



wwPDB EM Validation Summary Report ⓘ

Oct 22, 2024 – 12:19 AM JST

PDB ID : 8IHP
EMDB ID : EMD-35451
Title : Structure of Semliki Forest virus VLP in complex with the receptor VLDLR-LA3
Authors : Cao, D.; Ma, B.; Cao, Z.; Zhang, X.; Xiang, Y.
Deposited on : 2023-02-23
Resolution : 3.00 Å(reported)

This is a wwPDB EM Validation Summary Report for a publicly released PDB 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 ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

EMDB validation analysis : **FAILED**
Mogul : 1.8.5 (274361), CSD as541be (2020)
MolProbity : 4.02b-467
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
MapQ : **FAILED**
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.39

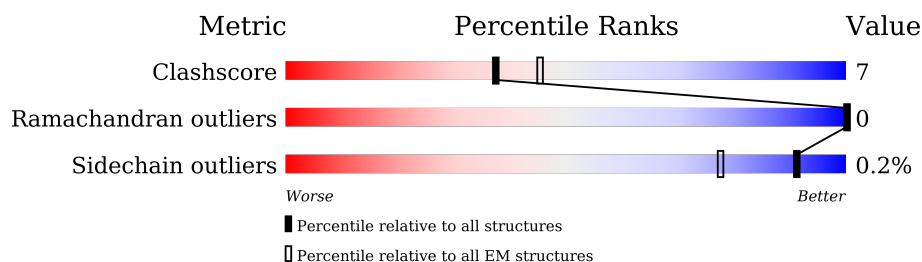
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 3.00 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	210492	15764
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. 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	422	80% 18% .
1	D	422	84% 15% .
1	G	422	80% 19% .
1	J	422	84% 15% .
2	B	438	86% 14%
2	E	438	81% 18%
2	H	438	88% 12%
2	K	438	87% 12%
3	C	162	77% 23%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	F	162	 70% 29% .
3	I	162	 75% 25%
3	L	162	 71% 29%
4	M	39	 79% 18% .
4	N	39	 90% 10%
4	O	39	 87% 13%

2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 32337 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 Spike glycoprotein E2.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	417	Total	C	N	O	S	0	0
			3250	2053	574	598	25		
1	D	417	Total	C	N	O	S	0	0
			3250	2053	574	598	25		
1	G	417	Total	C	N	O	S	0	0
			3250	2053	574	598	25		
1	J	417	Total	C	N	O	S	0	0
			3250	2053	574	598	25		

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	162	LYS	GLU	conflict	UNP P03315
D	162	LYS	GLU	conflict	UNP P03315
G	162	LYS	GLU	conflict	UNP P03315
J	162	LYS	GLU	conflict	UNP P03315

- Molecule 2 is a protein called Spike glycoprotein E1.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	437	Total	C	N	O	S	0	0
			3318	2103	555	635	25		
2	E	437	Total	C	N	O	S	0	0
			3318	2103	555	635	25		
2	H	437	Total	C	N	O	S	0	0
			3318	2103	555	635	25		
2	K	437	Total	C	N	O	S	0	0
			3318	2103	555	635	25		

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	323	ASP	ASN	variant	UNP P03315

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	323	ASP	ASN	variant	UNP P03315
H	323	ASP	ASN	variant	UNP P03315
K	323	ASP	ASN	variant	UNP P03315

- Molecule 3 is a protein called Capsid protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	C	162	Total	C	N	O	S	0	0
			1250	781	227	233	9		
3	F	162	Total	C	N	O	S	0	0
			1250	781	227	233	9		
3	I	162	Total	C	N	O	S	0	0
			1250	781	227	233	9		
3	L	162	Total	C	N	O	S	0	0
			1250	781	227	233	9		

- Molecule 4 is a protein called Very low-density lipoprotein receptor.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	M	39	Total	C	N	O	S	0	0
			298	170	56	66	6		
4	N	39	Total	C	N	O	S	0	0
			298	170	56	66	6		
4	O	39	Total	C	N	O	S	0	0
			298	170	56	66	6		

- Molecule 5 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: C₈H₁₅NO₆).



Mol	Chain	Residues	Atoms				AltConf
5	A	1	Total	C	N	O	0
			14	8	1	5	
5	A	1	Total	C	N	O	0
			14	8	1	5	
5	B	1	Total	C	N	O	0
			14	8	1	5	
5	D	1	Total	C	N	O	0
			14	8	1	5	
5	D	1	Total	C	N	O	0
			14	8	1	5	
5	E	1	Total	C	N	O	0
			14	8	1	5	
5	G	1	Total	C	N	O	0
			14	8	1	5	
5	G	1	Total	C	N	O	0
			14	8	1	5	
5	H	1	Total	C	N	O	0
			14	8	1	5	
5	J	1	Total	C	N	O	0
			14	8	1	5	
5	J	1	Total	C	N	O	0
			14	8	1	5	
5	K	1	Total	C	N	O	0
			14	8	1	5	

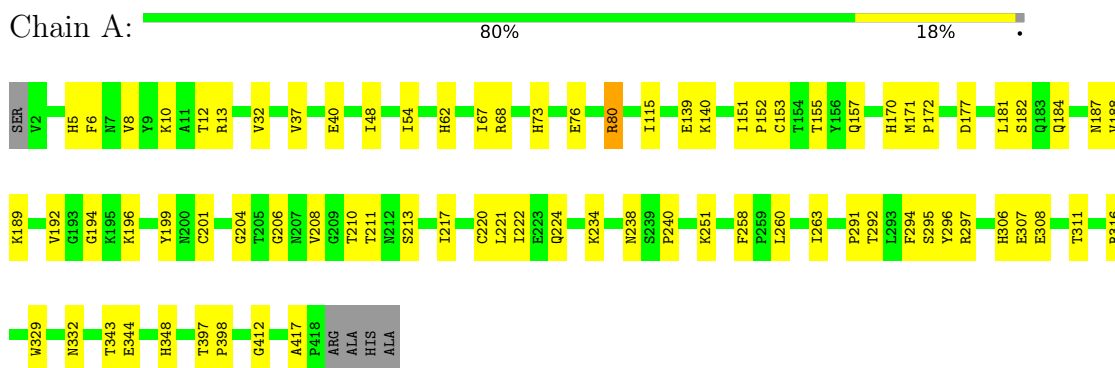
- Molecule 6 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		AltConf
6	M	1	Total 1	Ca 1	0
6	N	1	Total 1	Ca 1	0
6	O	1	Total 1	Ca 1	0

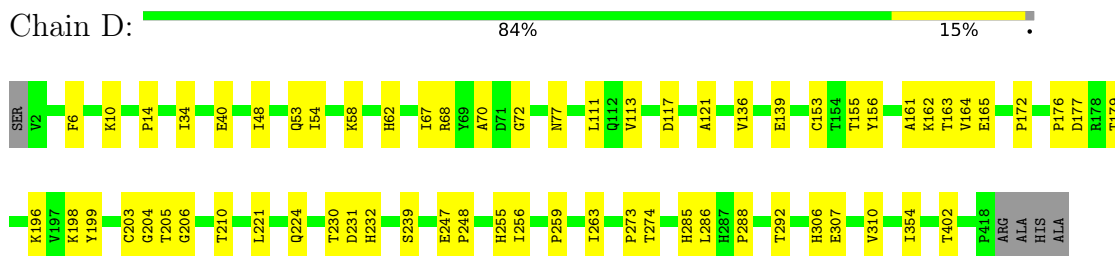
3 Residue-property plots [i](#)

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

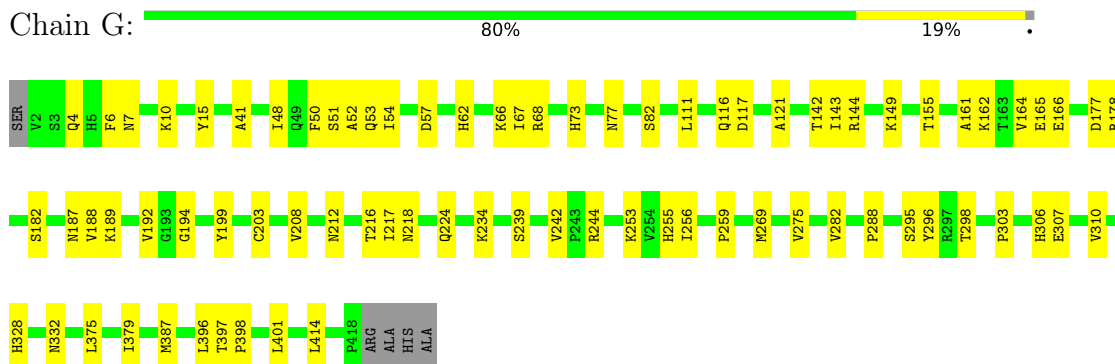
- Molecule 1: Spike glycoprotein E2




- Molecule 1: Spike glycoprotein E2

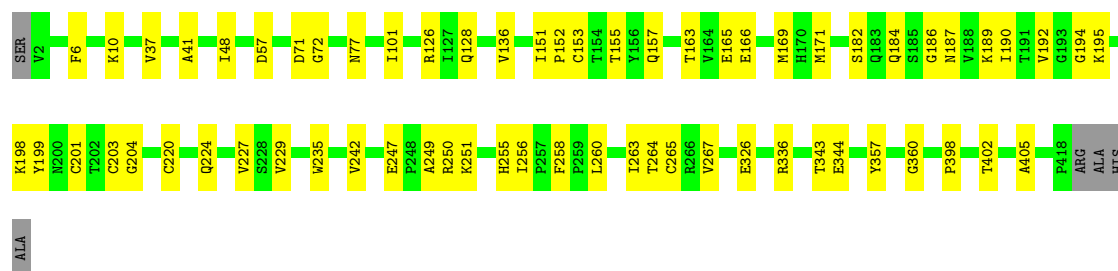


- Molecule 1: Spike glycoprotein E2



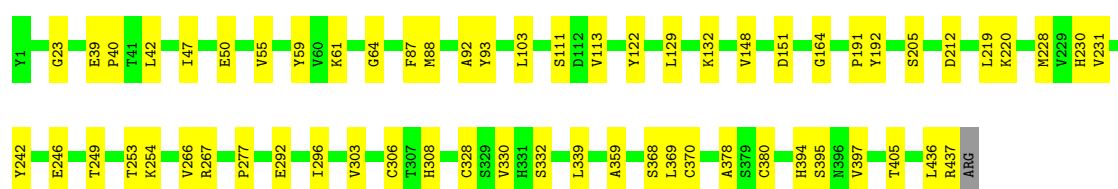
- Molecule 1: Spike glycoprotein E2

Chain J:  84% 15%




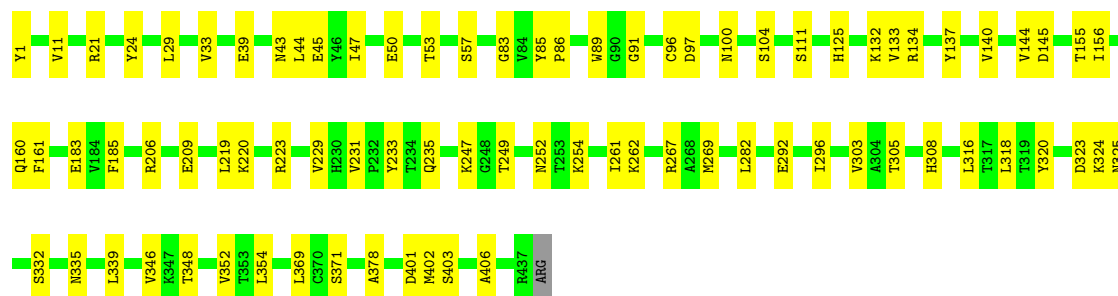
- Molecule 2: Spike glycoprotein E1

Chain B:  86% 14%

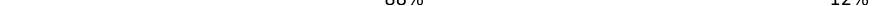


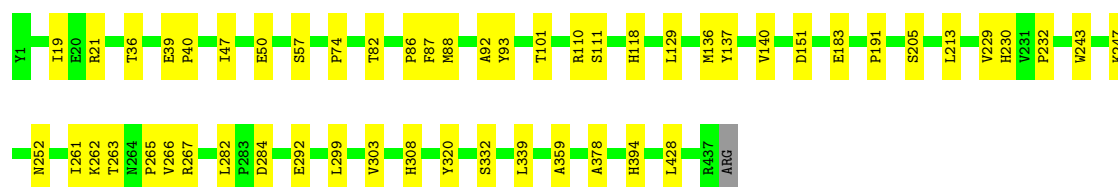
- Molecule 2: Spike glycoprotein E1

Chain E:  81% 18%




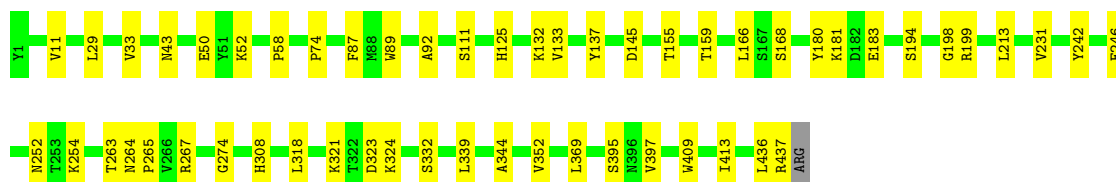
- Molecule 2: Spike glycoprotein E1

Chain H:  88% 12%

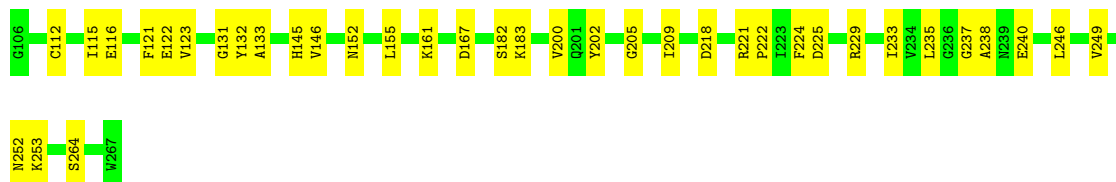
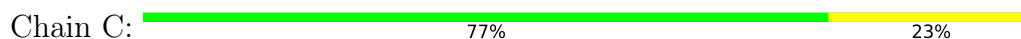


- Molecule 2: Spike glycoprotein E1

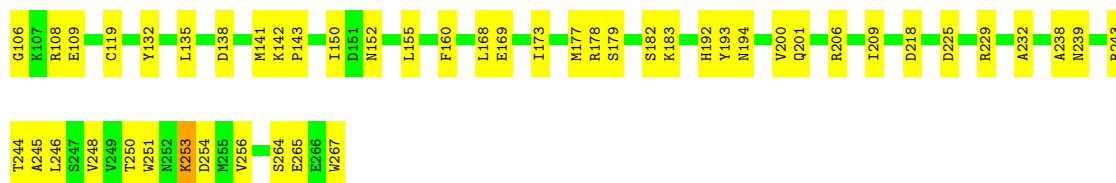
Chain K:  87% 12%



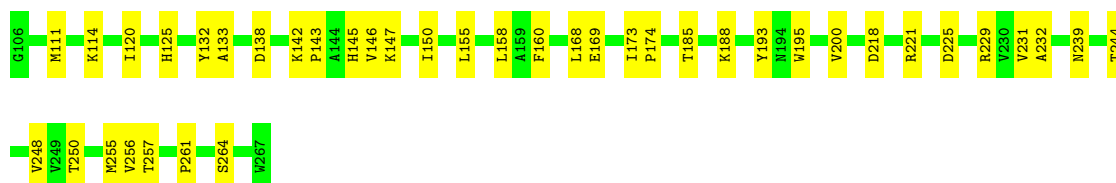
• Molecule 3: Capsid protein



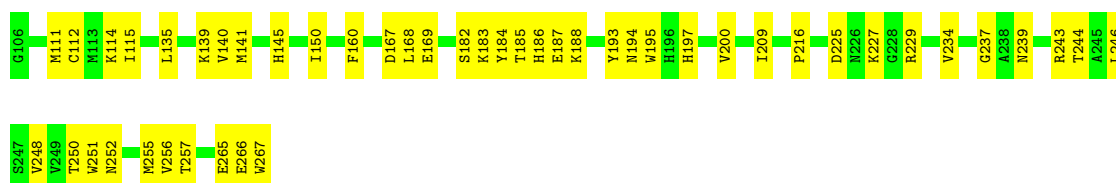
• Molecule 3: Capsid protein



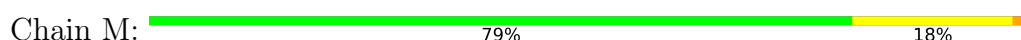
• Molecule 3: Capsid protein



• Molecule 3: Capsid protein

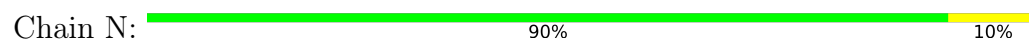


• Molecule 4: Very low-density lipoprotein receptor

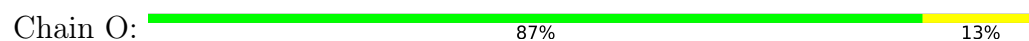




- Molecule 4: Very low-density lipoprotein receptor



- Molecule 4: Very low-density lipoprotein receptor



4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	651086	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 ($e^-/\text{\AA}^2$)	50	Depositor
Minimum defocus (nm)	1200	Depositor
Maximum defocus (nm)	1700	Depositor
Magnification	Not provided	
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: CA, NAG

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 >5	RMSZ	# Z >5
1	A	0.34	0/3343	0.50	0/4561
1	D	0.34	0/3343	0.49	0/4561
1	G	0.33	0/3343	0.49	0/4561
1	J	0.34	0/3343	0.50	0/4561
2	B	0.37	0/3404	0.49	0/4646
2	E	0.36	0/3404	0.49	0/4646
2	H	0.37	0/3404	0.49	0/4646
2	K	0.35	0/3404	0.50	0/4646
3	C	0.27	0/1278	0.45	0/1719
3	F	0.28	0/1278	0.48	0/1719
3	I	0.26	0/1278	0.47	0/1719
3	L	0.27	0/1278	0.46	0/1719
4	M	0.25	0/302	0.47	0/407
4	N	0.24	0/302	0.41	0/407
4	O	0.23	0/302	0.47	0/407
All	All	0.34	0/33006	0.49	0/44925

There are no bond length outliers.

There are no bond angle outliers.

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	3250	0	3160	54	0
1	D	3250	0	3160	46	0
1	G	3250	0	3160	49	0
1	J	3250	0	3160	52	0
2	B	3318	0	3236	44	0
2	E	3318	0	3234	54	0
2	H	3318	0	3234	32	0
2	K	3318	0	3234	37	0
3	C	1250	0	1232	22	0
3	F	1250	0	1232	31	0
3	I	1250	0	1232	26	0
3	L	1250	0	1232	31	0
4	M	298	0	249	5	0
4	N	298	0	249	2	0
4	O	298	0	249	5	0
5	A	28	0	26	1	0
5	B	14	0	13	0	0
5	D	28	0	26	0	0
5	E	14	0	13	0	0
5	G	28	0	26	0	0
5	H	14	0	13	0	0
5	J	28	0	26	0	0
5	K	14	0	13	0	0
6	M	1	0	0	0	0
6	N	1	0	0	0	0
6	O	1	0	0	0	0
All	All	32337	0	31409	467	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 7.

The worst 5 of 467 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)
2:B:296:ILE:HG21	2:B:369:LEU:HD23	1.53	0.90
2:B:296:ILE:HG13	2:B:369:LEU:HD23	1.58	0.85
1:D:155:THR:CG2	1:D:263:ILE:HD13	2.08	0.84
1:D:6:PHE:HA	1:D:62:HIS:HE1	1.43	0.82
1:J:402:THR:HB	1:J:405:ALA:HB2	1.63	0.80

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	415/422 (98%)	387 (93%)	28 (7%)	0	100	100
1	D	415/422 (98%)	384 (92%)	31 (8%)	0	100	100
1	G	415/422 (98%)	379 (91%)	36 (9%)	0	100	100
1	J	415/422 (98%)	374 (90%)	41 (10%)	0	100	100
2	B	435/438 (99%)	406 (93%)	29 (7%)	0	100	100
2	E	435/438 (99%)	406 (93%)	29 (7%)	0	100	100
2	H	435/438 (99%)	405 (93%)	30 (7%)	0	100	100
2	K	435/438 (99%)	405 (93%)	30 (7%)	0	100	100
3	C	160/162 (99%)	152 (95%)	8 (5%)	0	100	100
3	F	160/162 (99%)	150 (94%)	10 (6%)	0	100	100
3	I	160/162 (99%)	151 (94%)	9 (6%)	0	100	100
3	L	160/162 (99%)	151 (94%)	9 (6%)	0	100	100
4	M	37/39 (95%)	34 (92%)	3 (8%)	0	100	100
4	N	37/39 (95%)	35 (95%)	2 (5%)	0	100	100
4	O	37/39 (95%)	33 (89%)	4 (11%)	0	100	100
All	All	4151/4205 (99%)	3852 (93%)	299 (7%)	0	100	100

There are no Ramachandran outliers to report.

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	357/360 (99%)	356 (100%)	1 (0%)	91	96
1	D	357/360 (99%)	356 (100%)	1 (0%)	91	96
1	G	357/360 (99%)	356 (100%)	1 (0%)	91	96
1	J	357/360 (99%)	357 (100%)	0	100	100
2	B	371/372 (100%)	371 (100%)	0	100	100
2	E	371/372 (100%)	371 (100%)	0	100	100
2	H	371/372 (100%)	371 (100%)	0	100	100
2	K	371/372 (100%)	371 (100%)	0	100	100
3	C	132/132 (100%)	132 (100%)	0	100	100
3	F	132/132 (100%)	131 (99%)	1 (1%)	79	90
3	I	132/132 (100%)	132 (100%)	0	100	100
3	L	132/132 (100%)	131 (99%)	1 (1%)	79	90
4	M	35/35 (100%)	34 (97%)	1 (3%)	37	70
4	N	35/35 (100%)	34 (97%)	1 (3%)	37	70
4	O	35/35 (100%)	35 (100%)	0	100	100
All	All	3545/3561 (100%)	3538 (100%)	7 (0%)	91	97

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	144	ARG
3	L	243	ARG
4	N	133	ARG
4	M	133	ARG
3	F	253	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
4	O	126	GLN
1	J	116	GLN
2	E	230	HIS
1	D	62	HIS
1	G	187	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 oligosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 15 ligands modelled in this entry, 3 are monoatomic - leaving 12 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
5	NAG	J	501	1	14,14,15	0.24	0	17,19,21	0.58	0
5	NAG	D	501	1	14,14,15	0.37	0	17,19,21	0.56	0
5	NAG	K	500	2	14,14,15	0.17	0	17,19,21	0.49	0
5	NAG	B	500	2	14,14,15	0.20	0	17,19,21	0.46	0
5	NAG	A	501	1	14,14,15	0.38	0	17,19,21	1.26	1 (5%)
5	NAG	D	500	1	14,14,15	0.21	0	17,19,21	0.45	0
5	NAG	A	500	1	14,14,15	0.21	0	17,19,21	0.45	0
5	NAG	J	500	1	14,14,15	0.16	0	17,19,21	0.44	0
5	NAG	E	500	2	14,14,15	0.39	0	17,19,21	0.34	0
5	NAG	G	500	1	14,14,15	0.25	0	17,19,21	0.43	0
5	NAG	G	501	1	14,14,15	0.26	0	17,19,21	0.53	0
5	NAG	H	500	2	14,14,15	0.48	0	17,19,21	0.37	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	NAG	J	501	1	-	3/6/23/26	0/1/1/1
5	NAG	D	501	1	-	0/6/23/26	0/1/1/1
5	NAG	K	500	2	-	2/6/23/26	0/1/1/1
5	NAG	B	500	2	-	2/6/23/26	0/1/1/1
5	NAG	A	501	1	-	5/6/23/26	0/1/1/1
5	NAG	D	500	1	-	1/6/23/26	0/1/1/1
5	NAG	A	500	1	-	2/6/23/26	0/1/1/1
5	NAG	J	500	1	-	2/6/23/26	0/1/1/1
5	NAG	E	500	2	-	0/6/23/26	0/1/1/1
5	NAG	G	500	1	-	2/6/23/26	0/1/1/1
5	NAG	G	501	1	-	2/6/23/26	0/1/1/1
5	NAG	H	500	2	-	1/6/23/26	0/1/1/1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	501	NAG	C2-N2-C7	4.30	129.03	122.90

There are no chirality outliers.

5 of 22 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	B	500	NAG	C4-C5-C6-O6
5	G	500	NAG	O5-C5-C6-O6
5	K	500	NAG	O5-C5-C6-O6
5	B	500	NAG	O5-C5-C6-O6
5	A	500	NAG	O5-C5-C6-O6

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	501	NAG	1	0

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.