WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB EM Validation Summary Report (i)

Jun 11, 2024 – 04:06 PM JST

PDB ID	:	8WRL
EMDB ID	:	EMD-37779
Title	:	XBB.1.5 RBD in complex with ACE2
Authors	:	Feng, L.L.; Feng, L.L.
Deposited on	:	2023-10-15
Resolution	:	3.36 Å(reported)

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

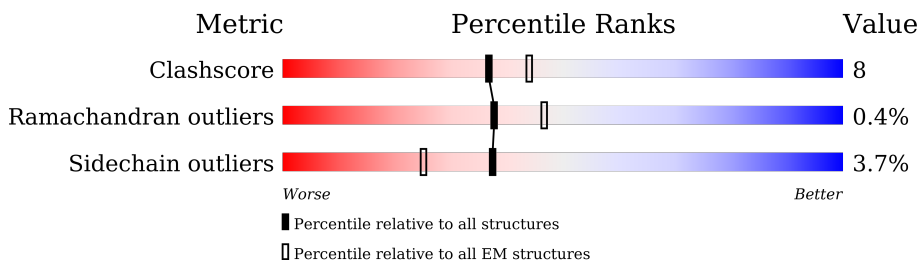|		
EMDB validation analysis	:	FAILED
Mogul	:	1.8.5 (274361), CSD as541be (2020)
MolProbity	:	4.02b-467
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	FAILED
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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 3.36 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	Percentile Ranks	Value
Clashscore		8
Ramachandran outliers		0.4%
Sidechain outliers		3.7%

Worse　　　　　　　　　　　　　　　*Better*

▊ Percentile relative to all structures
▯ Percentile relative to all EM structures

Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826

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 <=5%

Mol	Chain	Length	Quality of chain
1	A	594	76%　　　　　　　22%　　•
2	B	238	71%　　　11%　　18%

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 6473 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 Processed angiotensin-converting enzyme 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	594	Total 4845	C 3099	N 803	O 914	S 29	0	0

- Molecule 2 is a protein called Spike protein S1.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	196	Total 1558	C 1006	N 262	O 282	S 8	0	0

There are 41 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	300	MET	-	initiating methionine	UNP P0DTC2
B	301	PRO	-	expression tag	UNP P0DTC2
B	302	LEU	-	expression tag	UNP P0DTC2
B	303	LEU	-	expression tag	UNP P0DTC2
B	304	LEU	-	expression tag	UNP P0DTC2
B	305	LEU	-	expression tag	UNP P0DTC2
B	306	LEU	-	expression tag	UNP P0DTC2
B	307	PRO	-	expression tag	UNP P0DTC2
B	308	LEU	-	expression tag	UNP P0DTC2
B	309	LEU	-	expression tag	UNP P0DTC2
B	310	TRP	-	expression tag	UNP P0DTC2
B	311	ALA	-	expression tag	UNP P0DTC2
B	312	GLY	-	expression tag	UNP P0DTC2
B	313	ALA	-	expression tag	UNP P0DTC2
B	314	LEU	-	expression tag	UNP P0DTC2
B	315	ALA	-	expression tag	UNP P0DTC2
B	316	MET	-	expression tag	UNP P0DTC2
B	317	ALA	-	expression tag	UNP P0DTC2
B	318	ALA	-	expression tag	UNP P0DTC2
B	339	HIS	GLY	variant	UNP P0DTC2
B	346	THR	ARG	variant	UNP P0DTC2

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	368	ILE	LEU	variant	UNP P0DTC2
B	371	PHE	SER	variant	UNP P0DTC2
B	373	PRO	SER	variant	UNP P0DTC2
B	375	PHE	SER	variant	UNP P0DTC2
B	376	ALA	THR	variant	UNP P0DTC2
B	405	ASN	ASP	variant	UNP P0DTC2
B	408	SER	ARG	variant	UNP P0DTC2
B	417	ASN	LYS	variant	UNP P0DTC2
B	440	LYS	ASN	variant	UNP P0DTC2
B	445	PRO	VAL	variant	UNP P0DTC2
B	446	SER	GLY	variant	UNP P0DTC2
B	460	LYS	ASN	variant	UNP P0DTC2
B	477	ASN	SER	variant	UNP P0DTC2
B	478	LYS	THR	variant	UNP P0DTC2
B	484	ALA	GLU	variant	UNP P0DTC2
B	486	PRO	PHE	variant	UNP P0DTC2
B	490	SER	PHE	variant	UNP P0DTC2
B	498	ARG	GLN	variant	UNP P0DTC2
B	501	TYR	ASN	variant	UNP P0DTC2
B	505	HIS	TYR	variant	UNP P0DTC2

- Molecule 3 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$) (labeled as "Ligand of Interest" by depositor).
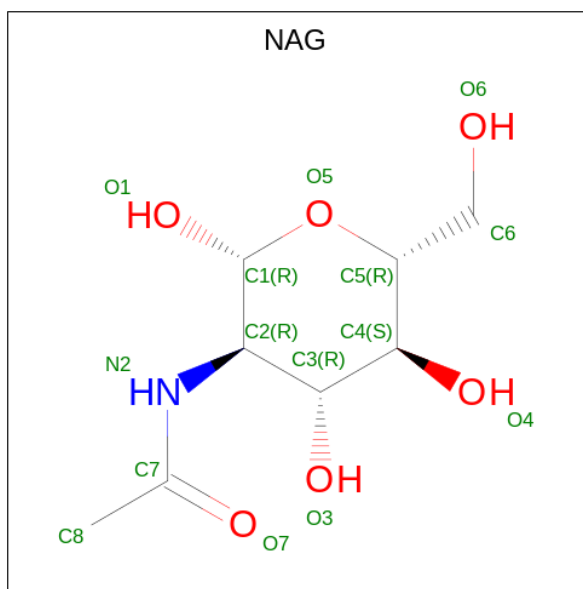
Mol	Chain	Residues	Atoms				AltConf
			Total	C	N	O	
3	A	1	14	8	1	5	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				AltConf
3	A	1	Total	C	N	O	0
			14	8	1	5	
3	A	1	Total	C	N	O	0
			14	8	1	5	
3	A	1	Total	C	N	O	0
			14	8	1	5	
3	B	1	Total	C	N	O	0
			14	8	1	5	

# 3   Residue-property plots ⓘ

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: Processed angiotensin-converting enzyme 2

Chain A:
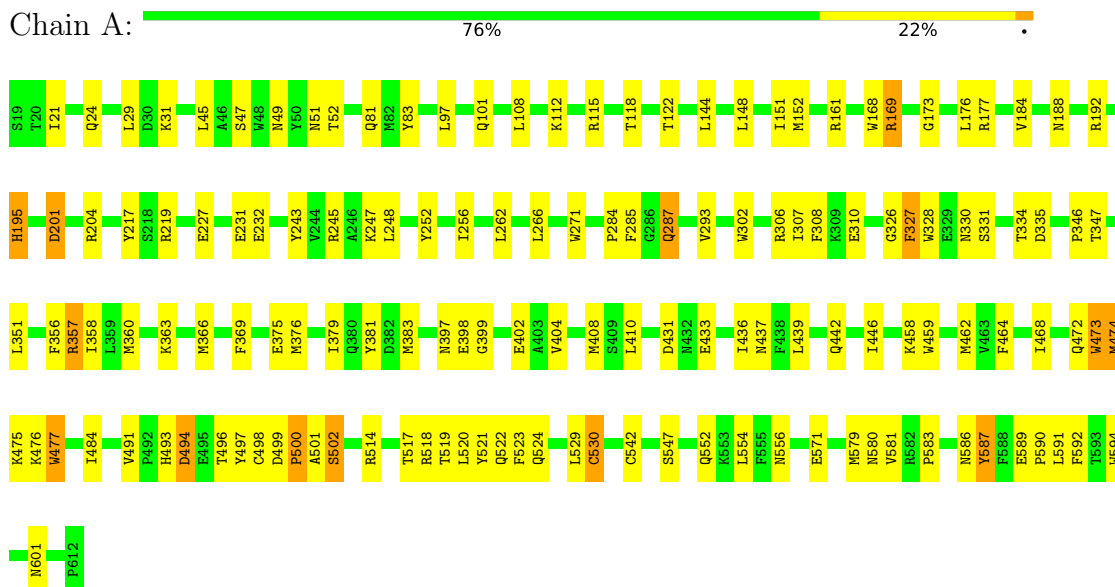
• Molecule 2: Spike protein S1

Chain B:
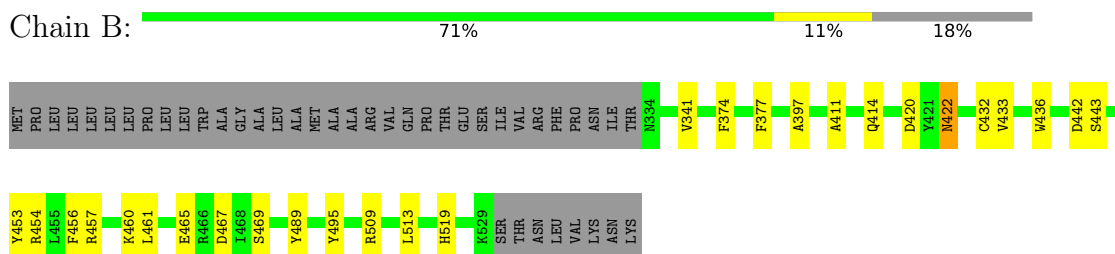
4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	226753	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 $(e^-/\text{Å}^2)$	52	Depositor
Minimum defocus (nm)	1200	Depositor
Maximum defocus (nm)	1800	Depositor
Magnification	Not provided	
Image detector	FEI FALCON IV (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: 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	0.35	1/4981 (0.0%)	0.57	1/6767 (0.0%)
2	B	0.31	0/1607	0.56	0/2188
All	All	0.34	1/6588 (0.0%)	0.56	1/8955 (0.0%)

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	2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	331	SER	CA-CB	-5.24	1.45	1.52

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	497	TYR	CB-CA-C	-5.16	100.08	110.40

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	169	ARG	Sidechain
1	A	357	ARG	Sidechain

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	4845	0	4621	91	0
2	B	1558	0	1493	15	0
3	A	56	0	52	0	0
3	B	14	0	13	0	0
All	All	6473	0	6179	105	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 8.

The worst 5 of 105 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:A:168:TRP:HE1	1:A:502:SER:HB2	1.36	0.90
1:A:472:GLN:HG2	1:A:475:LYS:HD2	1.68	0.76
1:A:379:ILE:O	1:A:383:MET:HG3	1.92	0.70
1:A:446:ILE:HG21	1:A:523:PHE:HE2	1.61	0.65
1:A:474:MET:HE3	1:A:498:CYS:HA	1.80	0.63

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	592/594 (100%)	558 (94%)	31 (5%)	3 (0%)	29	63
2	B	194/238 (82%)	171 (88%)	23 (12%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	786/832 (94%)	729 (93%)	54 (7%)	3 (0%)	38 68

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	502	SER
1	A	360	MET
1	A	496	THR

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	525/525 (100%)	505 (96%)	20 (4%)	33 63
2	B	169/205 (82%)	163 (96%)	6 (4%)	35 64
All	All	694/730 (95%)	668 (96%)	26 (4%)	37 63

5 of 26 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	500	PRO
1	A	571	GLU
2	B	489	TYR
1	A	547	SER
1	A	580	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
2	B	422	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 monosaccharides in this entry.

5.6 Ligand geometry ⓘ

5 ligands are modelled in this entry.

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	$\#|Z| > 2$
3	NAG	B	601	2	14,14,15	0.25	0	17,19,21	0.45	0
3	NAG	A	701	1	14,14,15	0.26	0	17,19,21	0.48	0
3	NAG	A	703	1	14,14,15	0.41	0	17,19,21	0.38	0
3	NAG	A	704	1	14,14,15	0.36	0	17,19,21	0.36	0
3	NAG	A	702	1	14,14,15	0.32	0	17,19,21	0.47	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
3	NAG	B	601	2	-	1/6/23/26	0/1/1/1
3	NAG	A	701	1	-	0/6/23/26	0/1/1/1
3	NAG	A	703	1	-	2/6/23/26	0/1/1/1
3	NAG	A	704	1	-	2/6/23/26	0/1/1/1
3	NAG	A	702	1	-	2/6/23/26	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.
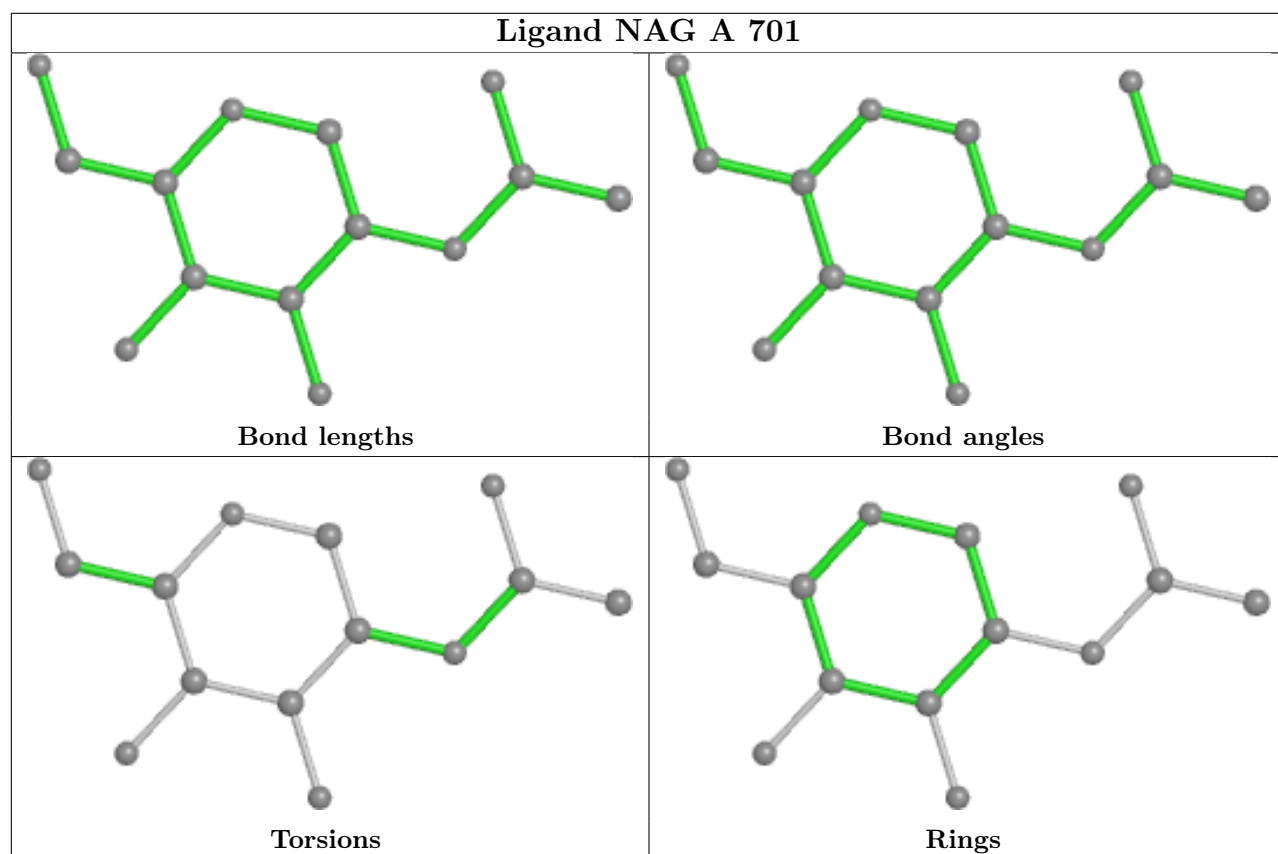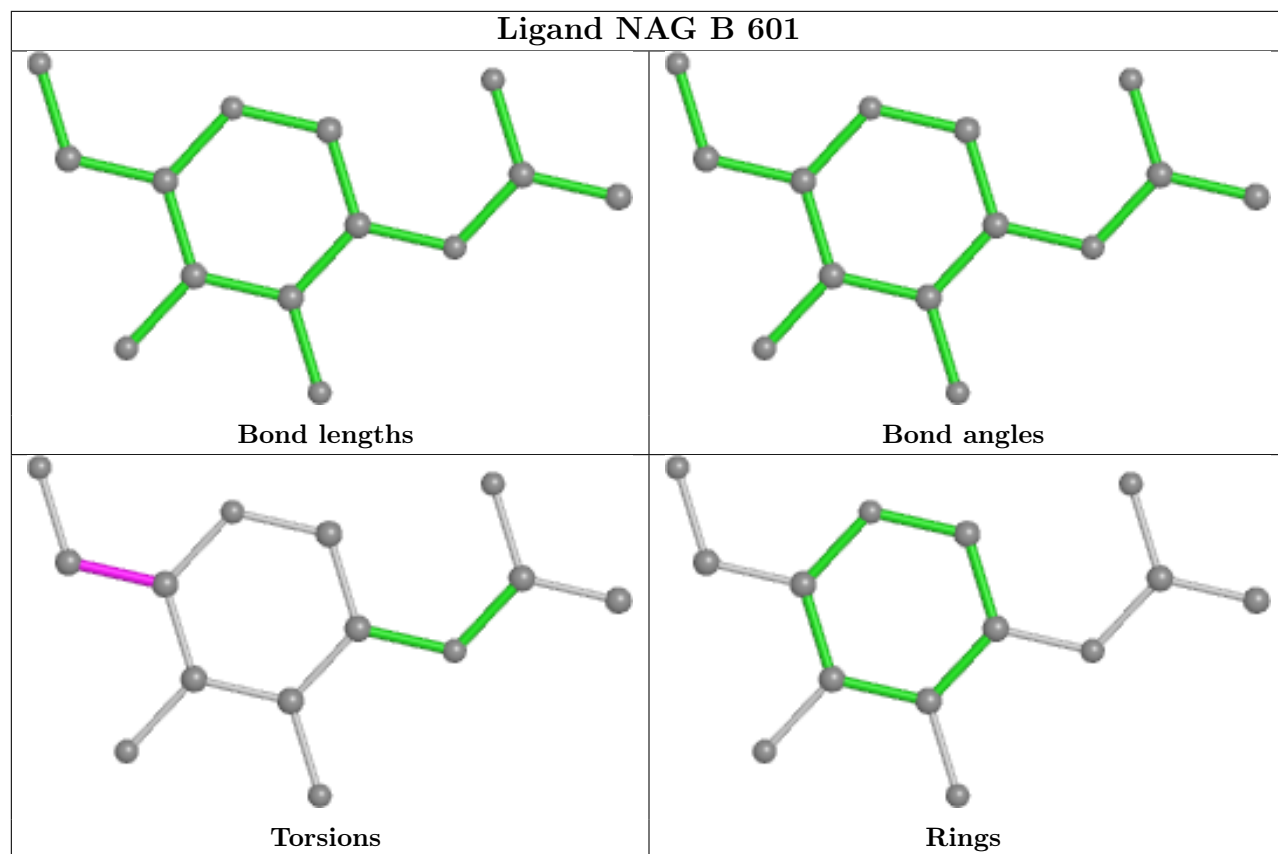
There are no chirality outliers.
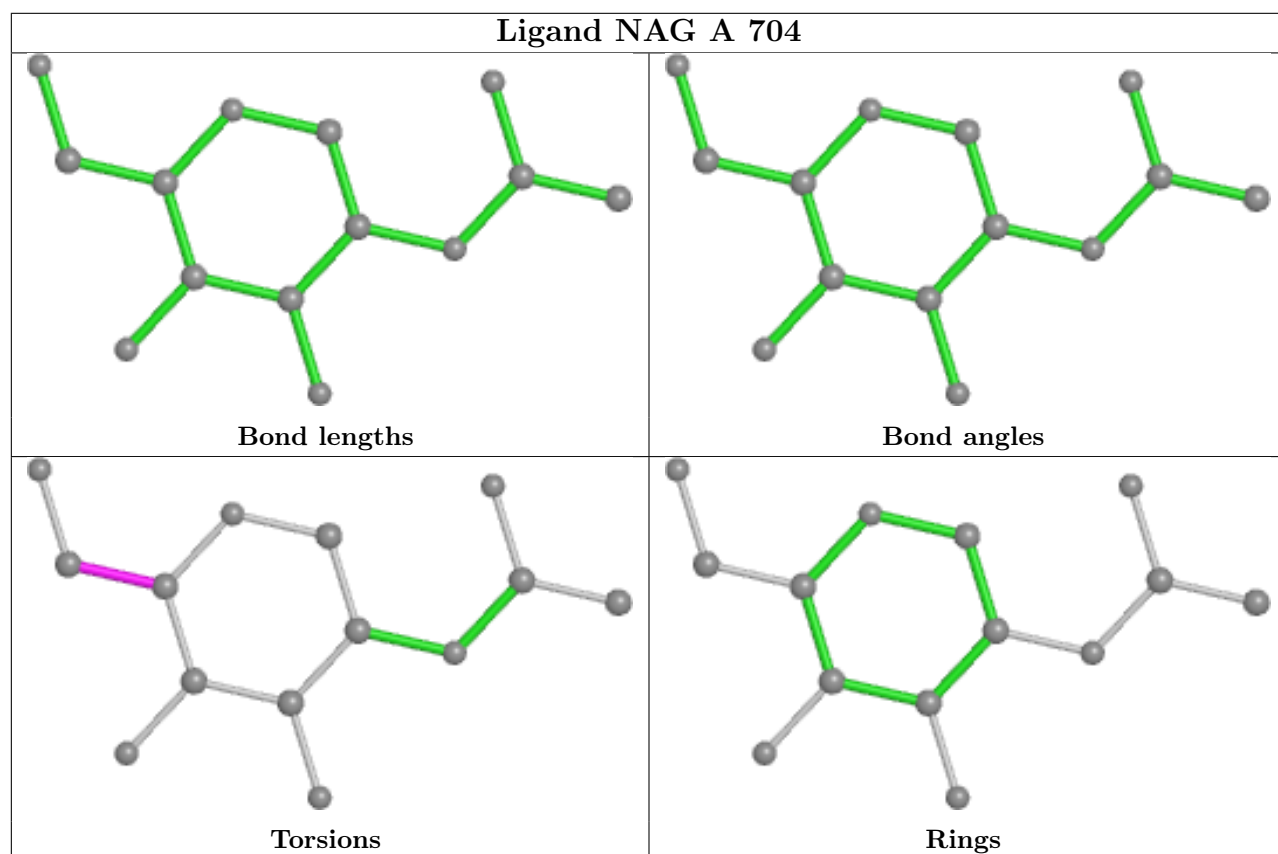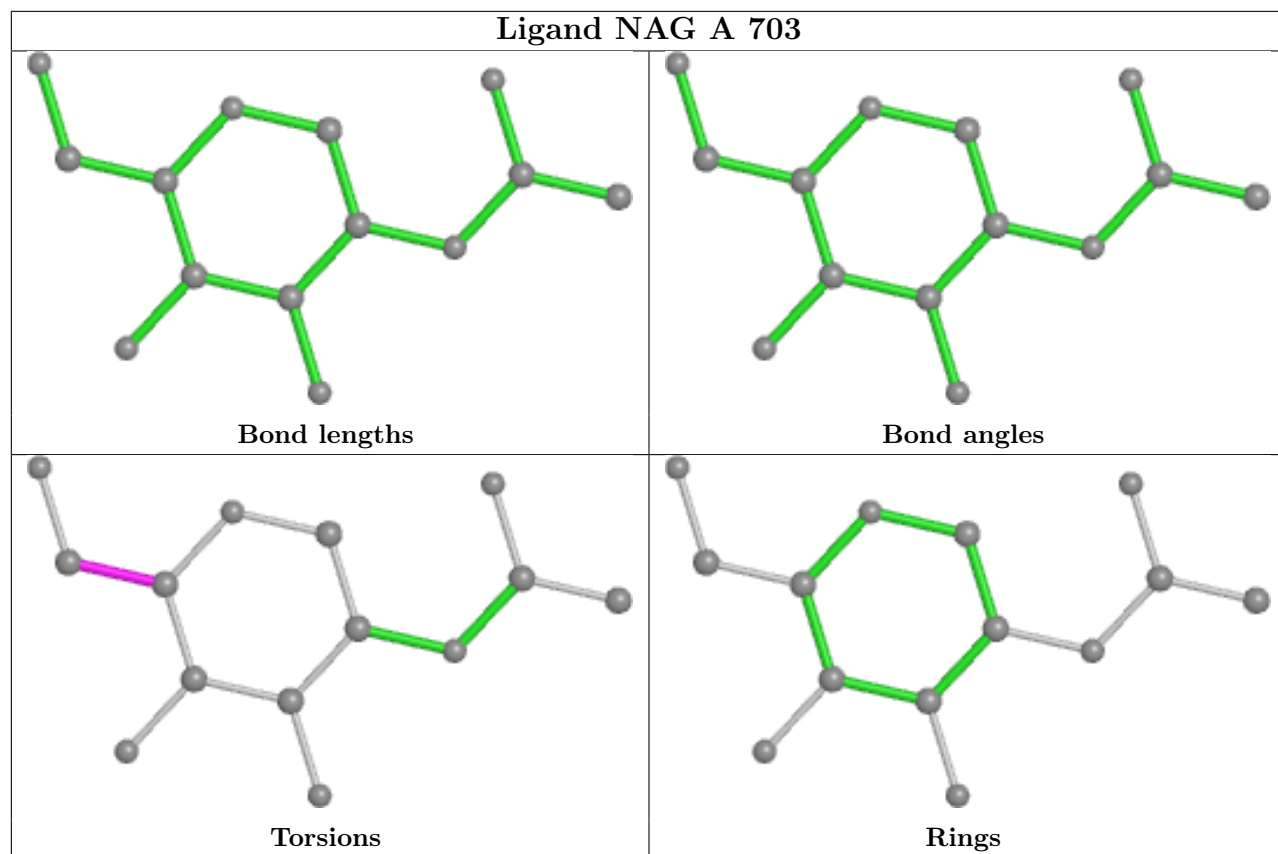
5 of 7 torsion outliers are listed below:

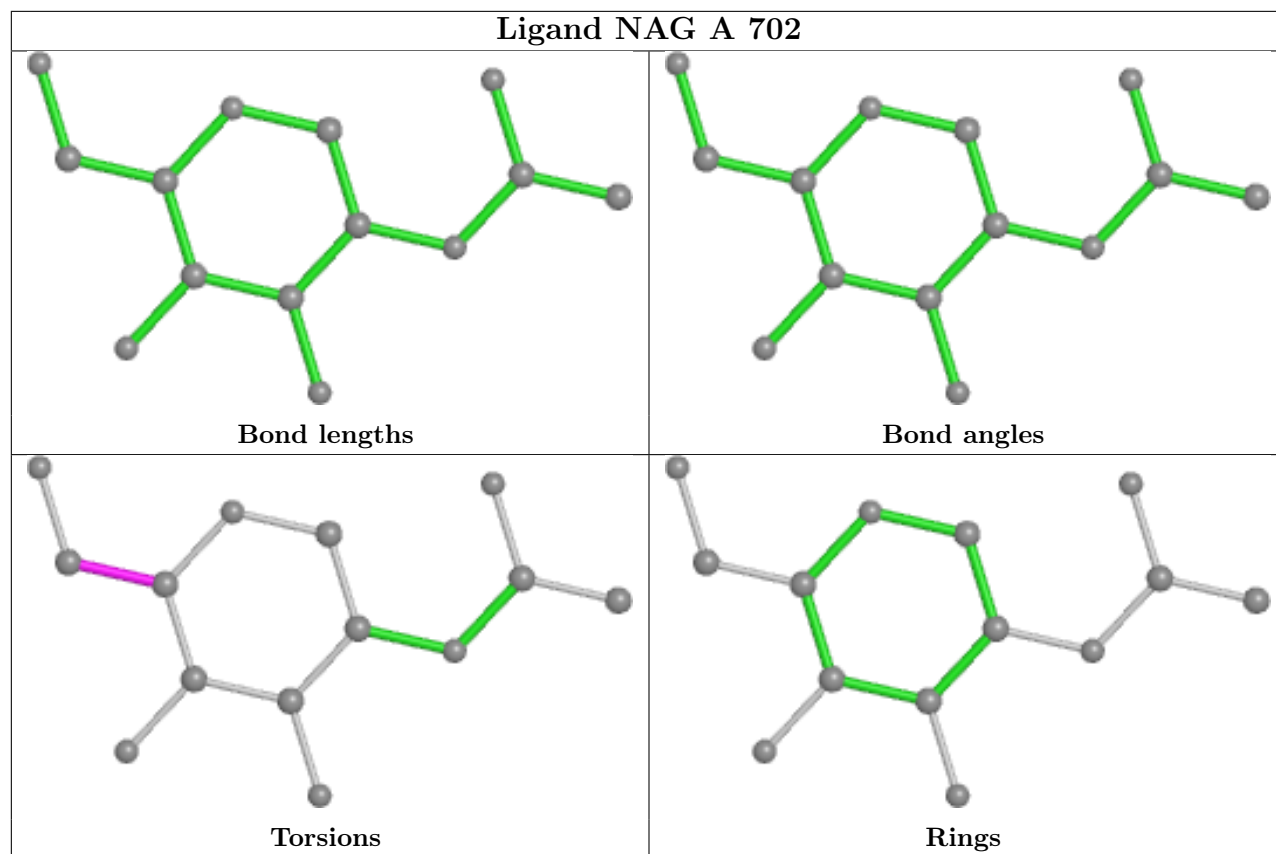| Mol	Chain	Res	Type	Atoms
3	A	702	NAG	O5-C5-C6-O6
3	A	704	NAG	O5-C5-C6-O6
3	A	702	NAG	C4-C5-C6-O6
3	A	704	NAG	C4-C5-C6-O6
3	A	703	NAG	C4-C5-C6-O6

There are no ring outliers.

No monomer is involved in short contacts.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

Ligand NAG B 601

Bond lengths

Bond angles

Torsions

Rings

Ligand NAG A 701

Bond lengths

Bond angles

Torsions

Rings

Ligand NAG A 703

Bond lengths

Bond angles

Torsions

Rings

Ligand NAG A 704

Bond lengths

Bond angles

Torsions

Rings

Ligand NAG A 702

Bond lengths

Bond angles

Torsions

Rings

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.