



wwPDB EM Map/Model Validation Report ⓘ

May 31, 2016 – 08:36 PM EDT

PDB ID : 3J2W
EMDB ID: : EMD-5577
Title : Electron cryo-microscopy of Chikungunya virus
Authors : Sun, S.; Xiang, Y.; Rossmann, M.G.
Deposited on : 2013-01-28
Resolution : 5.00 Å(reported)
Based on PDB ID : 3N43

This is a wwPDB EM Map/Model Validation Report for a publicly released PDB/EMDB entry.
For rigid body fitted models, validation errors reported here could stem from errors in the original structure(s) used in the fitting.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at
<http://wwpdb.org/validation/2016/EMValidationReportHelp>

MolProbity : 4.02b-467
Mogul : unknown
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20027674

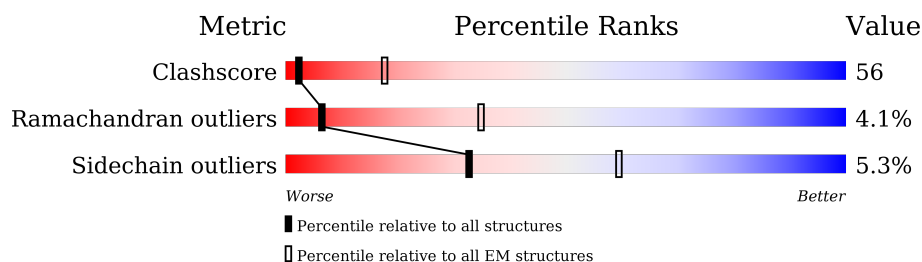
1 Overall quality at a glance

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.






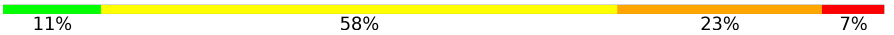
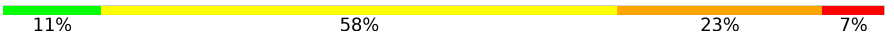
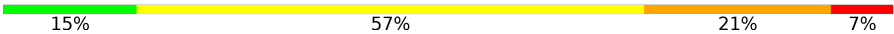
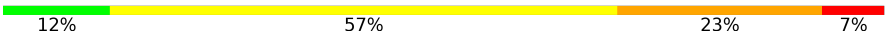




Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	114402	924
Ramachandran outliers	111179	726
Sidechain outliers	111093	686

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	393	78% 19% .
1	B	393	80% 17% .
1	C	393	79% 18% .
1	D	393	80% 17% .
2	M	336	72% 26% ..
2	N	336	65% 31% .
2	O	336	71% 25% .
3	P	336	67% 29% .
4	E	46	35% 52% 9% .

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	F	46	 37% 48% 13% .
4	G	46	 37% 48% 13% .
4	H	46	 35% 54% 7% .
5	Q	81	 11% 58% 23% 7%
5	R	81	 11% 58% 23% 7%
5	S	81	 15% 57% 21% 7%
5	T	81	 12% 57% 23% 7%
6	I	149	 68% 26% 5% .
6	J	149	 62% 34% .
6	K	149	 66% 30% .
6	L	149	 71% 25% .

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 30928 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 Glycoprotein E1.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	393	Total	C	N	O	S	0	0
			2992	1892	501	575	24		
1	B	393	Total	C	N	O	S	0	0
			2992	1892	501	575	24		
1	C	393	Total	C	N	O	S	0	0
			2992	1892	501	575	24		
1	D	393	Total	C	N	O	S	0	0
			2992	1892	501	575	24		

- Molecule 2 is a protein called Glycoprotein E2.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	M	336	Total	C	N	O	S	0	0
			2650	1653	480	497	20		
2	N	336	Total	C	N	O	S	0	0
			2650	1653	480	497	20		
2	O	336	Total	C	N	O	S	0	0
			2650	1653	480	497	20		

- Molecule 3 is a protein called Glycoprotein E2.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	P	336	Total	C	N	O	S	0	0
			2650	1652	480	497	21		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
P	3842	MET	LEU	CONFLICT	UNP Q1H8W5

- Molecule 4 is a protein called Glycoprotein E1.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	E	46	Total	C	N	O	S	0	0
			336	218	57	59	2		
4	F	46	Total	C	N	O	S	0	0
			336	218	57	59	2		
4	G	46	Total	C	N	O	S	0	0
			336	218	57	59	2		
4	H	46	Total	C	N	O	S	0	0
			336	218	57	59	2		

- Molecule 5 is a protein called Glycoprotein E2.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	Q	81	Total	C	N	O	S	0	0
			613	396	101	108	8		
5	R	81	Total	C	N	O	S	0	0
			613	396	101	108	8		
5	S	81	Total	C	N	O	S	0	0
			613	396	101	108	8		
5	T	81	Total	C	N	O	S	0	0
			613	396	101	108	8		

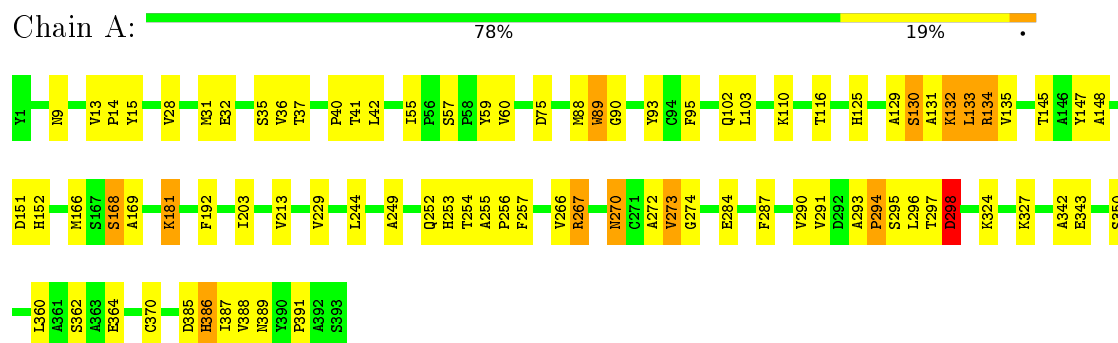
- Molecule 6 is a protein called Capsid protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	I	149	Total	C	N	O	S	0	0
			1141	723	201	211	6		
6	J	149	Total	C	N	O	S	0	0
			1141	723	201	211	6		
6	K	149	Total	C	N	O	S	0	0
			1141	723	201	211	6		
6	L	149	Total	C	N	O	S	0	0
			1141	723	201	211	6		

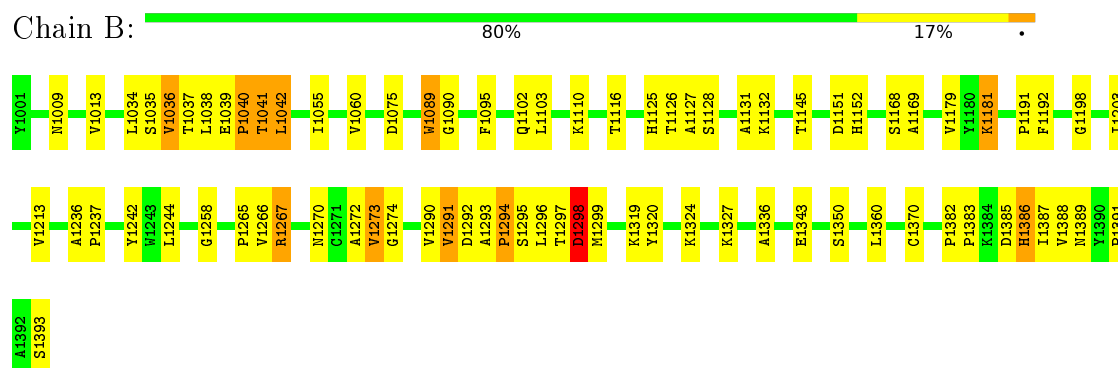
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

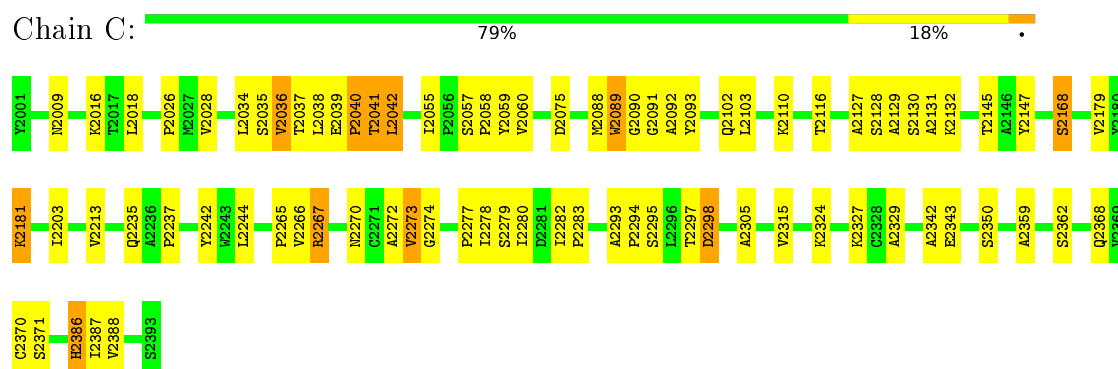
• Molecule 1: Glycoprotein E1



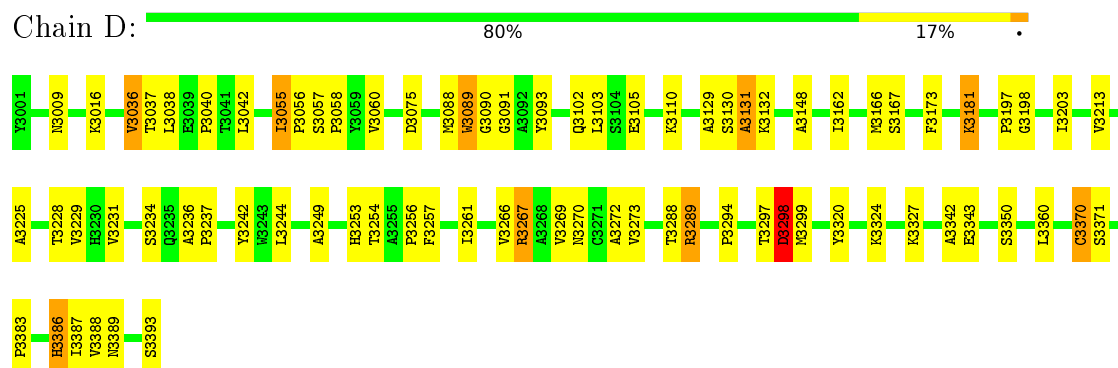
• Molecule 1: Glycoprotein E1



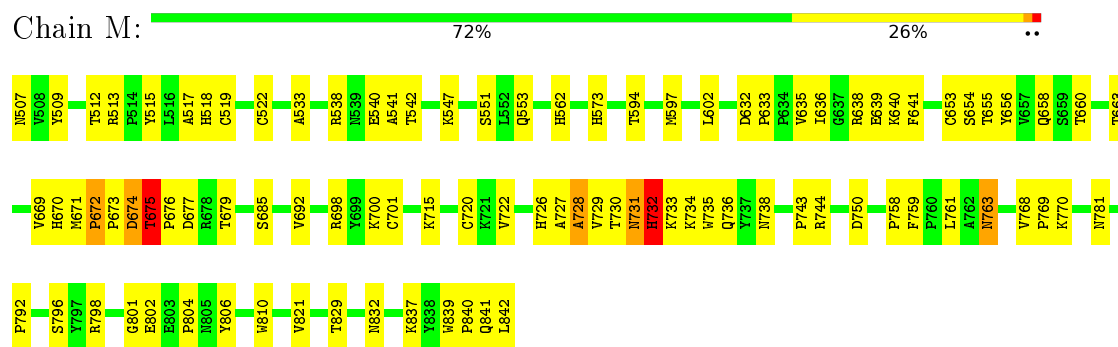
• Molecule 1: Glycoprotein E1



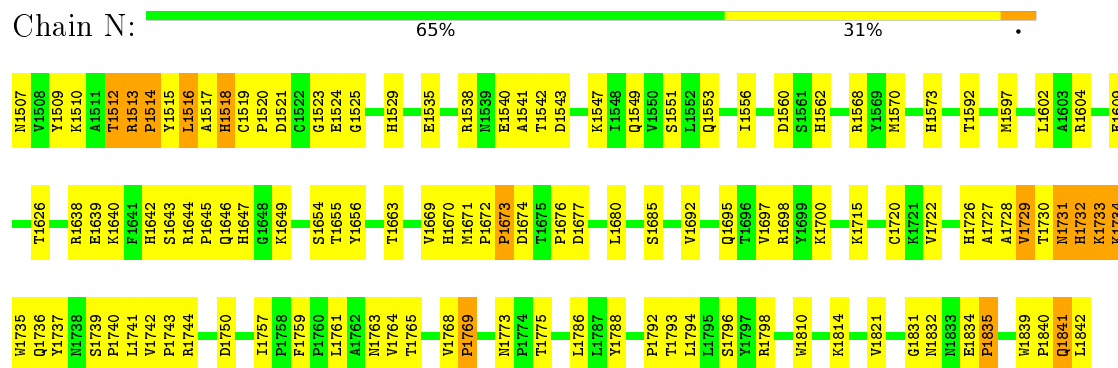
- Molecule 1: Glycoprotein E1



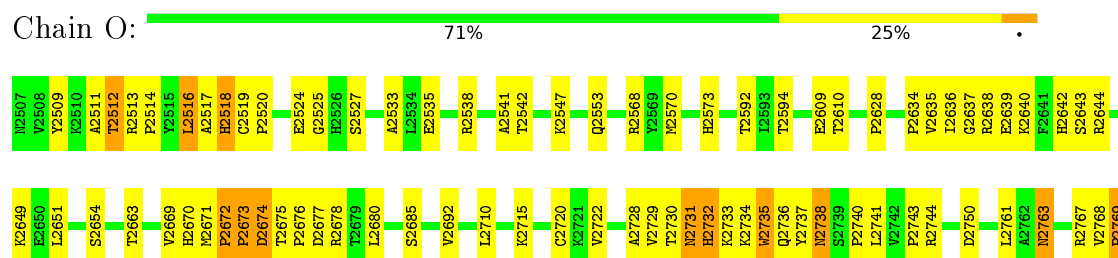
- Molecule 2: Glycoprotein E2

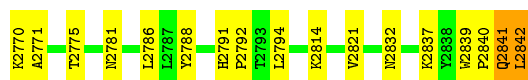


- Molecule 2: Glycoprotein E2



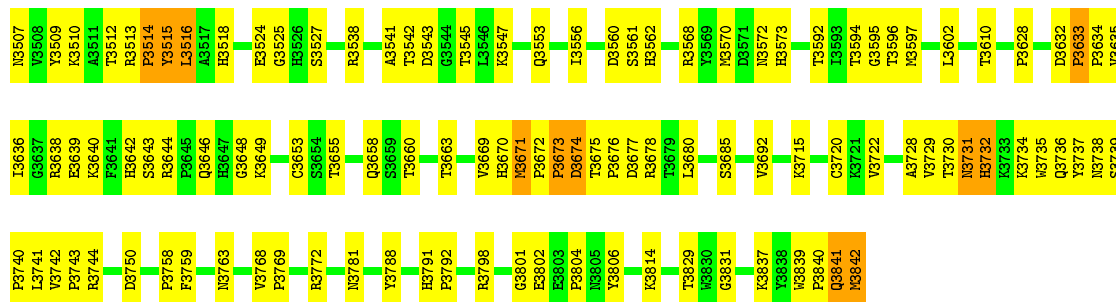
- Molecule 2: Glycoprotein E2





• Molecule 3: Glycoprotein E2

Chain P: 67% 29% .



• Molecule 4: Glycoprotein E1

Chain E: 35% 52% 9% .



• Molecule 4: Glycoprotein E1

Chain F: 37% 48% 13% .



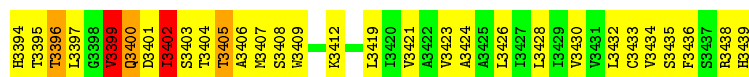
• Molecule 4: Glycoprotein E1

Chain G: 37% 48% 13% .



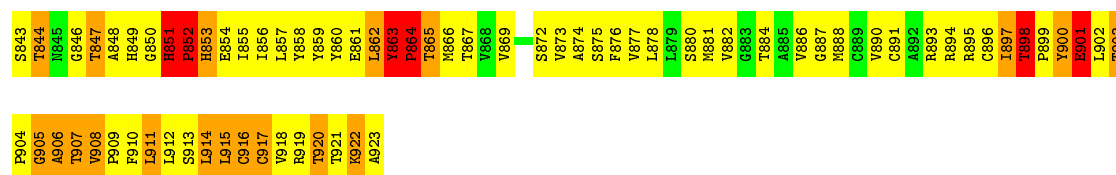
• Molecule 4: Glycoprotein E1

Chain H: 35% 54% 7% .



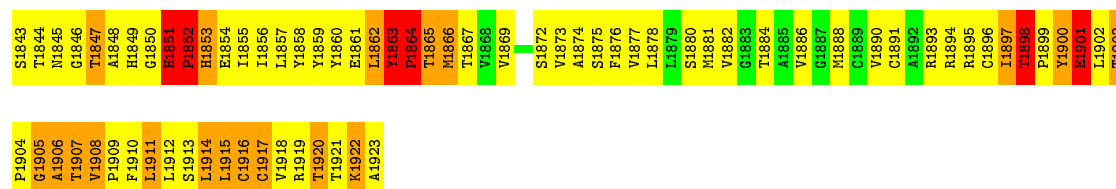
• Molecule 5: Glycoprotein E2

Chain Q: 11% 58% 23% 7% .



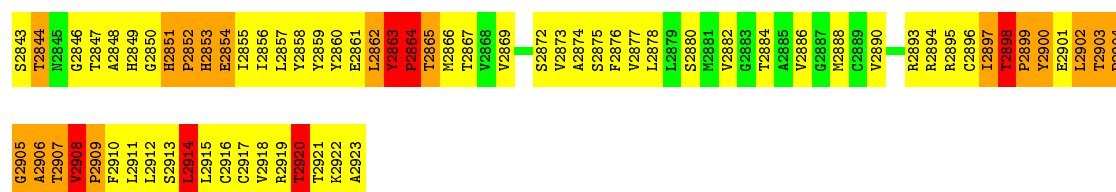
- Molecule 5: Glycoprotein E2

Chain R: 11% 58% 23% 7%



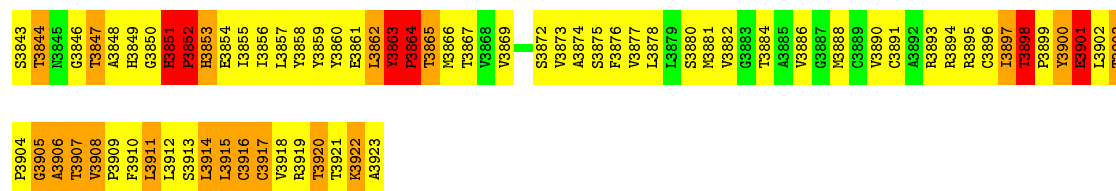
- Molecule 5: Glycoprotein E2

Chain S: 15% 57% 21% 7%



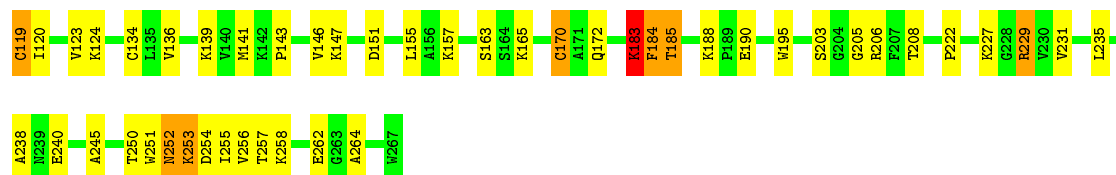
- Molecule 5: Glycoprotein E2

Chain T: 12% 57% 23% 7%



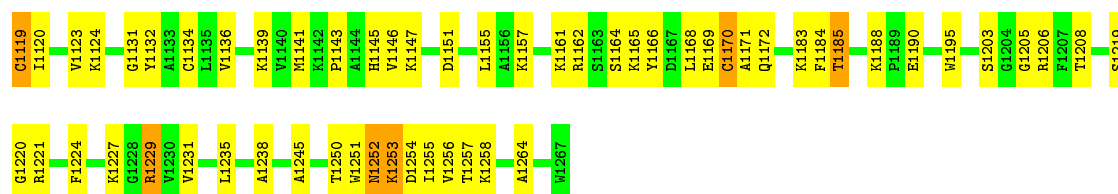
- Molecule 6: Capsid protein

Chain I: 68% 26% 5%

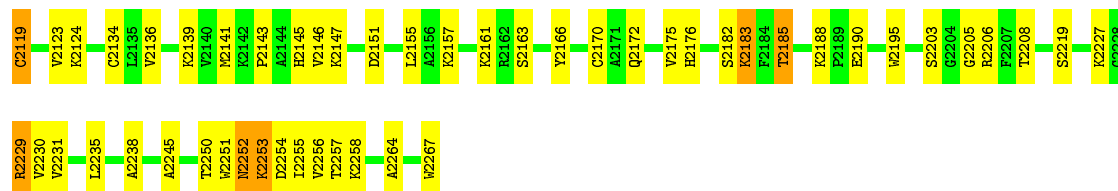


- Molecule 6: Capsid protein

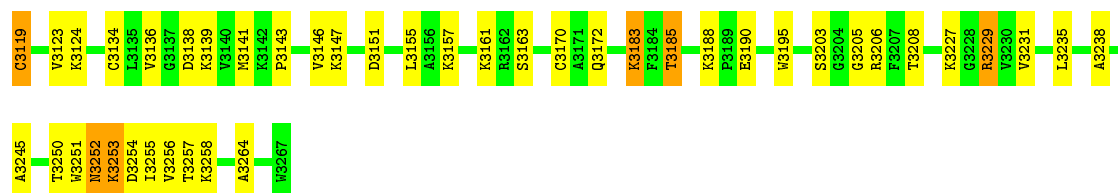
Chain J: 62% 34%



• Molecule 6: Capsid protein



• Molecule 6: Capsid protein



4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, 1	Depositor
Number of particles used	36236	Depositor
Resolution determination method	FSC at 0.5 cut-off	Depositor
CTF correction method	each particle	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	25	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	59000	Depositor
Image detector	film	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	0.57	1/3068 (0.0%)	0.73	2/4184 (0.0%)
1	B	0.35	0/3067	0.54	0/4181
1	C	0.39	1/3068 (0.0%)	0.66	3/4184 (0.1%)
1	D	0.52	1/3069 (0.0%)	0.57	3/4187 (0.1%)
2	M	0.35	0/2721	0.53	0/3704
2	N	0.34	0/2721	0.54	0/3704
2	O	0.34	0/2721	0.54	0/3704
3	P	0.34	0/2721	0.53	0/3703
4	E	0.20	0/340	0.38	0/464
4	F	0.20	0/339	0.38	0/461
4	G	0.20	0/339	0.38	0/461
4	H	0.20	0/340	0.38	0/464
5	Q	1.01	2/626 (0.3%)	1.94	7/855 (0.8%)
5	R	1.01	2/626 (0.3%)	2.00	7/855 (0.8%)
5	S	1.19	4/614 (0.7%)	2.35	12/819 (1.5%)
5	T	1.01	2/626 (0.3%)	1.94	7/855 (0.8%)
6	I	0.79	2/1169 (0.2%)	1.56	5/1577 (0.3%)
6	J	0.78	2/1168 (0.2%)	0.90	2/1574 (0.1%)
6	K	0.80	3/1169 (0.3%)	1.20	5/1577 (0.3%)
6	L	0.83	3/1169 (0.3%)	0.95	5/1577 (0.3%)
All	All	0.56	23/31681 (0.1%)	0.89	58/43090 (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
1	C	0	1
5	Q	0	1
5	R	0	1
5	S	0	4
5	T	0	1

Continued on next page...

Continued from previous page...

Mol	Chain	#Chirality outliers	#Planarity outliers
6	I	0	1
6	L	0	1
All	All	0	10

The worst 5 of 23 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	168	SER	C-N	25.07	1.91	1.34
5	R	1862	LEU	C-N	22.22	1.85	1.34
5	S	2862	LEU	C-N	22.18	1.85	1.34
5	T	3862	LEU	C-N	22.18	1.85	1.34
5	Q	862	LEU	C-N	22.17	1.85	1.34

The worst 5 of 58 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	I	183	LYS	O-C-N	-38.33	61.38	122.70
5	S	2914	LEU	O-C-N	-35.04	66.64	122.70
5	R	1862	LEU	O-C-N	-31.77	71.86	122.70
5	Q	862	LEU	O-C-N	-31.75	71.90	122.70
5	T	3862	LEU	O-C-N	-31.74	71.92	122.70

There are no chirality outliers.

5 of 10 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	2168	SER	Mainchain
5	Q	863	TYR	Mainchain
5	R	1863	TYR	Mainchain
5	S	2863	TYR	Mainchain
5	S	2914	LEU	Mainchain

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	2992	0	2890	212	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	2992	0	2894	157	0
1	C	2992	0	2890	207	0
1	D	2992	0	2892	260	0
2	M	2650	0	2571	205	0
2	N	2650	0	2566	405	0
2	O	2650	0	2571	277	0
3	P	2650	0	2566	421	0
4	E	336	0	360	138	0
4	F	336	0	359	124	0
4	G	336	0	359	93	0
4	H	336	0	361	105	0
5	Q	613	0	623	322	0
5	R	613	0	625	417	0
5	S	613	0	611	390	0
5	T	613	0	625	329	0
6	I	1141	0	1126	78	0
6	J	1141	0	1123	126	0
6	K	1141	0	1124	65	0
6	L	1141	0	1126	59	0
All	All	30928	0	30262	3436	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 56.

The worst 5 of 3436 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:S:2909:PRO:HD3	5:S:2910:PHE:CD2	1.22	1.69
2:N:1509:TYR:HB2	2:N:1562:HIS:CE1	1.16	1.65
2:N:1602:LEU:HD11	2:N:1759:PHE:CD2	1.17	1.65
2:O:2573:HIS:CE1	2:O:2729:VAL:HG21	1.31	1.61
1:C:2362:SER:HB3	4:G:2402:ILE:CD1	1.18	1.60

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	389/393 (99%)	363 (93%)	20 (5%)	6 (2%)	13	58
1	B	387/393 (98%)	357 (92%)	22 (6%)	8 (2%)	9	51
1	C	389/393 (99%)	357 (92%)	26 (7%)	6 (2%)	13	58
1	D	391/393 (100%)	368 (94%)	19 (5%)	4 (1%)	19	65
2	M	334/336 (99%)	299 (90%)	30 (9%)	5 (2%)	13	58
2	N	334/336 (99%)	295 (88%)	29 (9%)	10 (3%)	5	44
2	O	334/336 (99%)	297 (89%)	29 (9%)	8 (2%)	7	48
3	P	334/336 (99%)	291 (87%)	36 (11%)	7 (2%)	9	51
4	E	44/46 (96%)	35 (80%)	4 (9%)	5 (11%)	0	10
4	F	42/46 (91%)	35 (83%)	3 (7%)	4 (10%)	1	15
4	G	42/46 (91%)	35 (83%)	3 (7%)	4 (10%)	1	15
4	H	44/46 (96%)	35 (80%)	5 (11%)	4 (9%)	1	16
5	Q	77/81 (95%)	41 (53%)	12 (16%)	24 (31%)	0	0
5	R	77/81 (95%)	39 (51%)	15 (20%)	23 (30%)	0	0
5	S	63/81 (78%)	38 (60%)	6 (10%)	19 (30%)	0	0
5	T	77/81 (95%)	40 (52%)	13 (17%)	24 (31%)	0	0
6	I	147/149 (99%)	131 (89%)	14 (10%)	2 (1%)	14	58
6	J	145/149 (97%)	131 (90%)	14 (10%)	0	100	100
6	K	147/149 (99%)	133 (90%)	14 (10%)	0	100	100
6	L	147/149 (99%)	133 (90%)	14 (10%)	0	100	100
All	All	3944/4020 (98%)	3453 (88%)	328 (8%)	163 (4%)	6	35

5 of 163 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	130	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	273	VAL
1	A	298	ASP
1	B	1036	VAL
1	B	1273	VAL

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	329/329 (100%)	312 (95%)	17 (5%)	29	67
1	B	329/329 (100%)	313 (95%)	16 (5%)	31	68
1	C	329/329 (100%)	313 (95%)	16 (5%)	31	68
1	D	329/329 (100%)	314 (95%)	15 (5%)	33	69
2	M	298/298 (100%)	284 (95%)	14 (5%)	32	69
2	N	298/298 (100%)	285 (96%)	13 (4%)	35	70
2	O	298/298 (100%)	284 (95%)	14 (5%)	32	69
3	P	298/298 (100%)	282 (95%)	16 (5%)	27	66
4	E	38/38 (100%)	35 (92%)	3 (8%)	15	53
4	F	38/38 (100%)	34 (90%)	4 (10%)	8	39
4	G	38/38 (100%)	34 (90%)	4 (10%)	8	39
4	H	38/38 (100%)	35 (92%)	3 (8%)	15	53
5	Q	70/70 (100%)	67 (96%)	3 (4%)	35	71
5	R	70/70 (100%)	67 (96%)	3 (4%)	35	71
5	S	70/70 (100%)	68 (97%)	2 (3%)	50	79
5	T	70/70 (100%)	67 (96%)	3 (4%)	35	71
6	I	118/118 (100%)	109 (92%)	9 (8%)	16	55
6	J	118/118 (100%)	109 (92%)	9 (8%)	16	55
6	K	118/118 (100%)	109 (92%)	9 (8%)	16	55
6	L	118/118 (100%)	109 (92%)	9 (8%)	16	55

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	3412/3412 (100%)	3230 (95%)	182 (5%)	33 66

5 of 182 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	O	2692	VAL
1	D	3343	GLU
6	K	2185	THR
2	O	2720	CYS
1	D	3089	TRP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 62 such sidechains are listed below:

Mol	Chain	Res	Type
2	O	2693	ASN
1	D	3275	ASN
6	J	1252	ASN
2	O	2745	ASN
3	P	3562	HIS

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 ⓘ

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.