



wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 19, 2021 – 09:02 am BST

PDB ID : 7AGN
Title : Human carbonic anhydrase II in complex with 4-(2-aminoethylsulfanyl)-2,3,5,6-tetrafluoro-N-methyl-benzenesulfonamide
Authors : Paketuryte, V.; Smirnov, A.; Manakova, E.; Grazulis, S.
Deposited on : 2020-09-23
Resolution : 1.15 Å(reported)

This is a wwPDB X-ray Structure 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/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

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

MolProbity : 4.02b-467
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtrriage (Phenix) : **NOT EXECUTED**
EDS : **NOT EXECUTED**
buster-report : 1.1.7 (2018)
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.23.1

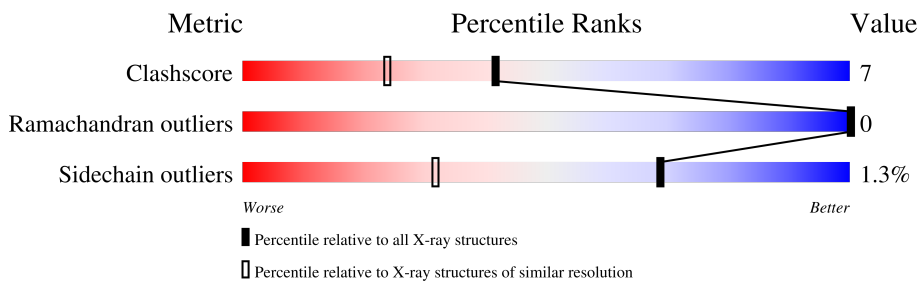
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.15 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)	Similar resolution (#Entries, resolution range(Å))
Clashscore	141614	1537 (1.18-1.10)
Ramachandran outliers	138981	1483 (1.18-1.10)
Sidechain outliers	138945	1480 (1.18-1.10)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower 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 $\leq 5\%$

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	260	
1	B	260	
1	C	260	
1	D	260	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	BCN	A	304	-	-	X	-
5	BCN	B	305	-	-	X	-
5	BCN	C	303	-	-	X	-
5	BCN	D	304	-	-	X	-

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 10132 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, 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.

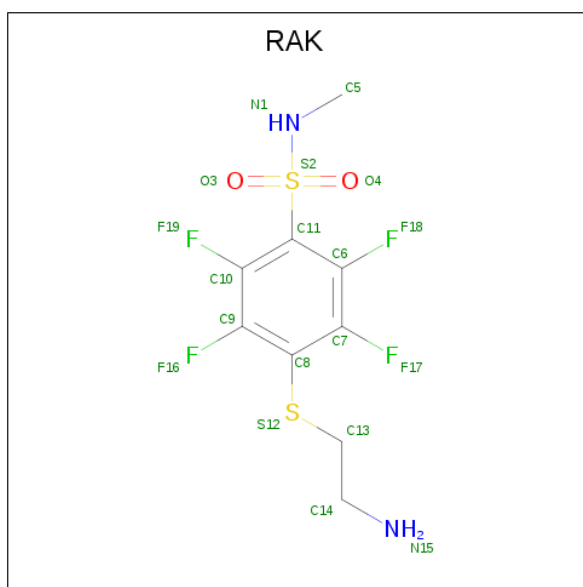
- Molecule 1 is a protein called carbonic anhydrase 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	258	Total 2165	C 1385	N 372	O 406	S 2	0	14	0
1	B	258	Total 2142	C 1373	N 365	O 402	S 2	0	12	0
1	C	258	Total 2165	C 1385	N 371	O 407	S 2	0	15	0
1	D	258	Total 2196	C 1403	N 376	O 415	S 2	0	19	0

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total 1	Zn 1	0	0
2	B	1	Total 1	Zn 1	0	0
2	C	1	Total 1	Zn 1	0	0
2	D	1	Total 1	Zn 1	0	0

- Molecule 3 is 4-(2-azanylethylsulfanyl)-2,3,5,6-tetrakis(fluoranyl)-{N}-methyl-benzenesulfonamide (three-letter code: RAK) (formula: C₉H₁₀F₄N₂O₂S₂) (labeled as "Ligand of Interest" by depositor).

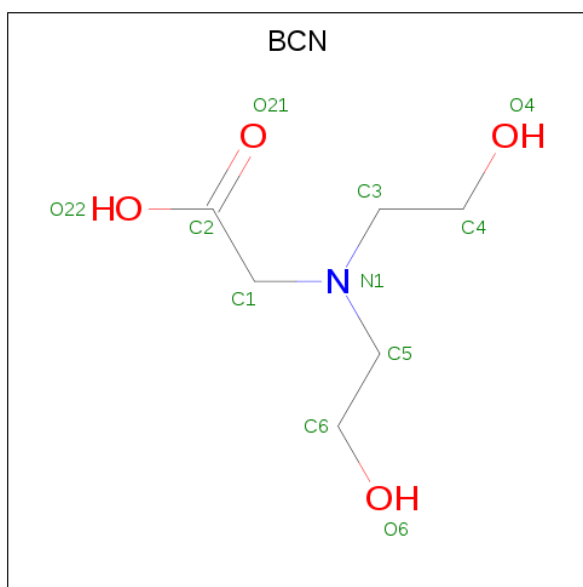


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	
			Total	C	F	N	O			S
3	A	1	Total	C	F	N	O	S	0	0
			19	9	4	2	2	2		
3	B	1	Total	C	F	N	O	S	0	0
			19	9	4	2	2	2		
3	C	1	Total	C	F	N	O	S	0	0
			19	9	4	2	2	2		
3	D	1	Total	C	F	N	O	S	0	0
			19	9	4	2	2	2		

- Molecule 4 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total	Na	0	0
			1	1		
4	B	2	Total	Na	0	0
			2	2		
4	D	1	Total	Na	0	0
			1	1		

- Molecule 5 is BICINE (three-letter code: BCN) (formula: C₆H₁₃NO₄).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total	C	N	O	0	0
			11	6	1	4		
5	B	1	Total	C	N	O	0	0
			11	6	1	4		
5	C	1	Total	C	N	O	0	0
			11	6	1	4		
5	D	1	Total	C	N	O	0	0
			11	6	1	4		

- Molecule 6 is water.


Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	344	Total	O	0	0
			344	344		
6	B	336	Total	O	0	0
			336	336		
6	C	331	Total	O	0	0
			331	331		
6	D	325	Total	O	0	0
			325	325		

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.

Note EDS was not executed.

- Molecule 1: carbonic anhydrase 2

Chain A:  92% 7% ..




- Molecule 1: carbonic anhydrase 2

Chain B:  94% 5% .




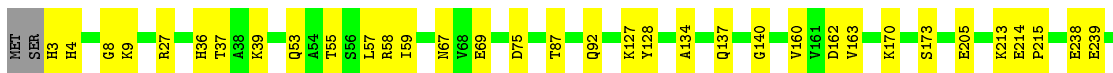
- Molecule 1: carbonic anhydrase 2

Chain C:  91% 8% ..



- Molecule 1: carbonic anhydrase 2

Chain D:  85% 14% .



4 Data and refinement statistics

Xtrriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	73.89Å 41.14Å 84.26Å 90.00° 109.51° 90.00°	Depositor
Resolution (Å)	69.65 – 1.15	Depositor
% Data completeness (in resolution range)	87.9 (69.65-1.15)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	0.10	Depositor
Refinement program	REFMAC 5.8.0189	Depositor
R, R_{free}	0.147 , 0.186	Depositor
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	10132	wwPDB-VP
Average B, all atoms (Å ²)	16.0	wwPDB-VP

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, BCN, NA, RAK

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	1.02	0/2230	1.07	0/3030
1	B	1.01	0/2207	1.06	0/2999
1	C	0.96	0/2231	1.07	0/3035
1	D	0.97	0/2261	1.02	0/3074
All	All	0.99	0/8929	1.06	0/12138

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	2165	0	2065	27	1
1	B	2142	0	2044	21	0
1	C	2165	0	2052	28	0
1	D	2196	0	2085	48	1
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
3	A	19	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	19	0	0	0	0
3	C	19	0	0	0	0
3	D	19	0	0	0	0
4	A	1	0	0	0	0
4	B	2	0	0	0	0
4	D	1	0	0	0	0
5	A	11	0	12	6	0
5	B	11	0	12	9	0
5	C	11	0	12	10	0
5	D	11	0	12	13	0
6	A	344	0	0	8	1
6	B	336	0	0	7	2
6	C	331	0	0	15	2
6	D	325	0	0	19	0
All	All	10132	0	8294	121	5

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 121 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:D:36[B]:HIS:CD2	6:D:408:HOH:O	1.99	1.15
1:C:10[B]:HIS:CD2	6:C:524:HOH:O	2.02	1.13
1:C:10[B]:HIS:CD2	6:C:425:HOH:O	2.10	1.03
1:D:8:GLY:HA3	6:D:406:HOH:O	1.60	1.02
1:A:58[A]:ARG:NH1	1:D:239:GLU:CG	2.28	0.97

All (5) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
6:C:480:HOH:O	6:C:634:HOH:O[1_545]	1.93	0.27
6:B:403:HOH:O	6:B:459:HOH:O[1_565]	2.05	0.15
1:A:73:SER:OG	6:A:419:HOH:O[1_565]	2.07	0.13
1:D:53:GLN:OE1	6:B:410:HOH:O[1_565]	2.07	0.13
6:C:404:HOH:O	6:C:473:HOH:O[1_565]	2.10	0.10

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 X-ray entries followed by that with respect to entries of similar resolution.

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	270/260 (104%)	257 (95%)	13 (5%)	0	100	100
1	B	268/260 (103%)	256 (96%)	12 (4%)	0	100	100
1	C	271/260 (104%)	264 (97%)	7 (3%)	0	100	100
1	D	275/260 (106%)	264 (96%)	11 (4%)	0	100	100
All	All	1084/1040 (104%)	1041 (96%)	43 (4%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

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	232/225 (103%)	228 (98%)	4 (2%)	60	22
1	B	229/225 (102%)	225 (98%)	4 (2%)	60	22
1	C	232/225 (103%)	227 (98%)	5 (2%)	52	12
1	D	235/225 (104%)	233 (99%)	2 (1%)	78	46
All	All	928/900 (103%)	913 (98%)	15 (2%)	69	24

5 of 15 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	261	LYS
1	D	4	HIS
1	C	55[A]	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	92	GLN
1	C	250[A]	PRO

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	4	HIS
1	C	255	GLN
1	D	255	GLN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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 monosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 16 ligands modelled in this entry, 8 are monoatomic - leaving 8 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	BCN	D	304	-	7,10,10	0.68	0	8,11,11	3.58	3 (37%)
5	BCN	C	303	-	7,10,10	0.83	1 (14%)	8,11,11	2.72	4 (50%)
3	RAK	C	302	2	18,19,19	1.14	1 (5%)	25,28,28	1.61	5 (20%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
3	RAK	D	302	2	18,19,19	1.52	2 (11%)	25,28,28	2.09	10 (40%)
5	BCN	A	304	-	7,10,10	0.70	0	8,11,11	3.09	2 (25%)
5	BCN	B	305	-	7,10,10	0.86	0	8,11,11	3.79	5 (62%)
3	RAK	A	302	2	18,19,19	2.02	2 (11%)	25,28,28	1.61	6 (24%)
3	RAK	B	302	2	18,19,19	1.84	4 (22%)	25,28,28	1.90	5 (20%)

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	BCN	D	304	-	-	5/8/10/10	-
5	BCN	C	303	-	-	5/8/10/10	-
3	RAK	C	302	2	-	2/13/13/13	0/1/1/1
3	RAK	D	302	2	-	2/13/13/13	0/1/1/1
5	BCN	A	304	-	-	4/8/10/10	-
5	BCN	B	305	-	-	4/8/10/10	-
3	RAK	A	302	2	-	0/13/13/13	0/1/1/1
3	RAK	B	302	2	-	1/13/13/13	0/1/1/1

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	302	RAK	C5-N1	-5.85	1.40	1.46
3	A	302	RAK	O3-S2	4.45	1.48	1.43
3	D	302	RAK	O4-S2	4.45	1.48	1.43
3	B	302	RAK	C5-N1	-4.36	1.41	1.46
3	B	302	RAK	O3-S2	4.16	1.48	1.43

The worst 5 of 40 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	D	304	BCN	C2-C1-N1	-8.59	101.25	113.48
5	A	304	BCN	C2-C1-N1	-7.16	103.28	113.48
5	B	305	BCN	C2-C1-N1	-6.82	103.76	113.48
3	B	302	RAK	C10-C11-C6	-5.27	112.98	116.67
3	D	302	RAK	C10-C11-C6	-4.75	113.34	116.67

There are no chirality outliers.

5 of 23 torsion outliers are listed below:

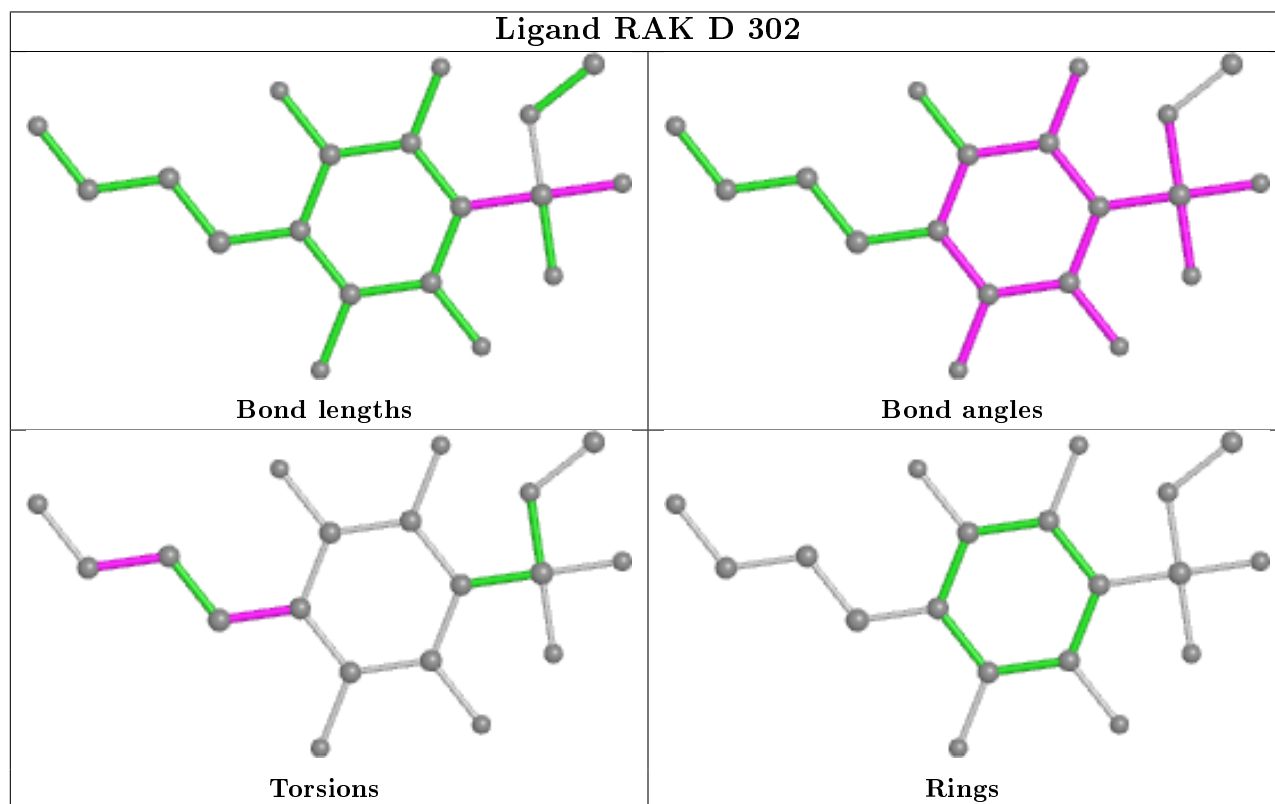
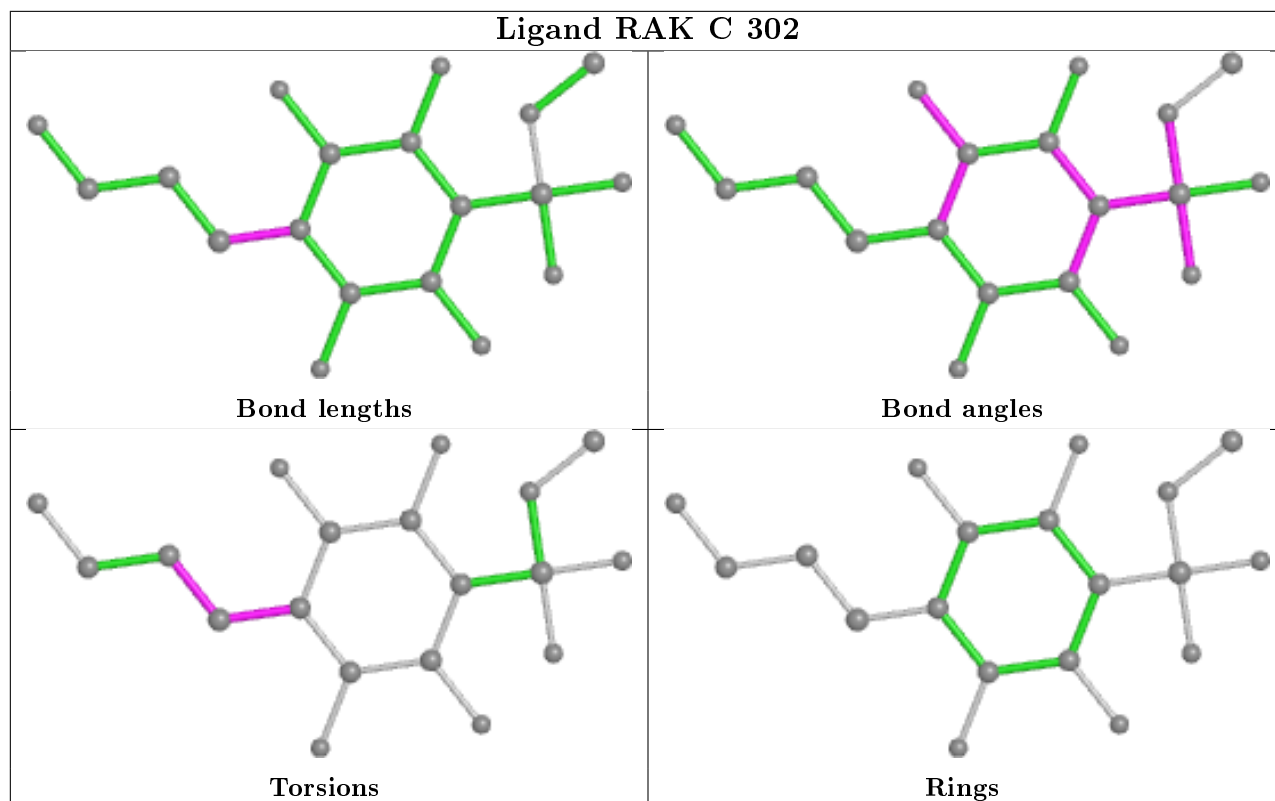
Mol	Chain	Res	Type	Atoms
3	B	302	RAK	S12-C13-C14-N15
3	C	302	RAK	C14-C13-S12-C8
3	D	302	RAK	S12-C13-C14-N15
5	B	305	BCN	C6-C5-N1-C3
5	C	303	BCN	C6-C5-N1-C1

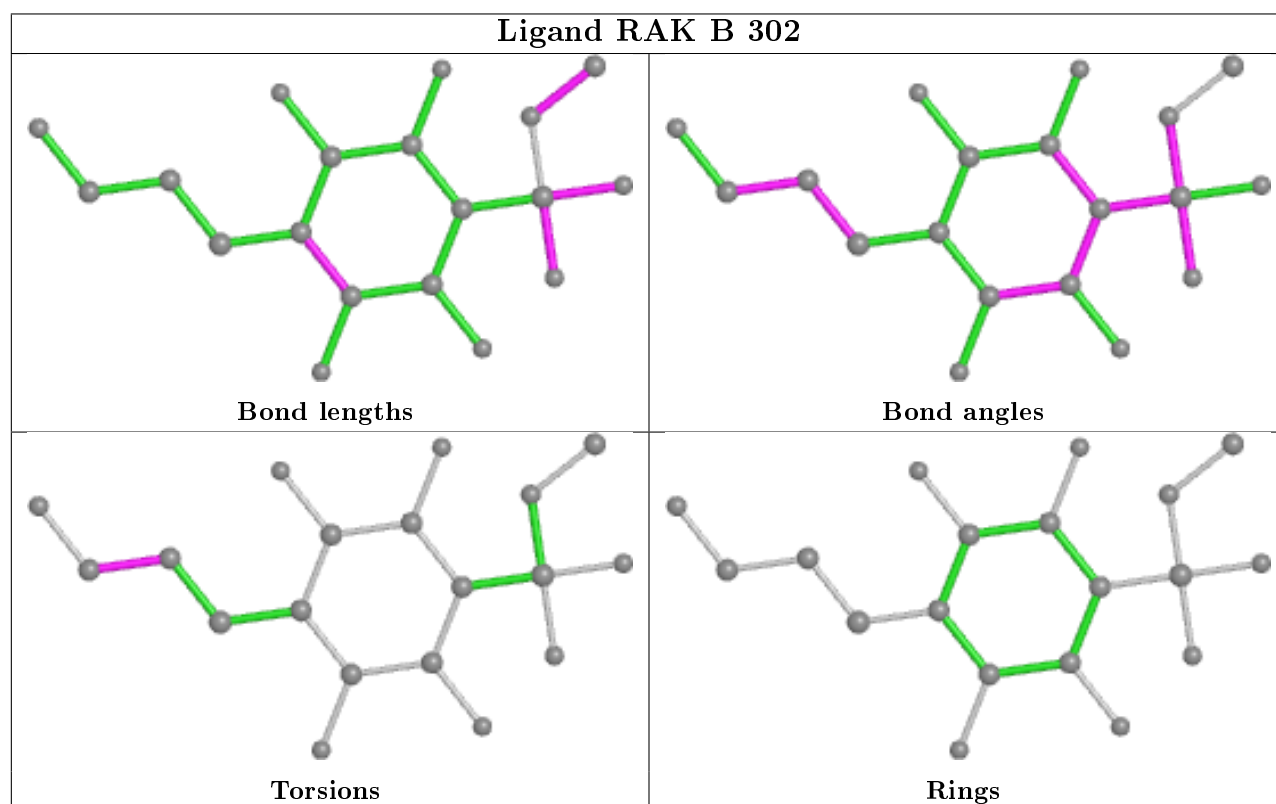
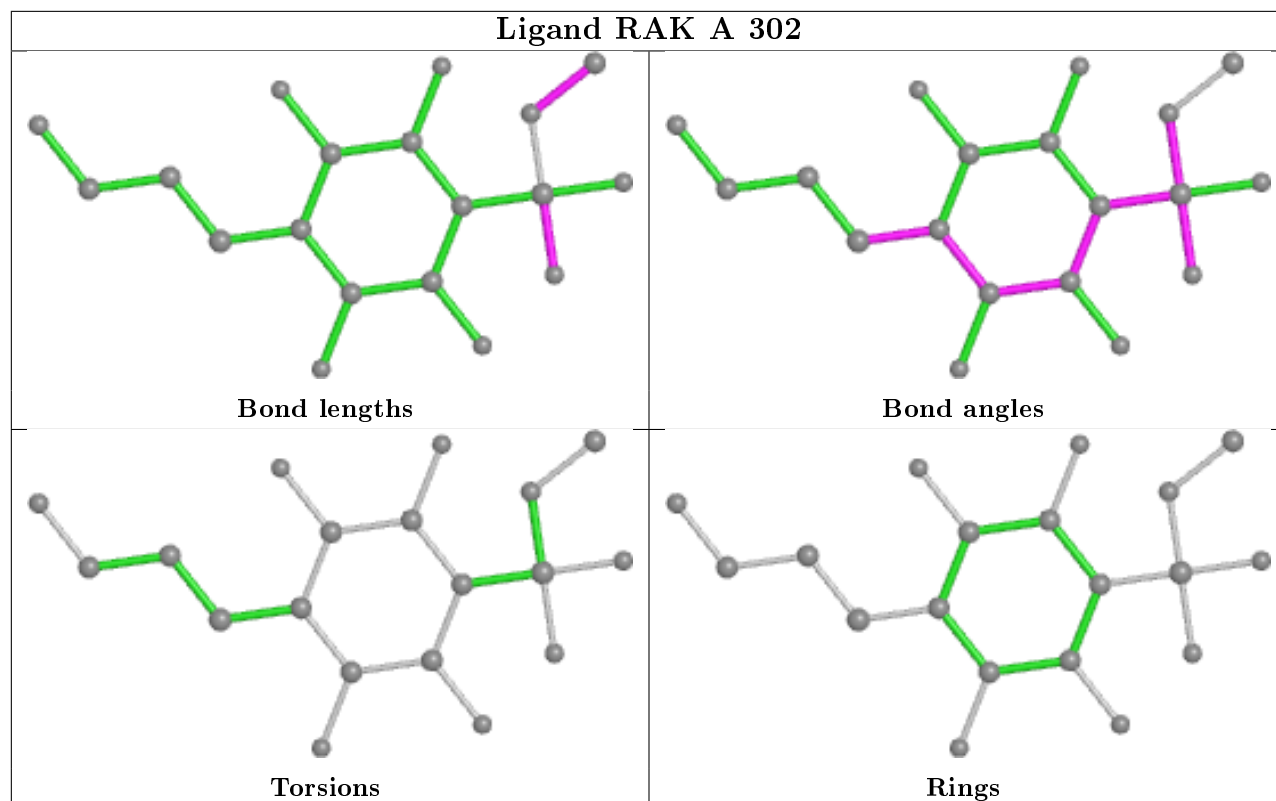
There are no ring outliers.

4 monomers are involved in 38 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	D	304	BCN	13	0
5	C	303	BCN	10	0
5	A	304	BCN	6	0
5	B	305	BCN	9	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 [\(i\)](#)

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

EDS was not executed - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

6.4 Ligands

EDS was not executed - this section is therefore empty.

6.5 Other polymers

EDS was not executed - this section is therefore empty.