UNP P89946
H	?	-	HIS	deletion	UNP P89946
H	?	-	HIS	deletion	UNP P89946
H	?	-	HIS	deletion	UNP P89946
H	?	-	HIS	deletion	UNP P89946
H	?	-	HIS	deletion	UNP P89946
H	?	-	HIS	deletion	UNP P89946
H	?	-	HIS	deletion	UNP P89946
H	?	-	HIS	deletion	UNP P89946
H	?	-	HIS	deletion	UNP P89946
H	?	-	HIS	deletion	UNP P89946
H	?	-	HIS	deletion	UNP P89946
K	?	-	HIS	deletion	UNP P89946
K	?	-	HIS	deletion	UNP P89946
K	?	-	HIS	deletion	UNP P89946
K	?	-	HIS	deletion	UNP P89946
K	?	-	HIS	deletion	UNP P89946
K	?	-	HIS	deletion	UNP P89946
K	?	-	HIS	deletion	UNP P89946
K	?	-	HIS	deletion	UNP P89946
K	?	-	HIS	deletion	UNP P89946
K	?	-	HIS	deletion	UNP P89946
K	?	-	HIS	deletion	UNP P89946
K	?	-	HIS	deletion	UNP P89946
K	?	-	HIS	deletion	UNP P89946
K	?	-	HIS	deletion	UNP P89946
K	?	-	HIS	deletion	UNP P89946

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
K	?	-	HIS	deletion	UNP P89946

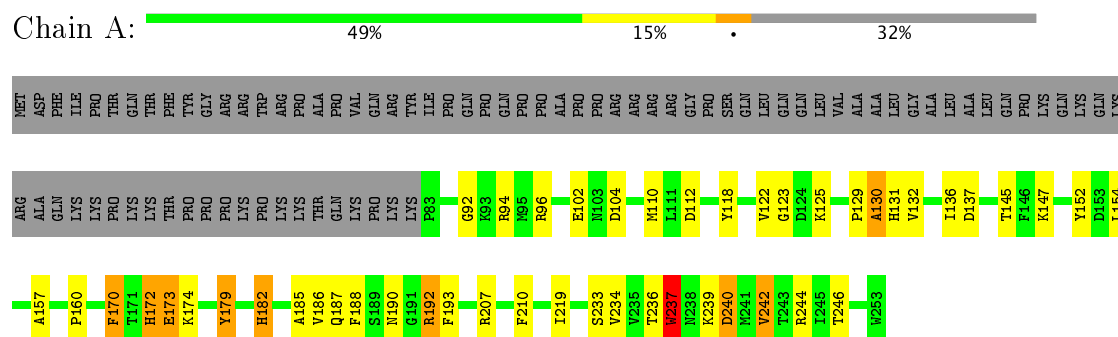
- Molecule 3 is a protein called E2 ENVELOPE GLYCOPROTEIN.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	C	421	Total	C	N	O	S	0	0
			3239	2054	552	610	23		
3	F	421	Total	C	N	O	S	0	0
			3239	2054	552	610	23		
3	I	421	Total	C	N	O	S	0	0
			3239	2054	552	610	23		
3	L	421	Total	C	N	O	S	0	0
			3239	2054	552	610	23		

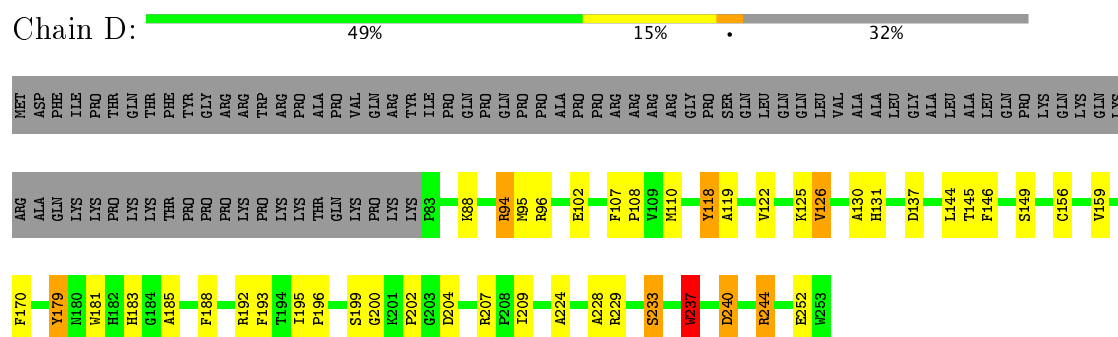
3 Residue-property plots

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.

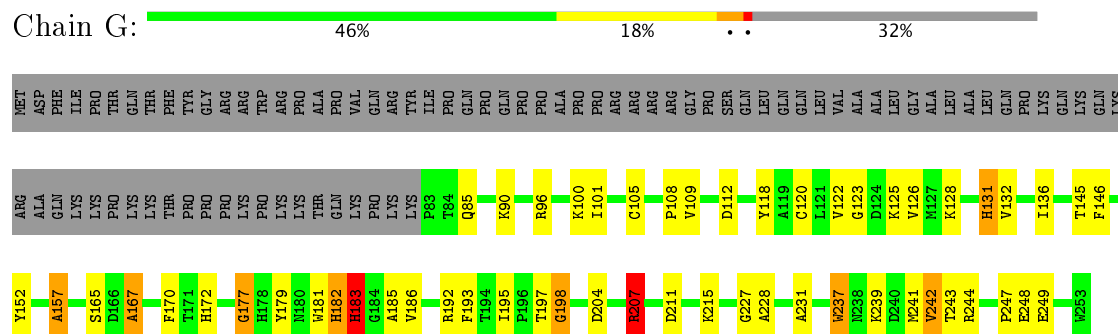
• Molecule 1: CAPSID PROTEIN



• Molecule 1: CAPSID PROTEIN

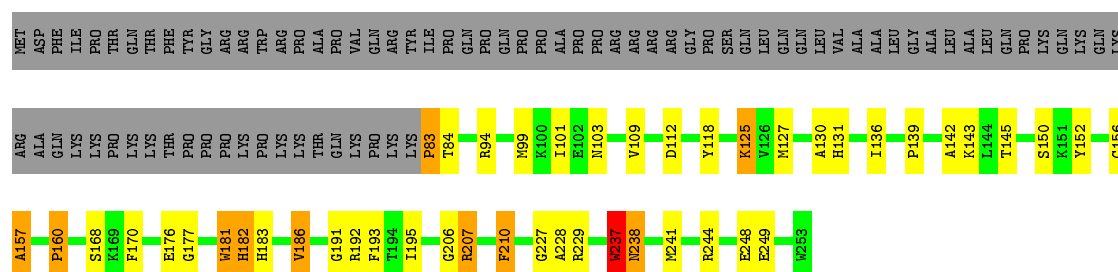


• Molecule 1: CAPSID PROTEIN



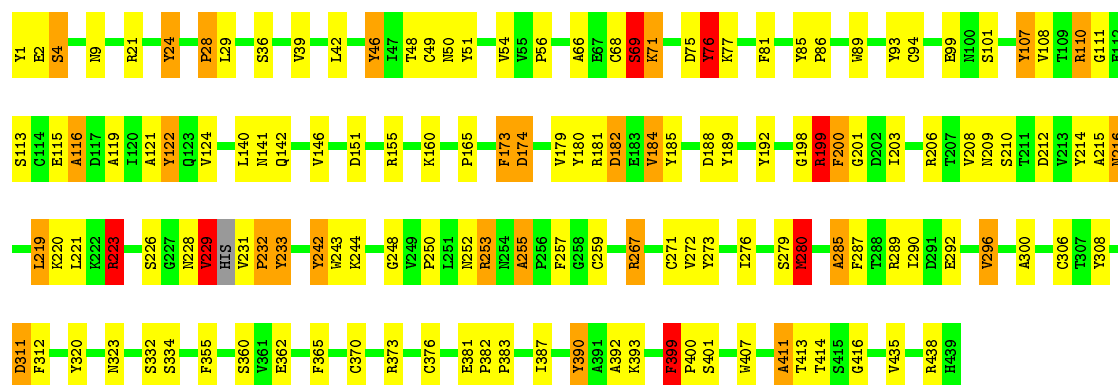
• Molecule 1: CAPSID PROTEIN





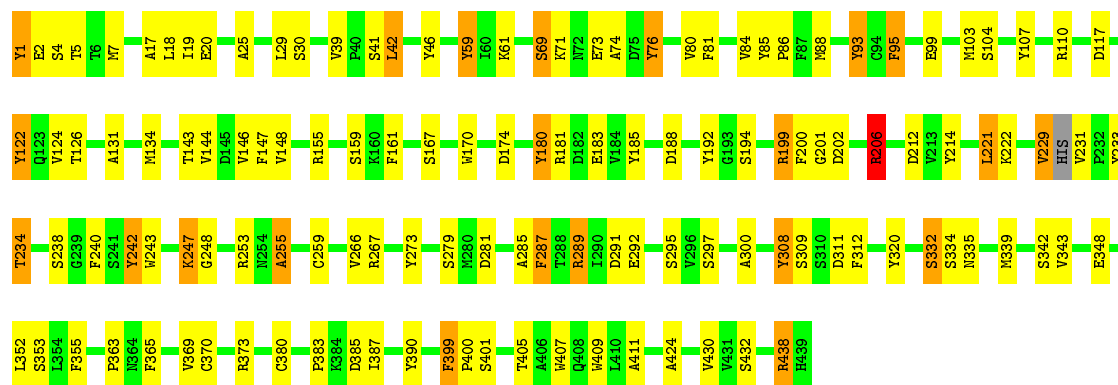
• Molecule 2: E1 ENVELOPE GLYCOPROTEIN

Chain B: 67% 25% 6% •



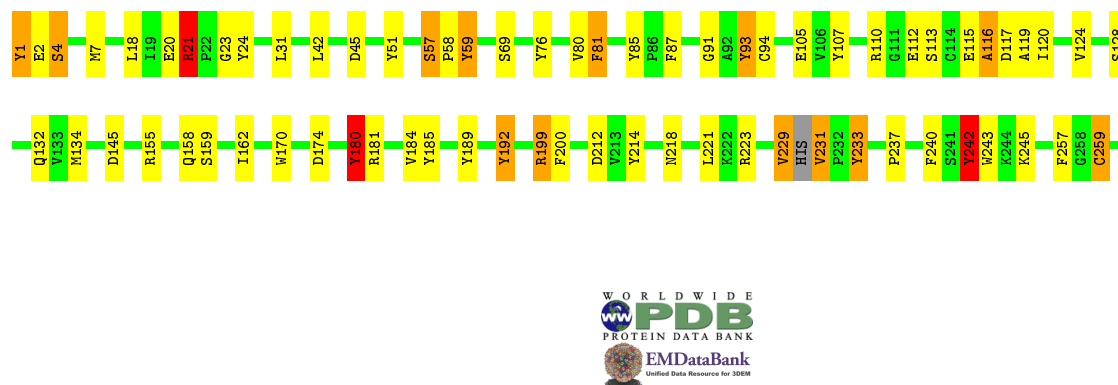
• Molecule 2: E1 ENVELOPE GLYCOPROTEIN

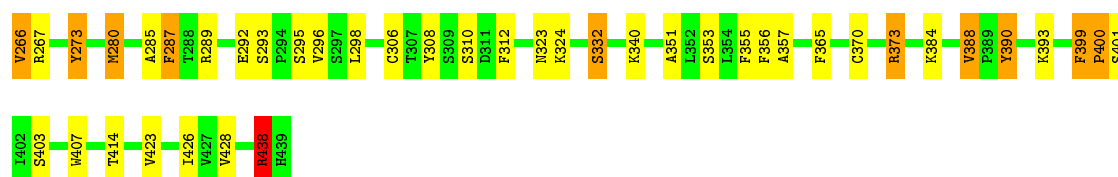
Chain E: 69% 25% 5% •



• Molecule 2: E1 ENVELOPE GLYCOPROTEIN

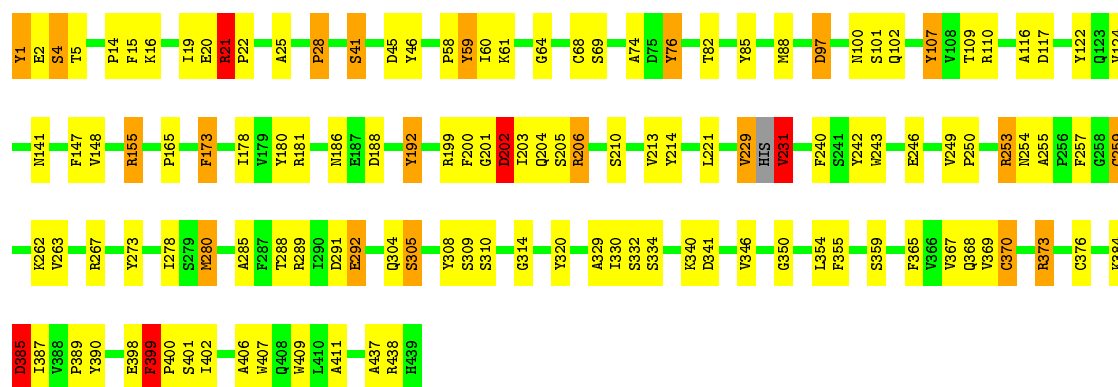
Chain H: 74% 20% 5% •





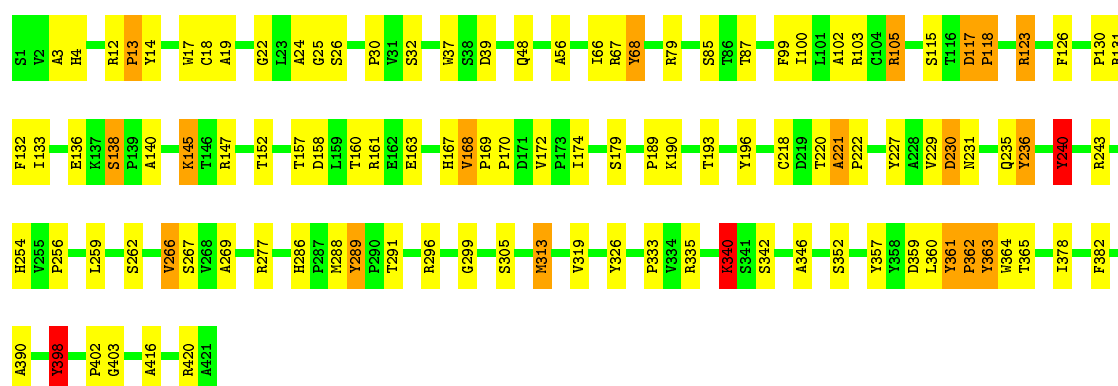
• Molecule 2: E1 ENVELOPE GLYCOPROTEIN

Chain K: 70% 24% 5%



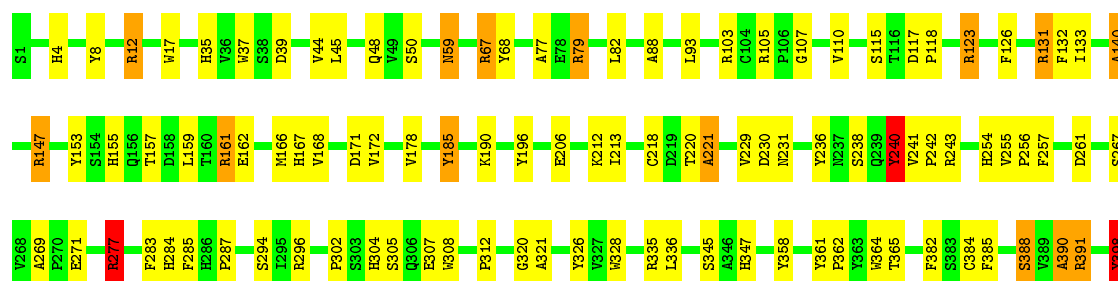
• Molecule 3: E2 ENVELOPE GLYCOPROTEIN

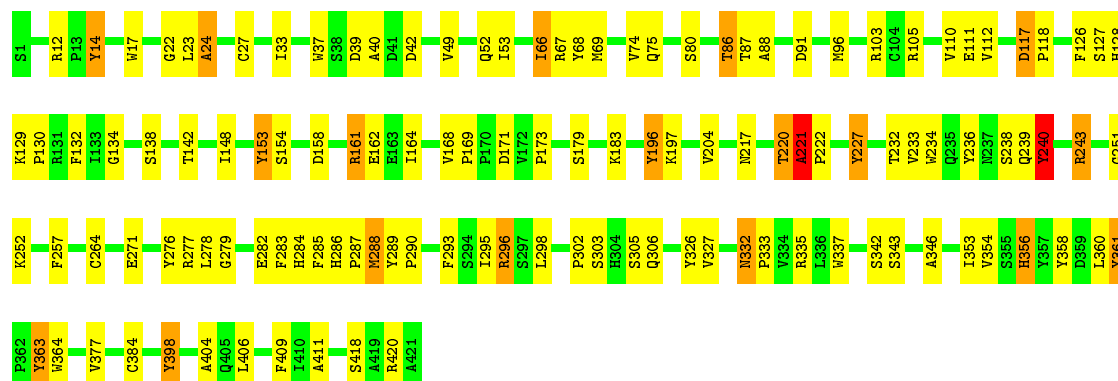
Chain C: 73% 22%



• Molecule 3: E2 ENVELOPE GLYCOPROTEIN

Chain F: 74% 22%





4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, I	Depositor
Number of particles used	5169	Depositor
Resolution determination method	Not provided	Depositor
CTF correction method	INDIVIDUAL PARTICLES	Depositor
Microscope	JEOL 3200FSC	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	20	Depositor
Minimum defocus (nm)	500	Depositor
Maximum defocus (nm)	3500	Depositor
Magnification	50000	Depositor
Image detector	GENERIC GATAN	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	3.19	16/1346 (1.2%)	2.07	35/1812 (1.9%)
1	D	3.19	11/1346 (0.8%)	2.09	43/1812 (2.4%)
1	G	3.16	17/1346 (1.3%)	2.03	33/1812 (1.8%)
1	J	3.19	18/1346 (1.3%)	2.09	31/1812 (1.7%)
2	B	1.75	29/3296 (0.9%)	2.05	102/4494 (2.3%)
2	E	1.75	28/3296 (0.8%)	1.98	84/4494 (1.9%)
2	H	1.74	26/3296 (0.8%)	2.05	93/4494 (2.1%)
2	K	1.78	29/3296 (0.9%)	1.99	83/4494 (1.8%)
3	C	2.79	40/3326 (1.2%)	1.94	75/4537 (1.7%)
3	F	2.71	44/3326 (1.3%)	1.99	82/4537 (1.8%)
3	I	2.80	48/3326 (1.4%)	2.00	80/4537 (1.8%)
3	L	2.77	35/3326 (1.1%)	2.01	95/4537 (2.1%)
All	All	2.49	341/31872 (1.1%)	2.01	836/43372 (1.9%)

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
1	A	0	4
1	D	0	7
1	G	0	8
1	J	0	4
2	B	0	20
2	E	0	21
2	H	0	19
2	K	0	20
3	C	0	11
3	F	0	6
3	I	0	11
3	L	0	9
All	All	0	140

The worst 5 of 341 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	237	TRP	CE2-CZ2	91.21	2.94	1.39
1	J	237	TRP	CE2-CZ2	88.72	2.90	1.39
1	G	237	TRP	CE2-CZ2	87.75	2.88	1.39
1	A	237	TRP	CE2-CZ2	87.70	2.88	1.39
3	I	398	TYR	CD2-CE2	78.02	2.56	1.39

The worst 5 of 836 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	L	105	ARG	NE-CZ-NH1	19.59	130.09	120.30
3	C	398	TYR	CB-CG-CD2	-18.30	110.02	121.00
1	G	237	TRP	NE1-CE2-CZ2	17.45	149.59	130.40
3	L	398	TYR	CG-CD2-CE2	-16.81	107.85	121.30
3	F	398	TYR	CG-CD2-CE2	-15.85	108.62	121.30

There are no chirality outliers.

5 of 140 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	125	LYS	Peptide
1	A	145	THR	Peptide
1	A	192	ARG	Sidechain
1	A	207	ARG	Sidechain
2	B	2	GLU	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	1316	0	1322	30	0
1	D	1316	0	1322	23	0
1	G	1316	0	1322	28	0
1	J	1316	0	1322	26	0
2	B	3225	0	3161	34	0
2	E	3225	0	3161	33	0
2	H	3225	0	3161	29	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	K	3225	0	3161	32	0
3	C	3239	0	3208	57	0
3	F	3239	0	3208	58	0
3	I	3239	0	3208	58	0
3	L	3239	0	3208	58	0
All	All	31120	0	30764	289	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.

The worst 5 of 289 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)
3:I:240:TYR:CD1	3:I:240:TYR:CE1	1.81	1.68
3:I:240:TYR:CD2	3:I:240:TYR:CE2	1.80	1.67
3:F:240:TYR:CE2	3:F:240:TYR:CD2	1.79	1.65
1:A:237:TRP:CD2	1:A:237:TRP:CE3	1.82	1.64
1:G:237:TRP:CD2	1:G:237:TRP:CE3	1.74	1.64

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	169/253 (67%)	136 (80%)	24 (14%)	9 (5%)	2	26
1	D	169/253 (67%)	148 (88%)	15 (9%)	6 (4%)	4	34
1	G	169/253 (67%)	133 (79%)	21 (12%)	15 (9%)	1	15
1	J	169/253 (67%)	132 (78%)	22 (13%)	15 (9%)	1	15
2	B	422/427 (99%)	354 (84%)	47 (11%)	21 (5%)	2	27

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	E	422/427 (99%)	352 (83%)	42 (10%)	28 (7%)	1	21
2	H	422/427 (99%)	362 (86%)	47 (11%)	13 (3%)	5	38
2	K	422/427 (99%)	366 (87%)	33 (8%)	23 (6%)	2	25
3	C	419/421 (100%)	359 (86%)	36 (9%)	24 (6%)	2	24
3	F	419/421 (100%)	365 (87%)	38 (9%)	16 (4%)	4	33
3	I	419/421 (100%)	364 (87%)	38 (9%)	17 (4%)	3	31
3	L	419/421 (100%)	360 (86%)	40 (10%)	19 (4%)	3	29
All	All	4040/4404 (92%)	3431 (85%)	403 (10%)	206 (5%)	4	27

5 of 206 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	136	ILE
1	A	182	HIS
2	B	69	SER
2	B	86	PRO
2	B	89	TRP

5.3.2 Protein sidechains ⓘ

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	145/217 (67%)	140 (97%)	5 (3%)	42	70
1	D	145/217 (67%)	140 (97%)	5 (3%)	42	70
1	G	145/217 (67%)	142 (98%)	3 (2%)	59	80
1	J	145/217 (67%)	138 (95%)	7 (5%)	30	62
2	B	364/366 (100%)	338 (93%)	26 (7%)	17	50
2	E	364/366 (100%)	355 (98%)	9 (2%)	53	77
2	H	364/366 (100%)	350 (96%)	14 (4%)	38	67
2	K	364/366 (100%)	346 (95%)	18 (5%)	29	61
3	C	366/366 (100%)	352 (96%)	14 (4%)	38	67

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	F	366/366 (100%)	353 (96%)	13 (4%)	40	68
3	I	366/366 (100%)	343 (94%)	23 (6%)	21	54
3	L	366/366 (100%)	351 (96%)	15 (4%)	35	65
All	All	3500/3796 (92%)	3348 (96%)	152 (4%)	38	64

5 of 152 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	F	307	GLU
2	H	340	LYS
3	L	129	LYS
3	F	365	THR
2	H	42	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 35 such sidechains are listed below:

Mol	Chain	Res	Type
3	F	284	HIS
1	G	183	HIS
3	L	109	GLN
3	F	347	HIS
1	G	138	ASN

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 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.