



wwPDB/EMDataBank EM Map/Model Validation Summary Report ⓘ

Aug 20, 2017 – 11:36 AM EDT

PDB ID : 2CGT
EMDB ID: : EMD-1202
Title : GROEL-ADP-gp31 COMPLEX
Authors : Clare, D.K.; Bakkes, P.J.; van Heerikhuizen, H.; van der Vies, S.M.; Saibil, H.R.
Deposited on : unknown
Resolution : 8.20 Å(reported)
Based on PDB ID : 1A0N, 1G31

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

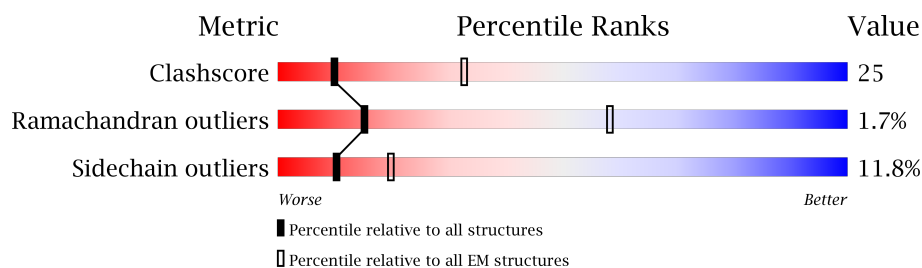
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 8.20 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.








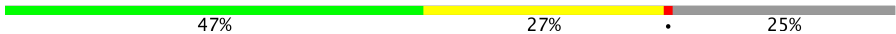
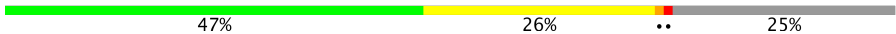
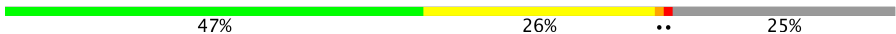
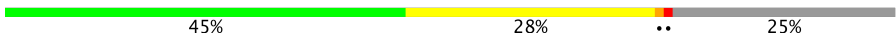
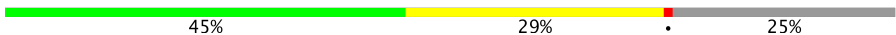
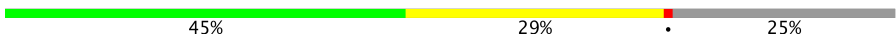
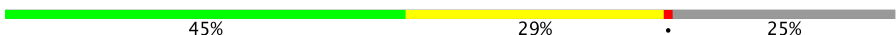
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	125131	1336
Ramachandran outliers	121729	1120
Sidechain outliers	121581	1026

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	547	
1	B	547	
1	C	547	
1	D	547	
1	E	547	
1	F	547	
1	G	547	
1	H	547	
1	I	547	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	J	547	 55% 33% 5% • 6%
1	K	547	 55% 32% 6% • 6%
1	L	547	 55% 33% 5% • 6%
1	M	547	 55% 33% 6% • 6%
1	N	547	 55% 33% 5% • 6%
2	O	111	 47% 27% • 25%
2	P	111	 47% 26% •• 25%
2	Q	111	 47% 26% •• 25%
2	R	111	 45% 28% •• 25%
2	S	111	 45% 29% • 25%
2	T	111	 45% 29% • 25%
2	U	111	 45% 29% • 25%

2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 57953 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 60 KDA GROEL.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	524	Total	C	N	O	S	0	0
			3855	2397	665	773	20		
1	B	524	Total	C	N	O	S	0	0
			3855	2397	665	773	20		
1	C	524	Total	C	N	O	S	0	0
			3855	2397	665	773	20		
1	D	524	Total	C	N	O	S	0	0
			3855	2397	665	773	20		
1	E	524	Total	C	N	O	S	0	0
			3855	2397	665	773	20		
1	F	524	Total	C	N	O	S	0	0
			3855	2397	665	773	20		
1	G	524	Total	C	N	O	S	0	0
			3855	2397	665	773	20		
1	H	515	Total	C	N	O	S	0	0
			3783	2354	656	753	20		
1	I	515	Total	C	N	O	S	0	0
			3783	2354	656	753	20		
1	J	515	Total	C	N	O	S	0	0
			3783	2354	656	753	20		
1	K	515	Total	C	N	O	S	0	0
			3783	2354	656	753	20		
1	L	515	Total	C	N	O	S	0	0
			3783	2354	656	753	20		
1	M	515	Total	C	N	O	S	0	0
			3783	2354	656	753	20		
1	N	515	Total	C	N	O	S	0	0
			3783	2354	656	753	20		

- Molecule 2 is a protein called CAPSID ASSEMBLY PROTEIN GP31.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	O	83	Total	C	N	O	S	0	0
			641	417	106	114	4		

Continued on next page...

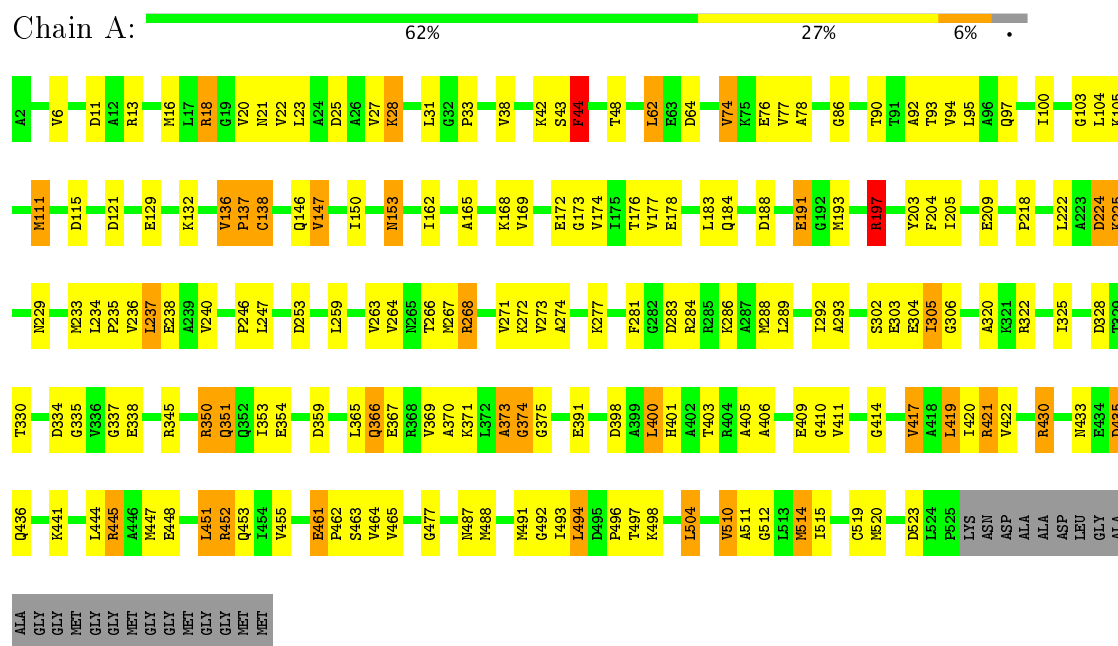
Continued from previous page...

Mol	Chain	Residues	Atoms					AltConf	Trace
2	P	83	Total 641	C 417	N 106	O 114	S 4	0	0
2	Q	83	Total 641	C 417	N 106	O 114	S 4	0	0
2	R	83	Total 641	C 417	N 106	O 114	S 4	0	0
2	S	83	Total 641	C 417	N 106	O 114	S 4	0	0
2	T	83	Total 641	C 417	N 106	O 114	S 4	0	0
2	U	83	Total 641	C 417	N 106	O 114	S 4	0	0

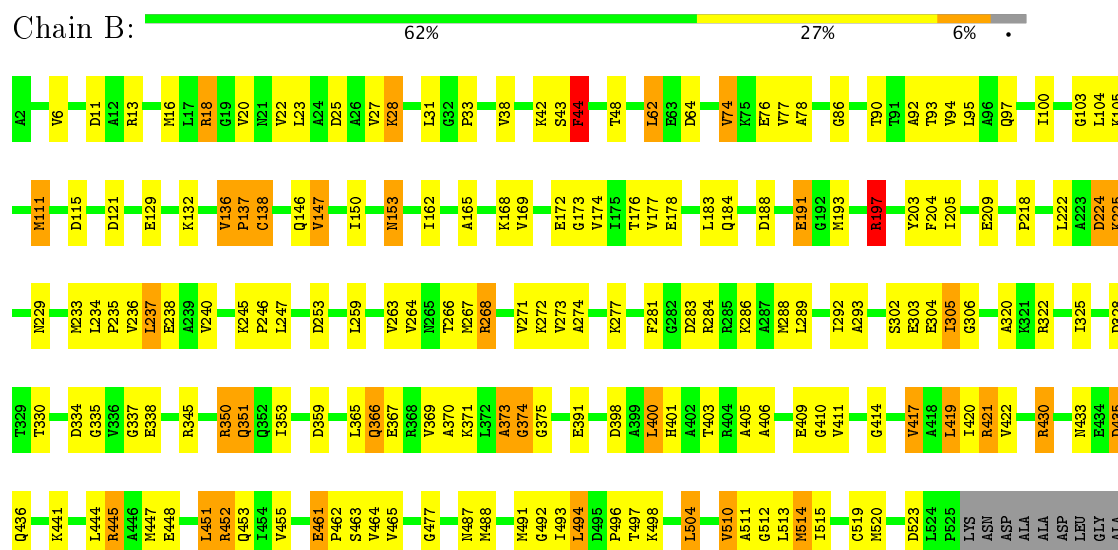
3 Residue-property plots

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: 60 KDA GROEL



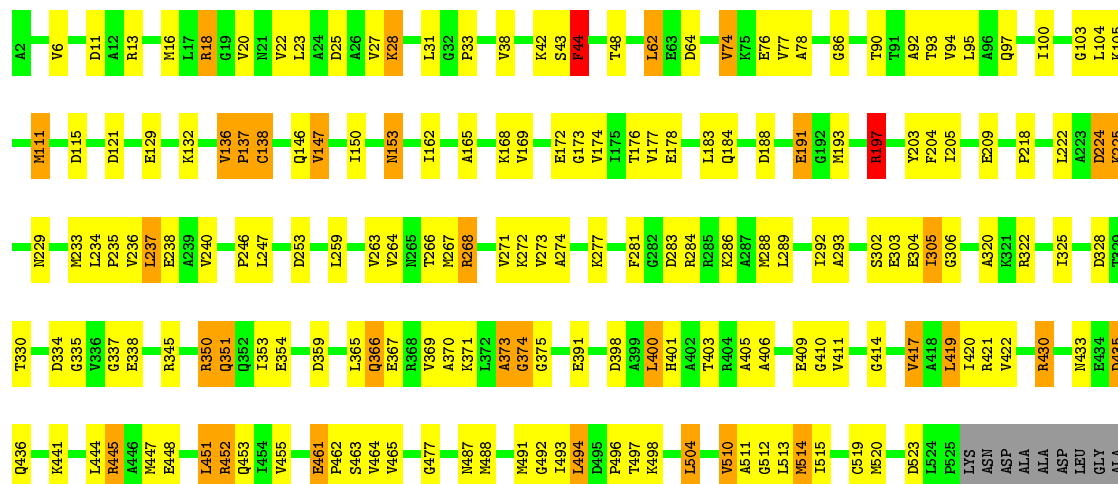
• Molecule 1: 60 KDA GROEL



ALA
GLY
GLY
MET
GLY
GLY
MET
GLY
GLY
MET
MET

• Molecule 1: 60 KDA GROEL

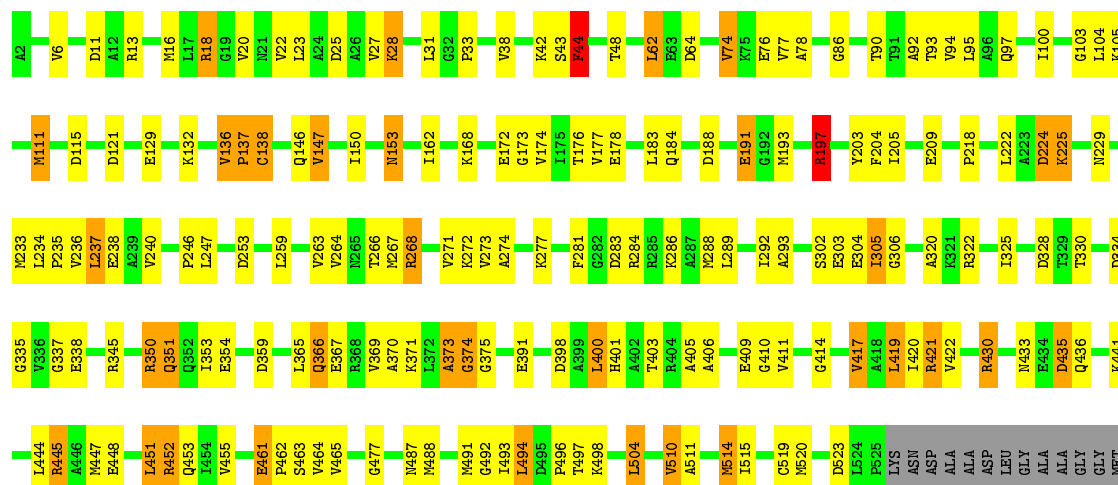
Chain C:  62% 27% 6% .



ALA
GLY
GLY
MET
GLY
GLY
MET
GLY
GLY
MET
MET

• Molecule 1: 60 KDA GROEL

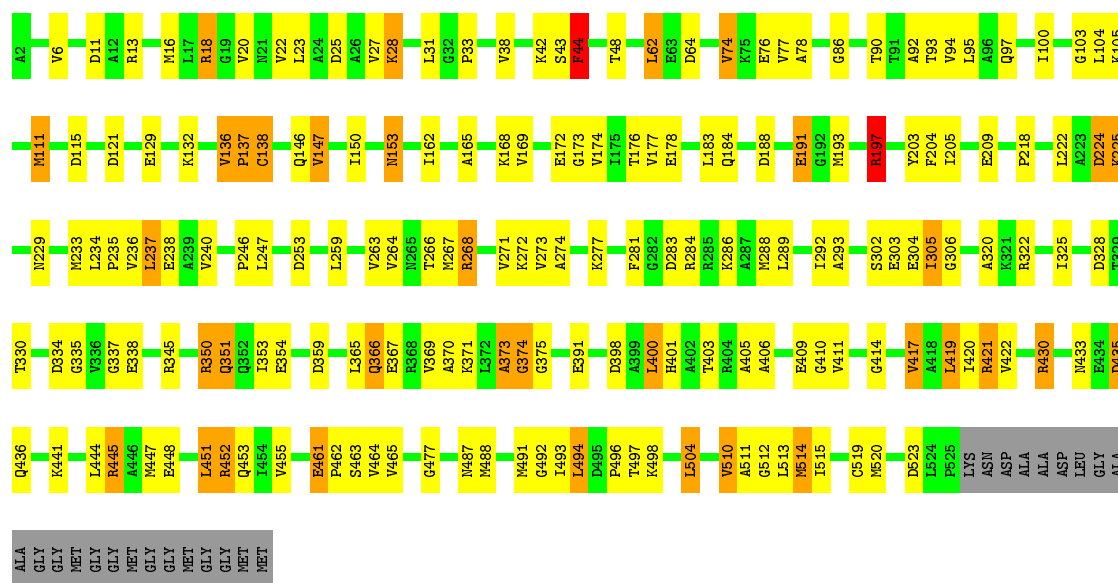
Chain D:  63% 26% 6% .

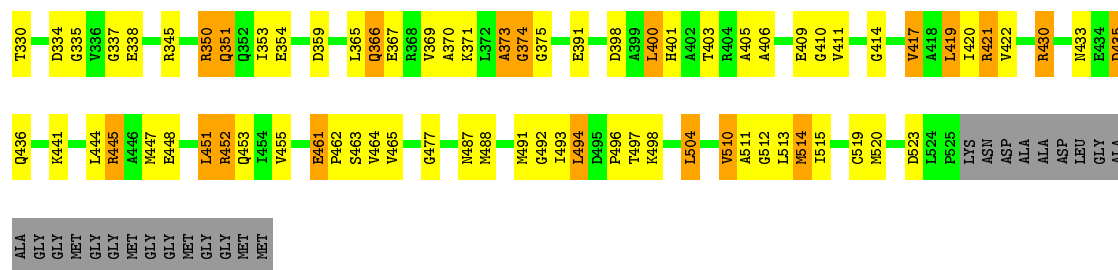


GLY
GLY
MET
GLY
GLY
MET
GLY
GLY
MET
MET

• Molecule 1: 60 KDA GROEL

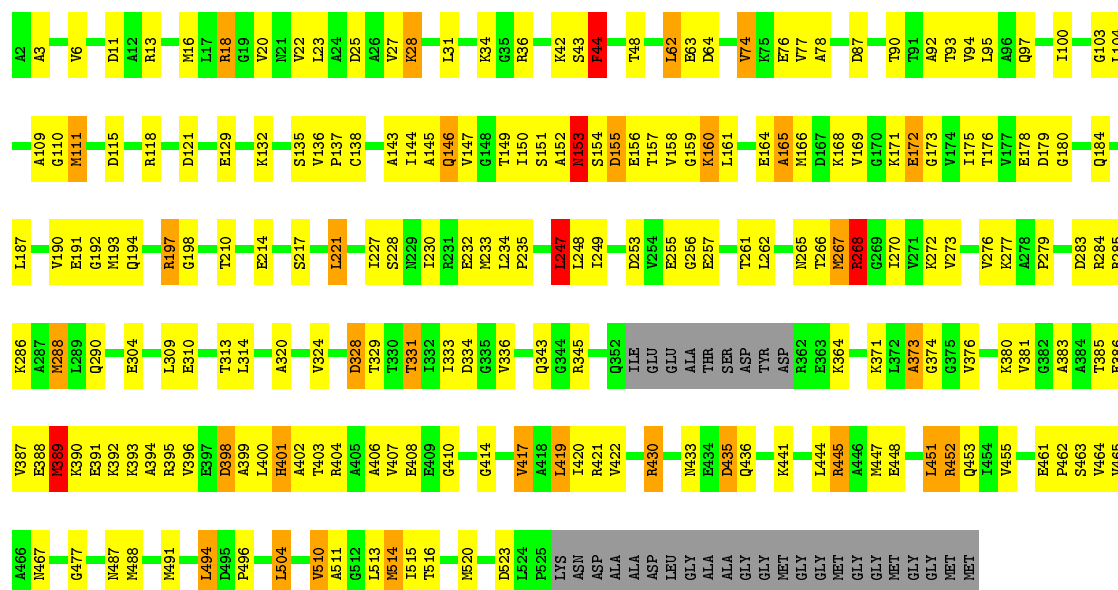
Chain E:  62% 27% 6% .





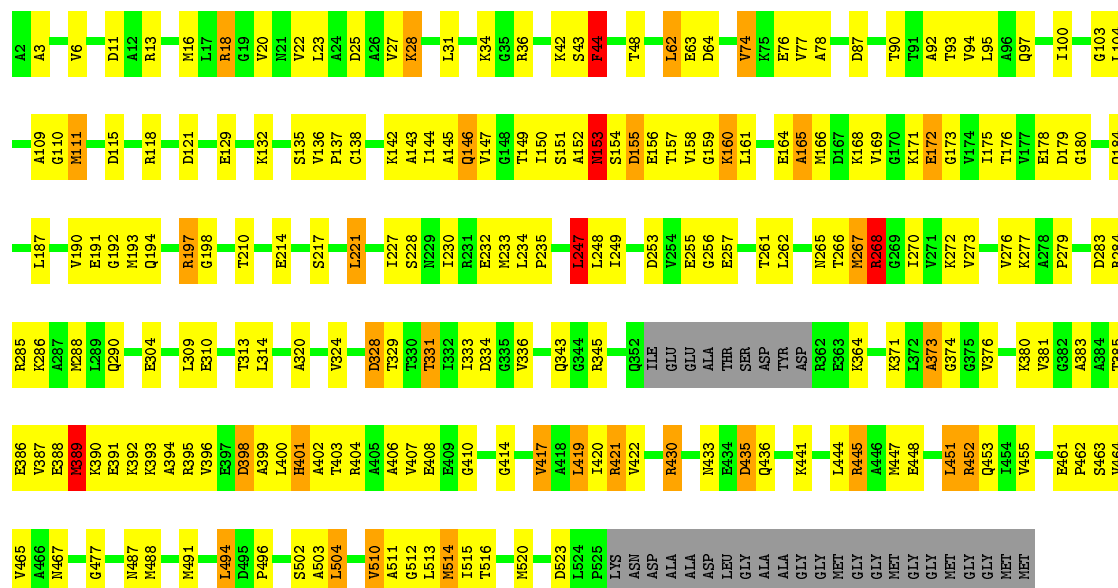
Molecule 1: 60 KDA GROEL

Chain H: 56% 32% 5% • 6%

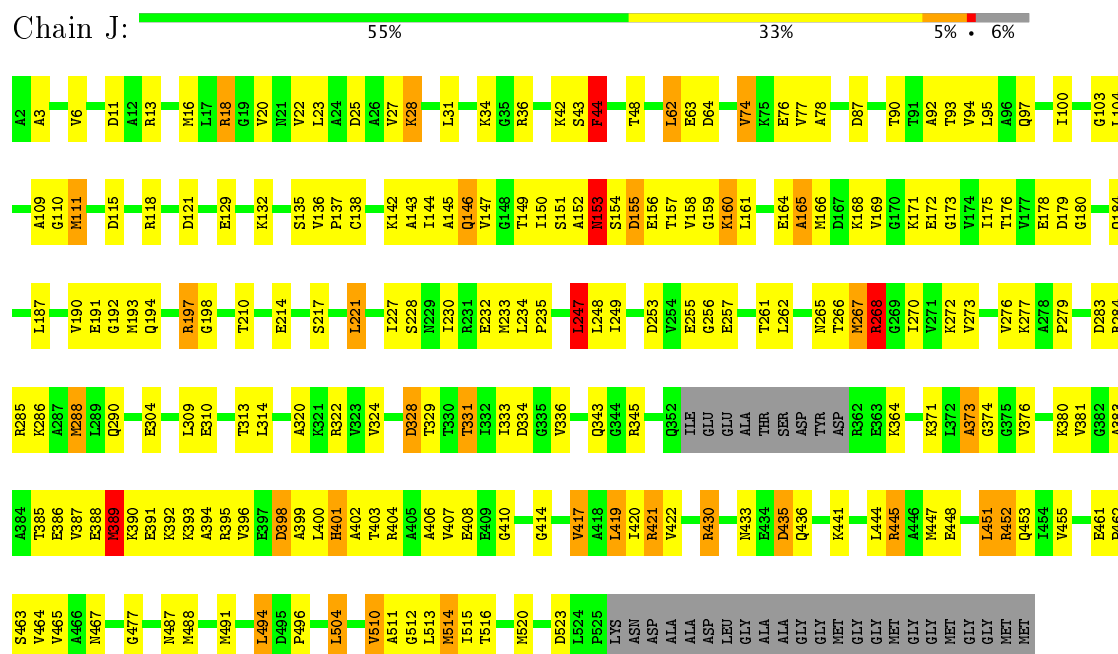


Molecule 1: 60 KDA GROEL

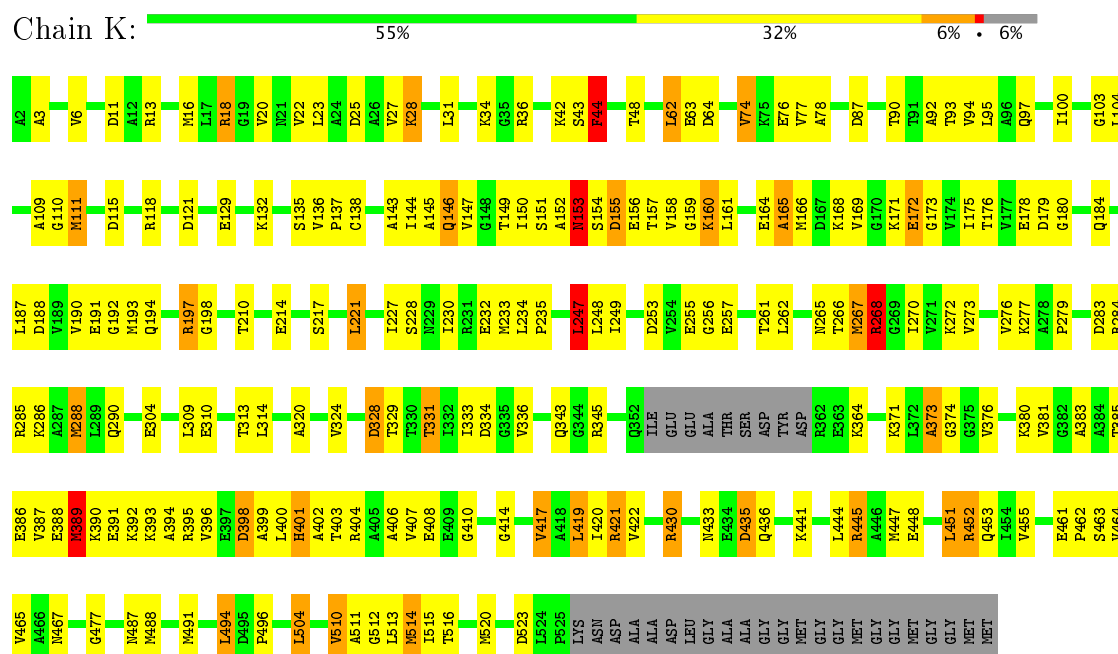
Chain I: 55% 33% 5% • 6%



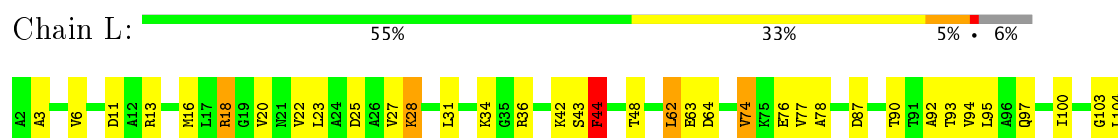
- Molecule 1: 60 KDA GROEL

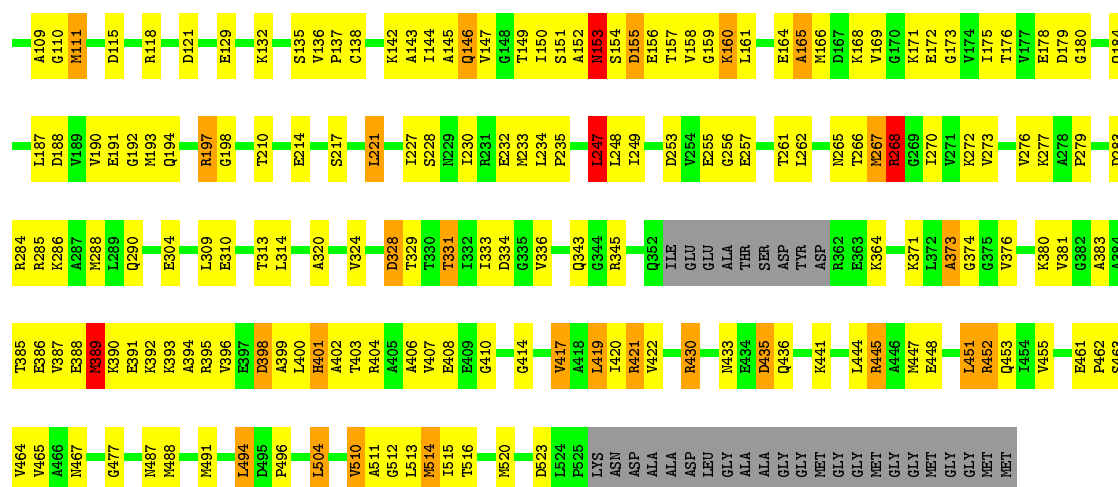


- Molecule 1: 60 KDA GROEL

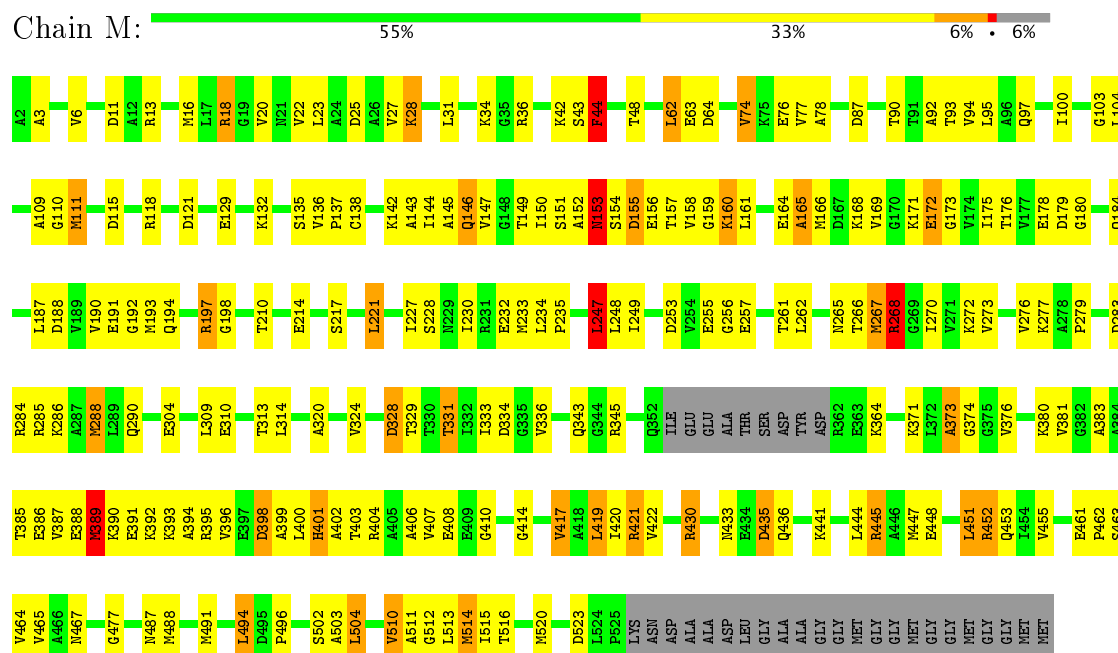


- Molecule 1: 60 KDA GROEL

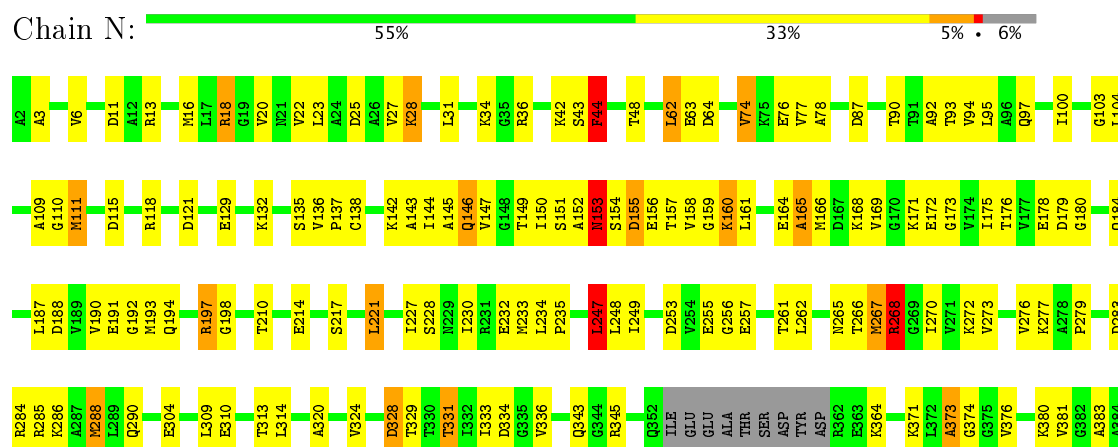


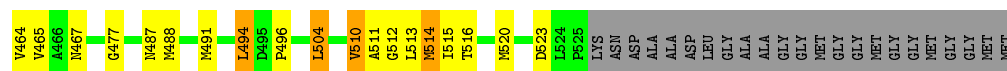
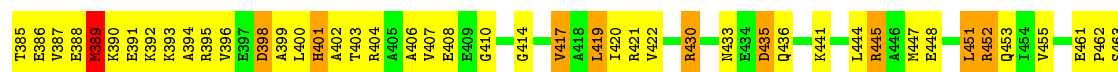


• Molecule 1: 60 KDA GROEL

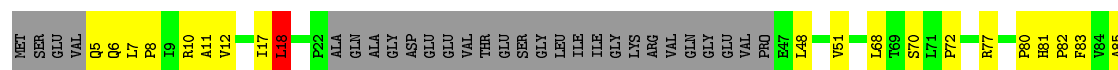


• Molecule 1: 60 KDA GROEL

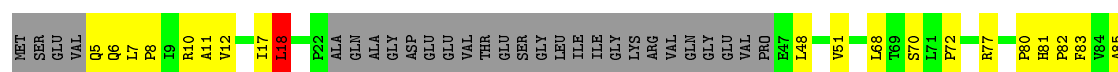




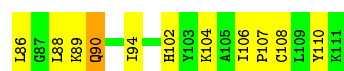
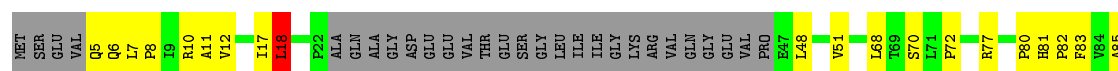
• Molecule 2: CAPSID ASSEMBLY PROTEIN GP31



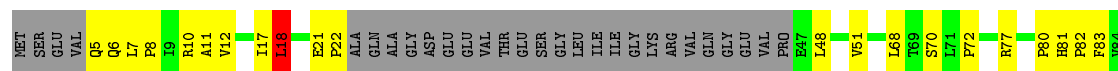
• Molecule 2: CAPSID ASSEMBLY PROTEIN GP31



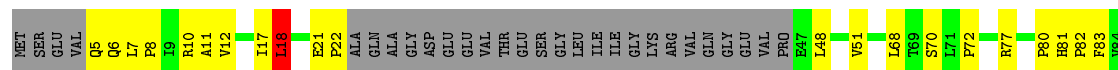
• Molecule 2: CAPSID ASSEMBLY PROTEIN GP31

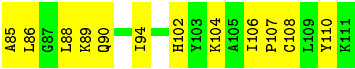


• Molecule 2: CAPSID ASSEMBLY PROTEIN GP31

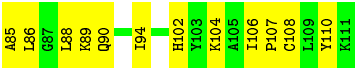


• Molecule 2: CAPSID ASSEMBLY PROTEIN GP31

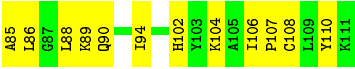




● Molecule 2: CAPSID ASSEMBLY PROTEIN GP31



● Molecule 2: CAPSID ASSEMBLY PROTEIN GP31



4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, D7	Depositor
Number of particles used	10300	Depositor
Resolution determination method	Not provided	Depositor
CTF correction method	FULL CORRECTION ON 2D CLASS AVERAGES	Depositor
Microscope	FEI TECNAI F20	Depositor
Voltage (kV)	200	Depositor
Electron dose ($e^-/\text{\AA}^2$)	15	Depositor
Minimum defocus (nm)	1300	Depositor
Maximum defocus (nm)	3300	Depositor
Magnification	50000	Depositor
Image detector	KODAK SO-163 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	1.07	3/3882 (0.1%)	1.23	27/5240 (0.5%)
1	B	1.07	3/3882 (0.1%)	1.23	27/5240 (0.5%)
1	C	1.07	3/3882 (0.1%)	1.23	27/5240 (0.5%)
1	D	1.07	3/3882 (0.1%)	1.23	27/5240 (0.5%)
1	E	1.07	3/3882 (0.1%)	1.23	27/5240 (0.5%)
1	F	1.07	3/3882 (0.1%)	1.23	27/5240 (0.5%)
1	G	1.07	3/3882 (0.1%)	1.23	27/5240 (0.5%)
1	H	0.68	1/3806 (0.0%)	1.04	15/5130 (0.3%)
1	I	0.69	1/3806 (0.0%)	1.04	15/5130 (0.3%)
1	J	0.69	1/3806 (0.0%)	1.04	15/5130 (0.3%)
1	K	0.69	1/3806 (0.0%)	1.04	15/5130 (0.3%)
1	L	0.69	1/3806 (0.0%)	1.04	15/5130 (0.3%)
1	M	0.69	1/3806 (0.0%)	1.04	15/5130 (0.3%)
1	N	0.69	1/3806 (0.0%)	1.04	15/5130 (0.3%)
2	O	0.59	0/657	0.79	1/894 (0.1%)
2	P	0.59	0/657	0.79	1/894 (0.1%)
2	Q	0.59	0/657	0.79	1/894 (0.1%)
2	R	0.59	0/657	0.79	1/894 (0.1%)
2	S	0.59	0/657	0.79	1/894 (0.1%)
2	T	0.59	0/657	0.79	1/894 (0.1%)
2	U	0.59	0/657	0.79	1/894 (0.1%)
All	All	0.88	28/58415 (0.0%)	1.12	301/78848 (0.4%)

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	1
1	B	0	1
1	C	0	1
1	D	0	1

Continued on next page...

Continued from previous page...

Mol	Chain	#Chirality outliers	#Planarity outliers
1	E	0	1
1	F	0	1
1	G	0	1
1	H	0	1
1	I	0	1
1	J	0	1
1	K	0	1
1	L	0	1
1	M	0	1
1	N	0	1
All	All	0	14

The worst 5 of 28 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	F	191	GLU	C-N	-42.62	0.56	1.33
1	A	191	GLU	C-N	-42.60	0.56	1.33
1	C	191	GLU	C-N	-42.60	0.56	1.33
1	E	191	GLU	C-N	-42.60	0.56	1.33
1	D	191	GLU	C-N	-42.59	0.56	1.33

The worst 5 of 301 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	H	373	ALA	O-C-N	-33.61	66.06	123.20
1	J	373	ALA	O-C-N	-33.61	66.06	123.20
1	N	373	ALA	O-C-N	-33.61	66.06	123.20
1	K	373	ALA	O-C-N	-33.60	66.07	123.20
1	L	373	ALA	O-C-N	-33.60	66.07	123.20

There are no chirality outliers.

5 of 14 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	191	GLU	Mainchain
1	B	191	GLU	Mainchain
1	C	191	GLU	Mainchain
1	D	191	GLU	Mainchain
1	E	191	GLU	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	3855	0	3971	193	0
1	B	3855	0	3971	194	0
1	C	3855	0	3971	197	0
1	D	3855	0	3971	189	0
1	E	3855	0	3971	199	0
1	F	3855	0	3971	196	0
1	G	3855	0	3971	198	0
1	H	3783	0	3916	226	0
1	I	3783	0	3916	233	0
1	J	3783	0	3916	229	0
1	K	3783	0	3916	234	0
1	L	3783	0	3916	228	0
1	M	3783	0	3916	236	0
1	N	3783	0	3916	234	0
2	O	641	0	652	60	0
2	P	641	0	652	61	0
2	Q	641	0	652	61	0
2	R	641	0	652	61	0
2	S	641	0	652	60	0
2	T	641	0	652	60	0
2	U	641	0	652	61	0
All	All	57953	0	59773	2885	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 25.

The worst 5 of 2885 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:L:138:CYS:CB	1:L:410:GLY:HA2	1.38	1.52
1:K:138:CYS:CB	1:K:410:GLY:HA2	1.38	1.51
1:N:138:CYS:CB	1:N:410:GLY:HA2	1.38	1.51
1:M:138:CYS:CB	1:M:410:GLY:HA2	1.38	1.50
1:H:138:CYS:CB	1:H:410:GLY:HA2	1.38	1.49

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	520/547 (95%)	483 (93%)	31 (6%)	6 (1%)	15	57
1	B	520/547 (95%)	483 (93%)	31 (6%)	6 (1%)	15	57
1	C	520/547 (95%)	483 (93%)	31 (6%)	6 (1%)	15	57
1	D	520/547 (95%)	483 (93%)	31 (6%)	6 (1%)	15	57
1	E	520/547 (95%)	483 (93%)	31 (6%)	6 (1%)	15	57
1	F	520/547 (95%)	483 (93%)	31 (6%)	6 (1%)	15	57
1	G	520/547 (95%)	483 (93%)	31 (6%)	6 (1%)	15	57
1	H	505/547 (92%)	447 (88%)	45 (9%)	13 (3%)	6	40
1	I	505/547 (92%)	447 (88%)	45 (9%)	13 (3%)	6	40
1	J	505/547 (92%)	447 (88%)	45 (9%)	13 (3%)	6	40
1	K	505/547 (92%)	447 (88%)	45 (9%)	13 (3%)	6	40
1	L	505/547 (92%)	447 (88%)	45 (9%)	13 (3%)	6	40
1	M	505/547 (92%)	447 (88%)	45 (9%)	13 (3%)	6	40
1	N	505/547 (92%)	447 (88%)	45 (9%)	13 (3%)	6	40
2	O	79/111 (71%)	76 (96%)	3 (4%)	0	100	100
2	P	79/111 (71%)	76 (96%)	3 (4%)	0	100	100
2	Q	79/111 (71%)	76 (96%)	3 (4%)	0	100	100
2	R	79/111 (71%)	76 (96%)	3 (4%)	0	100	100
2	S	79/111 (71%)	76 (96%)	3 (4%)	0	100	100
2	T	79/111 (71%)	76 (96%)	3 (4%)	0	100	100
2	U	79/111 (71%)	76 (96%)	3 (4%)	0	100	100
All	All	7728/8435 (92%)	7042 (91%)	553 (7%)	133 (2%)	15	50

5 of 133 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	44	PHE
1	A	137	PRO
1	A	337	GLY
1	B	44	PHE
1	B	137	PRO

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	404/414 (98%)	352 (87%)	52 (13%)	5	25
1	B	404/414 (98%)	352 (87%)	52 (13%)	5	25
1	C	404/414 (98%)	352 (87%)	52 (13%)	5	25
1	D	404/414 (98%)	352 (87%)	52 (13%)	5	25
1	E	404/414 (98%)	352 (87%)	52 (13%)	5	25
1	F	404/414 (98%)	352 (87%)	52 (13%)	5	25
1	G	404/414 (98%)	352 (87%)	52 (13%)	5	25
1	H	396/414 (96%)	347 (88%)	49 (12%)	5	26
1	I	396/414 (96%)	347 (88%)	49 (12%)	5	26
1	J	396/414 (96%)	347 (88%)	49 (12%)	5	26
1	K	396/414 (96%)	347 (88%)	49 (12%)	5	26
1	L	396/414 (96%)	347 (88%)	49 (12%)	5	26
1	M	396/414 (96%)	347 (88%)	49 (12%)	5	26
1	N	396/414 (96%)	347 (88%)	49 (12%)	5	26
2	O	73/96 (76%)	71 (97%)	2 (3%)	50	74
2	P	73/96 (76%)	71 (97%)	2 (3%)	50	74
2	Q	73/96 (76%)	71 (97%)	2 (3%)	50	74
2	R	73/96 (76%)	71 (97%)	2 (3%)	50	74
2	S	73/96 (76%)	71 (97%)	2 (3%)	50	74

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	T	73/96 (76%)	71 (97%)	2 (3%)	50	74
2	U	73/96 (76%)	71 (97%)	2 (3%)	50	74
All	All	6111/6468 (94%)	5390 (88%)	721 (12%)	10	27

5 of 721 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	147	VAL
1	H	422	VAL
1	N	74	VAL
1	G	225	LYS
1	H	23	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 152 such sidechains are listed below:

Mol	Chain	Res	Type
1	I	97	GLN
1	K	146	GLN
2	S	81	HIS
1	I	153	ASN
1	J	146	GLN

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](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	G	3
1	J	3
1	D	3
1	K	3
1	E	3
1	H	3
1	B	3
1	I	3
1	C	3
1	A	3
1	N	3
1	L	3
1	F	3
1	M	3

The worst 5 of 42 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	H	409:GLU	C	410:GLY	N	4.21
1	I	409:GLU	C	410:GLY	N	4.21
1	J	409:GLU	C	410:GLY	N	4.21
1	K	409:GLU	C	410:GLY	N	4.21
1	L	409:GLU	C	410:GLY	N	4.21