



wwPDB/EMDataBank EM Map/Model Validation Summary Report ⓘ

Sep 7, 2019 – 02:07 PM EDT

PDB ID : 6O7E
EMDB ID: : EMD-0640
Title : Cryo-EM structure of Csm-crRNA-target RNA ternary complex in complex with AMPPNP in type III-A CRISPR-Cas system
Authors : Jia, N.; Patel, D.J.
Deposited on : 2019-03-07
Resolution : 3.20 Å(reported)

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

<https://www.wwpdb.org/validation/2017/EMValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

MolProbity : 4.02b-467
Mogul : 1.8.0 (224370), CSD as540be (2019)
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : 2.4

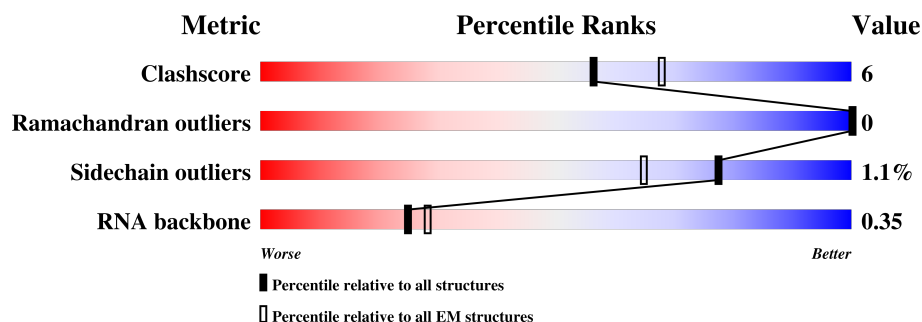
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 3.20 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	136327	1886
Ramachandran outliers	132723	1663
Sidechain outliers	132532	1531
RNA backbone	3747	458

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	791	86% 10% . .
2	B	187	77% 8% 15%
3	C	291	77% 18% .
3	D	291	84% 14% .
4	E	289	89% 9% . .
5	G	38	37% 29% 5% 29%
6	H	40	15% 23% 10% 53%
7	Y	378	79% 16% .

2 Entry composition

There are 10 unique types of molecules in this entry. The entry contains 17606 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 Csm1.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	771	Total	C	N	O	S	0	0
			6150	3941	1072	1118	19		

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-13	MET	-	initiating methionine	UNP B6YWB8
A	-12	GLY	-	expression tag	UNP B6YWB8
A	-11	SER	-	expression tag	UNP B6YWB8
A	-10	SER	-	expression tag	UNP B6YWB8
A	-9	HIS	-	expression tag	UNP B6YWB8
A	-8	HIS	-	expression tag	UNP B6YWB8
A	-7	HIS	-	expression tag	UNP B6YWB8
A	-6	HIS	-	expression tag	UNP B6YWB8
A	-5	HIS	-	expression tag	UNP B6YWB8
A	-4	HIS	-	expression tag	UNP B6YWB8
A	-3	SER	-	expression tag	UNP B6YWB8
A	-2	GLN	-	expression tag	UNP B6YWB8
A	-1	ASP	-	expression tag	UNP B6YWB8
A	0	PRO	-	expression tag	UNP B6YWB8
A	14	ALA	HIS	engineered mutation	UNP B6YWB8
A	15	ASN	ASP	engineered mutation	UNP B6YWB8

- Molecule 2 is a protein called Csm2.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	159	Total	C	N	O	S	0	0
			1128	714	202	208	4		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	0	SER	-	expression tag	UNP B6YWB9

- Molecule 3 is a protein called Csm3.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	C	278	Total	C	N	O	S	0	0
			2214	1406	389	413	6		
3	D	284	Total	C	N	O	S	0	0
			2262	1434	398	423	7		

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	0	SER	-	expression tag	UNP B6YWC0
D	0	SER	-	expression tag	UNP B6YWC0

- Molecule 4 is a protein called Csm4.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	E	286	Total	C	N	O	S	0	0
			2261	1467	380	410	4		

- Molecule 5 is a RNA chain called RNA (38-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
5	G	27	Total	C	N	O	P	0	0
			589	262	115	186	26		

- Molecule 6 is a RNA chain called RNA (40-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
6	H	19	Total	C	N	O	P	0	0
			398	178	69	132	19		

- Molecule 7 is a protein called Csm5.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	Y	361	Total	C	N	O	S	0	0
			2536	1603	453	471	9		

- Molecule 8 is ZINC ION (three-letter code: ZN) (formula: Zn).

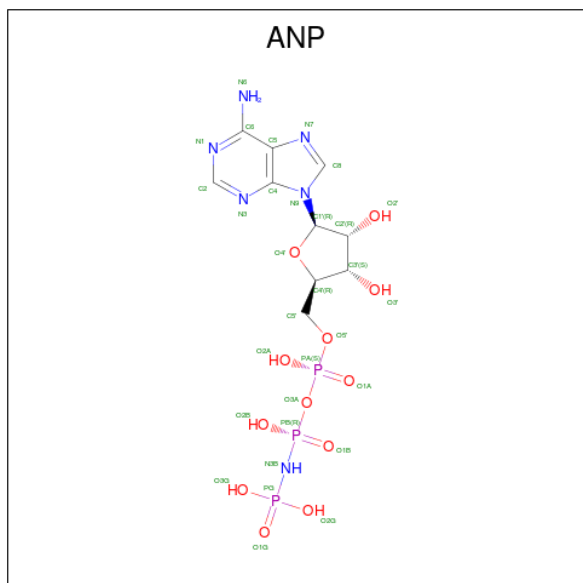
Mol	Chain	Residues	Atoms		AltConf
8	A	1	Total	Zn	0
			1	1	

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		AltConf
8	D	1	Total	Zn	0
			1	1	
8	C	1	Total	Zn	0
			1	1	

- Molecule 9 is PHOSPHOAMINOPHOSPHONIC ACID-ADENYLATE ESTER (three-letter code: ANP) (formula: $C_{10}H_{17}N_6O_{12}P_3$).



Mol	Chain	Residues	Atoms					AltConf
9	A	1	Total	C	N	O	P	0
			62	20	12	24	6	
9	A	1	Total	C	N	O	P	0
			62	20	12	24	6	

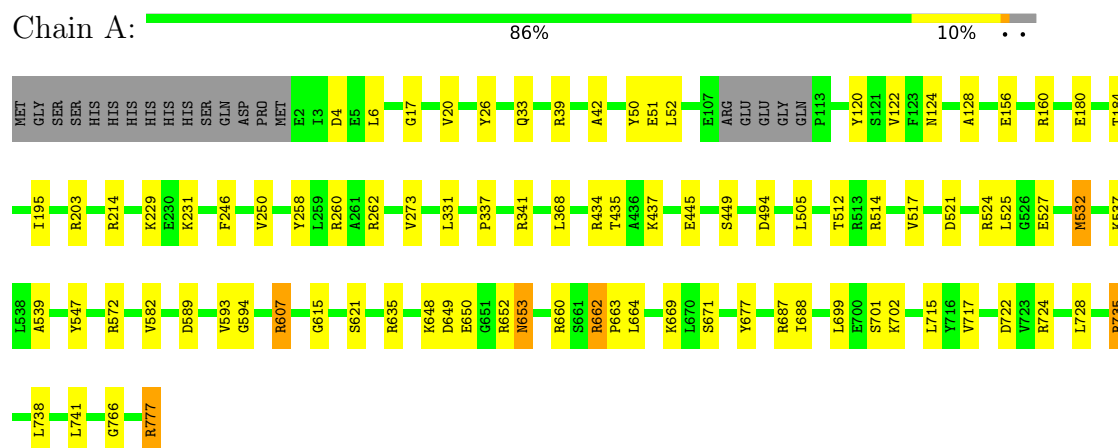
- Molecule 10 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		AltConf
10	A	3	Total	Mn	0
			3	3	

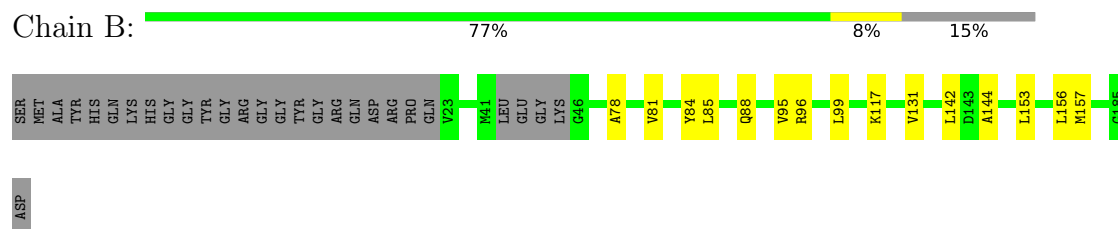
3 Residue-property plots

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.

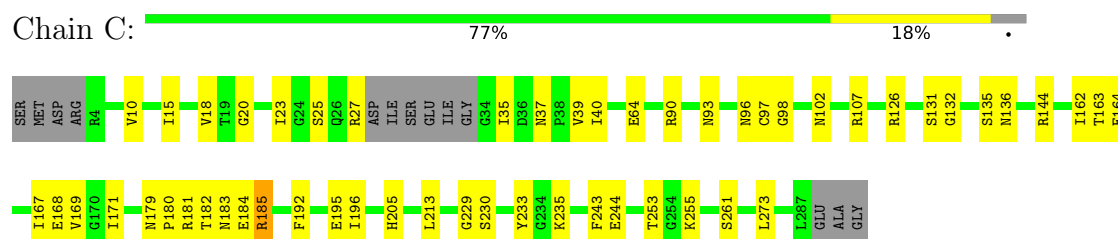
• Molecule 1: Csm1



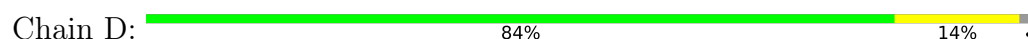
• Molecule 2: Csm2

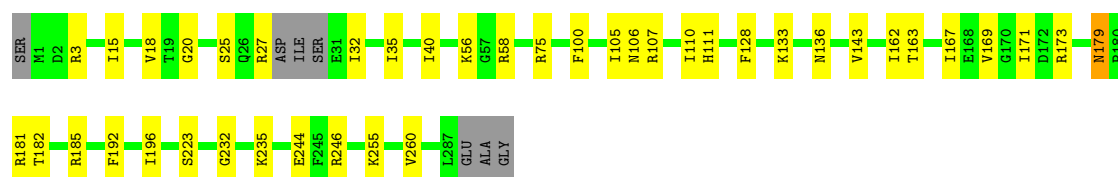


• Molecule 3: Csm3

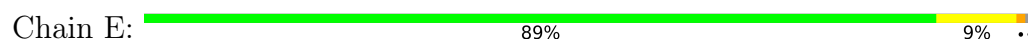


• Molecule 3: Csm3





• Molecule 4: Csm4



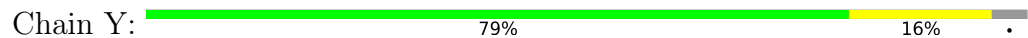
• Molecule 5: RNA (38-MER)



• Molecule 6: RNA (40-MER)



• Molecule 7: Csm5



4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of particles used	110194	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 ($e^-/\text{\AA}^2$)	1.35	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	Not provided	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor

5 Model quality [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, MN, ANP

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	0.29	0/6290	0.54	0/8489
2	B	0.25	0/1142	0.48	0/1540
3	C	0.29	0/2263	0.55	0/3051
3	D	0.28	0/2311	0.54	0/3114
4	E	0.30	0/2322	0.54	0/3146
5	G	0.33	0/661	0.84	0/1033
6	H	0.25	0/442	0.89	0/685
7	Y	0.27	0/2122	0.52	0/2875
All	All	0.29	0/17553	0.56	0/23933

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 [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	6150	0	6143	49	0
2	B	1128	0	1033	8	0
3	C	2214	0	2181	37	0
3	D	2262	0	2230	29	0
4	E	2261	0	2265	26	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	G	589	0	295	15	0
6	H	398	0	208	7	0
7	Y	2536	0	2142	51	0
8	A	1	0	0	0	0
8	C	1	0	0	0	0
8	D	1	0	0	0	0
9	A	62	0	26	3	0
10	A	3	0	0	0	0
All	All	17606	0	16523	190	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 6.

The worst 5 of 190 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)
7:Y:106:LEU:HD23	7:Y:111:TYR:CD1	1.91	1.04
7:Y:109:ARG:HH11	7:Y:218:GLY:HA2	1.23	1.04
7:Y:106:LEU:HD23	7:Y:111:TYR:HD1	1.23	0.93
7:Y:106:LEU:CD2	7:Y:111:TYR:CD1	2.54	0.91
7:Y:109:ARG:HB3	7:Y:109:ARG:CZ	2.11	0.79

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	767/791 (97%)	725 (94%)	42 (6%)	0	100	100
2	B	155/187 (83%)	140 (90%)	15 (10%)	0	100	100
3	C	274/291 (94%)	256 (93%)	18 (7%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	D	280/291 (96%)	264 (94%)	16 (6%)	0	100	100
4	E	284/289 (98%)	271 (95%)	13 (5%)	0	100	100
7	Y	259/378 (68%)	249 (96%)	10 (4%)	0	100	100
All	All	2019/2227 (91%)	1905 (94%)	114 (6%)	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	643/664 (97%)	635 (99%)	8 (1%)	74	90
2	B	95/154 (62%)	95 (100%)	0	100	100
3	C	237/248 (96%)	235 (99%)	2 (1%)	83	94
3	D	242/248 (98%)	238 (98%)	4 (2%)	63	86
4	E	237/240 (99%)	235 (99%)	2 (1%)	83	94
7	Y	212/245 (86%)	210 (99%)	2 (1%)	81	93
All	All	1666/1799 (93%)	1648 (99%)	18 (1%)	77	90

5 of 18 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	C	181	ARG
3	C	185	ARG
4	E	53	ARG
1	A	735	ARG
1	A	777	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
3	C	136	ASN
7	Y	18	ASN
3	D	106	ASN
1	A	668	HIS
3	C	179	ASN

5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
5	G	26/38 (68%)	5 (19%)	1 (3%)
6	H	18/40 (45%)	11 (61%)	0
All	All	44/78 (56%)	16 (36%)	1 (2%)

5 of 16 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
5	G	-7	U
5	G	5	C
5	G	6	A
5	G	9	G
5	G	16	U

All (1) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
5	G	5	C

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 8 ligands modelled in this entry, 6 are monoatomic - leaving 2 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	$\# Z > 2$
9	ANP	A	802	10	28,33,33	1.19	4 (14%)	29,52,52	1.09	2 (6%)
9	ANP	A	803	-	28,33,33	1.27	4 (14%)	29,52,52	1.24	3 (10%)

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
9	ANP	A	802	10	-	2/13/38/38	0/3/3/3
9	ANP	A	803	-	-	4/13/38/38	0/3/3/3

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
9	A	803	ANP	PG-O1G	3.38	1.50	1.46
9	A	803	ANP	PB-O1B	3.19	1.49	1.46
9	A	802	ANP	PG-O1G	3.14	1.49	1.46
9	A	802	ANP	PB-O1B	2.99	1.49	1.46
9	A	803	ANP	PB-O3A	-2.44	1.56	1.59

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
9	A	803	ANP	PA-O3A-PB	-4.04	118.33	132.46
9	A	802	ANP	PA-O3A-PB	-3.30	120.92	132.46
9	A	803	ANP	C4'-O4'-C1'	-2.44	107.28	109.83
9	A	803	ANP	C5-C6-N6	2.27	123.94	120.38
9	A	802	ANP	C5-C6-N6	2.14	123.75	120.38

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

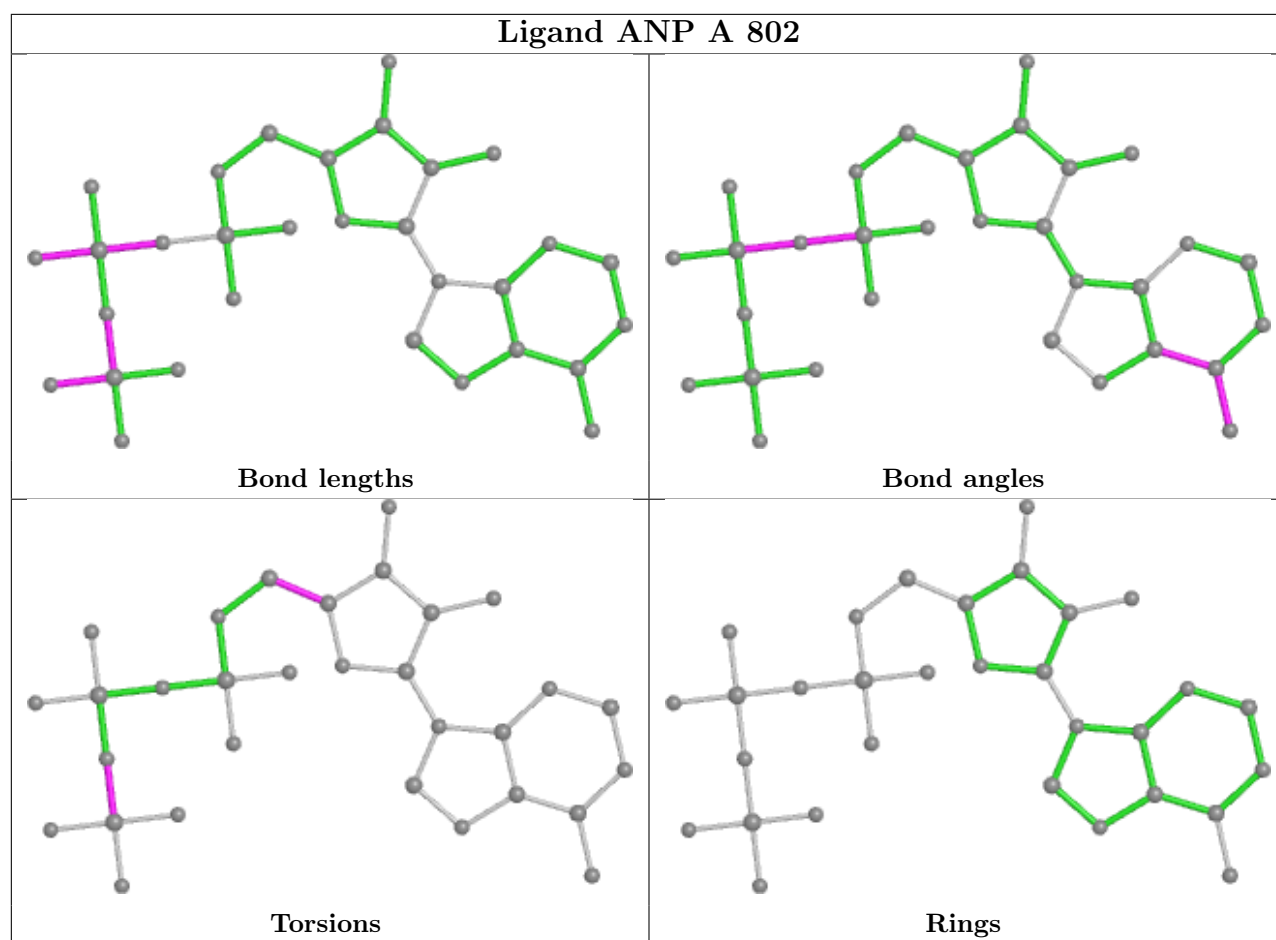
Mol	Chain	Res	Type	Atoms
9	A	802	ANP	PB-N3B-PG-O1G
9	A	803	ANP	PG-N3B-PB-O1B
9	A	803	ANP	O4'-C4'-C5'-O5'
9	A	803	ANP	C3'-C4'-C5'-O5'
9	A	803	ANP	PB-O3A-PA-O1A

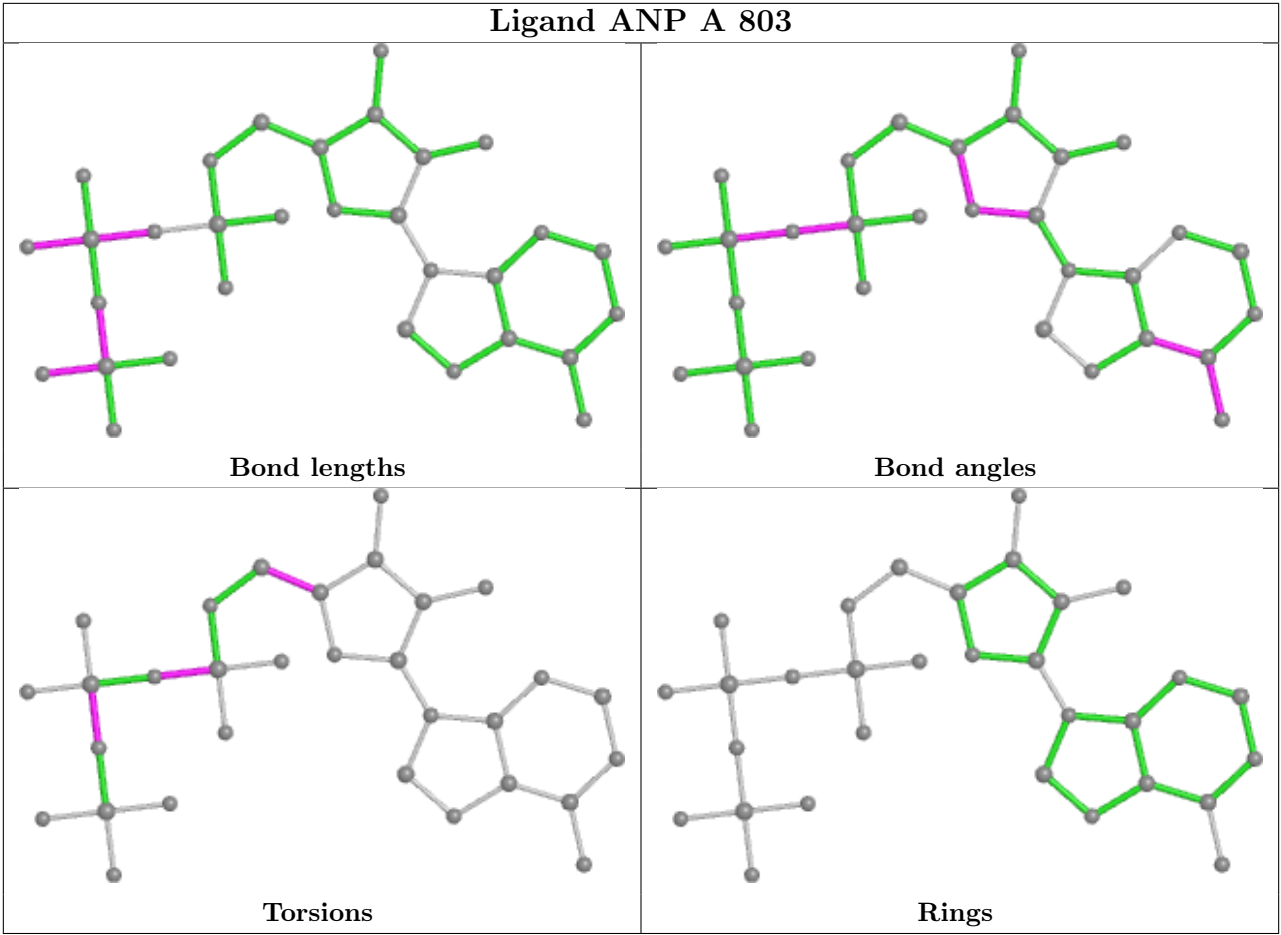
There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
9	A	802	ANP	1	0
9	A	803	ANP	2	0

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 than 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.





5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

The following chains have linkage breaks:

Mol	Chain	Number of breaks
7	Y	8

The worst 5 of 8 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	Y	169:UNK	C	175:UNK	N	17.56
1	Y	91:UNK	C	95:ARG	N	9.63
1	Y	191:UNK	C	197:ILE	N	7.06
1	Y	143:UNK	C	150:UNK	N	7.00
1	Y	156:UNK	C	159:UNK	N	6.48