Full wwPDB NMR Structure Validation Report (i)

Dec 17, 2023 – 12:45 PM EST

PDB ID	:	4AYK
Title	:	CATALYTIC FRAGMENT OF HUMAN FIBROBLAST COLLAGENASE COMPLEXED WITH CGS-27023A, NMR, 30 STRUCTURES
Authors	:	Powers, R.; Moy, F.J.
Deposited on	:	1999-02-01

This is a Full wwPDB NMR Structure Validation 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/NMRValidationReportHelp
with specific help available everywhere you see the (i) 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 (i)) were used in the production of this report:

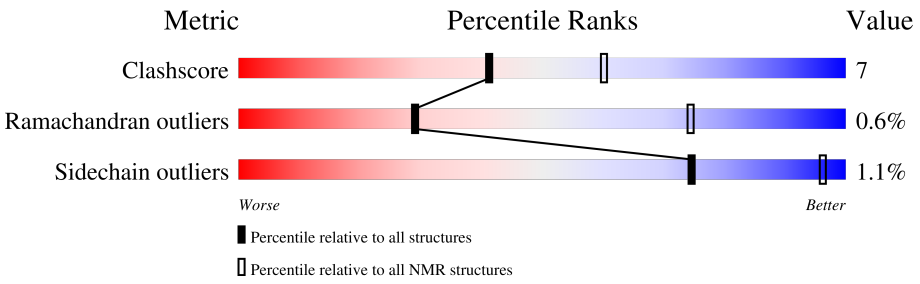
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
SOLUTION NMR

The overall completeness of chemical shifts assignment was not calculated.

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)	NMR archive (#Entries)
Clashscore	158937	12864
Ramachandran outliers	154571	11451
Sidechain outliers	154315	11428

The table below summarises the geometric issues observed across the polymeric chains and their fit to the experimental data. The red, orange, yellow and green segments indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. A cyan segment indicates the fraction of residues that are not part of the well-defined cores, and 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	169	82% 7% 11%

2 Ensemble composition and analysis (i)

This entry contains 30 models. Model 30 is the overall representative, medoid model (most similar to other models).

The following residues are included in the computation of the global validation metrics.

Well-defined (core) protein residues			
Well-defined core	**Residue range (total)**	**Backbone RMSD (Å)**	**Medoid model**
1	A:8-A:140, A:146-A:163 (151)	0.54	30

Ill-defined regions of proteins are excluded from the global statistics.

Ligands and non-protein polymers are included in the analysis.

The models can be grouped into 5 clusters and 3 single-model clusters were found.

Cluster number	Models
1	3, 4, 8, 9, 13, 14, 15, 19, 23, 30
2	5, 11, 16, 18, 21, 24, 26, 28
3	1, 2, 6, 12, 27
4	10, 25
5	17, 22
Single-model clusters	7; 20; 29

3 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 2618 atoms, of which 1253 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called PROTEIN (COLLAGENASE).

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	169	2565	834	1230	237	262	2	0

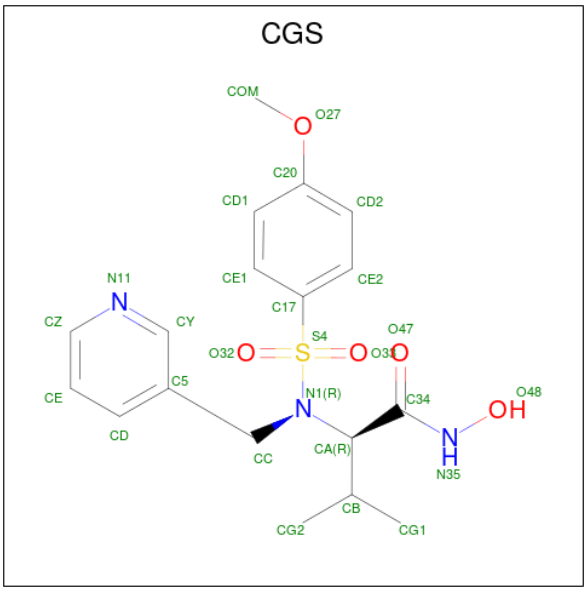
- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	
			Total	Zn
2	A	2	2	2

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	
			Total	Ca
3	A	1	1	1

- Molecule 4 is N-HYDROXY-2(R)-[[(4-METHOXYPHENYL)SULFONYL](3-PICOLYL)AMINO]-3-METHYLBUTANAMIDE HYDROCHLORIDE (three-letter code: CGS) (formula: $C_{18}H_{23}N_3O_5S$).



Mol	Chain	Residues	Atoms					
			Total	C	H	N	O	S
4	A	1	50	18	23	3	5	1

4 Residue-property plots (i)

4.1 Average score per residue in the NMR ensemble

These plots are provided for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic is the same as shown in the summary in section 1 of this report. The second graphic shows the sequence where residues are colour-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 outliers are shown as green connectors. Residues which are classified as ill-defined in the NMR ensemble, are shown in cyan with an underline colour-coded according to the previous scheme. Residues which were present in the experimental sample, but not modelled in the final structure are shown in grey.

• Molecule 1: PROTEIN (COLLAGENASE)

Chain A: 82% 7% 11%

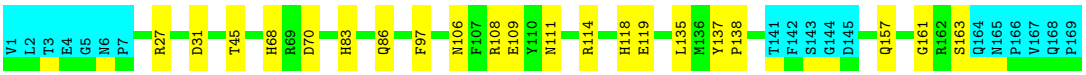
4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1

• Molecule 1: PROTEIN (COLLAGENASE)

Chain A: 77% 12% 11%



4.2.2 Score per residue for model 2

• Molecule 1: PROTEIN (COLLAGENASE)

Chain A: 79% 9% • 11%

4.2.3 Score per residue for model 3

- Molecule 1: PROTEIN (COLLAGENASE)

Chain A: 79% 10% • 11%

V1 L2 T3 E4 G5 N6 P7 E35 F38 Q39 L40 W41 T45 F49 H68 R69 D70 N80 N111 R114 H118 D131 I132 G133 T141 F142 S143 G144 D145 D151 Q157 G161 Q164 N165 P166 V167 Q168 P169

4.2.4 Score per residue for model 4

- Molecule 1: PROTEIN (COLLAGENASE)

Chain A: 71% 18% • 11%

V1 L2 T3 E4 G5 N6 P7 Q11 T12 H13 I18 E19 N20 F38 Q39 N43 T48 F49 T50 K51 D58 H68 R69 D70 N71 N80 F97 N106 F107 R108 V115 H118 E119 D131 I132 S139 Y140 T141 F142 S143 G144 D145 V146 Q147 Q157

G161 Q164 N165 P166 V167 Q168 P169

4.2.5 Score per residue for model 5

- Molecule 1: PROTEIN (COLLAGENASE)

Chain A: 79% 9% • 11%

V1 L2 T3 E4 G5 N6 P7 F38 V44 T45 F49 H68 R69 D70 H83 F97 N106 F107 R108 E109 Y110 N111 H118 E119 T141 F142 S143 G144 D145 Q157 G161 Q164 N165 P166 V167 Q168 P169

4.2.6 Score per residue for model 6

- Molecule 1: PROTEIN (COLLAGENASE)

Chain A: 73% 15% • 11%

V1 L2 T3 E4 G5 N6 P7 H13 H32 K36 T48 R65 D70 N71 N80 L81 A82 G90 I91 G92 F97 D98 E99 R102 N106 F107 R108 E109 Y110 N111 H118 E119 T141 F142 S143 G144 D145 V146 Q147 D151 Q157 G161 Q164

N165 P166 V167 Q168 P169

4.2.7 Score per residue for model 7

- Molecule 1: PROTEIN (COLLAGENASE)

Chain A: 82% 7% 11%

V1 L2 T3 E4 G5 N6 P7 D31 T45 H68 R69 D70 N80 H83 H118 M136 T141 F142 S143 G144 D145 I156 Q157 G161 R162 S163 Q164 N165 P166 V167 Q168 P169

4.2.8 Score per residue for model 8

- Molecule 1: PROTEIN (COLLAGENASE)

Chain A: 78% 11% • 11%

V1 L2 T3 E4 G5 N6 P7 F38 F49 H68 R69 D70 N80 F97 D98 E101 N106 F107 R108 N111 R114 H118 E119 Y137 P138 S139 Y140 T141 F142 S143 G144 D145 Q157 G161 Q164 N165 P166 V167 Q168 P169

4.2.9 Score per residue for model 9

- Molecule 1: PROTEIN (COLLAGENASE)

Chain A: 76% 12% • 11%

V1 L2 T3 E4 G5 N6 P7 E19 N20 P26 R27 F38 F49 H68 R69 D70 N80 F97 R102 W103 T104 N105 N106 F107 R108 E109 Y110 V115 H118 E119 Y140 T141 F142 S143 G144 D145 V146 Q157 G161 Q164 N165 P166 V167 Q168 P169

4.2.10 Score per residue for model 10

- Molecule 1: PROTEIN (COLLAGENASE)

Chain A: 81% 8% 11%

V1 L2 T3 E4 G5 N6 P7 H13 E19 N20 T48 H68 R69 D70 R102 W103 T104 Y110 H118 Y140 T141 F142 S143 G144 D145 V146 Q157 G161 Q164 N165 P166 V167 Q168 P169

4.2.11 Score per residue for model 11

- Molecule 1: PROTEIN (COLLAGENASE)

Chain A: 80% 9% 11%

V1 L2 T3 E4 G5 N6 P7 H13 T48 D67 H68 R69 D70 N71 N80 H83 P89 N106 F107 R108 R114 H118 T141 F142 S143 G144 D145 Q157 G161 R162 S163 Q164 N165 P166 V167 Q168 P169

4.2.12 Score per residue for model 12

- Molecule 1: PROTEIN (COLLAGENASE)

Chain A: 76% 11% • 11%

V1 L2 T3 E4 G5 N6 P7 P26 H32 K36 W41 T45 H68 R69 D70 N80 N106 F107 R108 E109 Y110 N111 R114 V115 H118 E119 Y137 T141 F142 S143 G144 D145 I156 Q157 G161 Q164 N165 P166 V167 Q168 P169

4.2.13 Score per residue for model 13

- Molecule 1: PROTEIN (COLLAGENASE)

Chain A: 75% 14% • 11%

V1 L2 T3 E4 G5 N6 P7 H13 L14 R17 L25 P26 H32 K36 T48 F49 V52 S53 E54 G55 R65 G66 D67 H68 R69 D70 N80 H83 N105 N106 N111 H118 T141 F142 S143 G144 D145 Q157 G161 Q164 N165 P166 V167

Q168 P169

4.2.14 Score per residue for model 14

- Molecule 1: PROTEIN (COLLAGENASE)

Chain A: 67% 21% • 11%

V1 L2 T3 E4 G5 N6 P7 H13 R17 I18 E19 N20 Y21 H32 K36 A37 F38 Q39 L40 W41 S42 N43 V44 T45 T48 F49 V52 S53 E54 G55 M60 D67 H68 R69 D70 N71 N80 H83 F97 N106 F107 N111 R114 H118 E119

I132 G133 P138 S139 Y140 T141 F142 S143 G144 D145 Q157 G161 Q164 N165 P166 V167 Q168 P169

4.2.15 Score per residue for model 15

- Molecule 1: PROTEIN (COLLAGENASE)

Chain A: 77% 12% • 11%

V1 L2 T3 E4 G5 N6 P7 H13 D31 E35 F38 Q39 T45 T48 F49 Q56 T104 N105 N106 F107 R108 E109 Y110 H118 D131 I132 G133 A134 L135 W136 Y137 T141 F142 S143 G144 D145 Q157 G161 R162 S163 Q164 N165 P166 V167 Q168 P169

4.2.16 Score per residue for model 16

• Molecule 1: PROTEIN (COLLAGENASE)

Chain A: 79% 9% • 11%

V1 L2 T3 E4 G5 N6 P7 H13 T48 R65 G66 D67 H68 R69 D70 N80 F97 N106 F107 H118 E119 L135 M136 Y137 P138 T141 F142 S143 G144 D145 Q157 G161 Q164 N165 P166 V167 Q168 P169

4.2.17 Score per residue for model 17

• Molecule 1: PROTEIN (COLLAGENASE)

Chain A: 75% 14% • 11%

V1 L2 T3 E4 G5 N6 P7 E10 H13 R27 A28 D31 F38 W41 T45 T48 F49 F74 N80 D98 N106 F107 R108 E109 H113 H118 S127 M136 Y137 P138 T141 F142 S143 G144 D145 V146 D151 Q157 G161 Q164 N165 P166 V167 Q168 P169

4.2.18 Score per residue for model 18

• Molecule 1: PROTEIN (COLLAGENASE)

Chain A: 76% 12% • 11%

V1 L2 T3 E4 G5 N6 P7 E19 N20 W41 T45 D67 H68 R69 D70 N80 H83 F97 D98 E99 D100 E101 R108 E109 R114 E119 T141 F142 S143 G144 D145 Q157 G161 R162 S163 Q164 N165 P166 V167 Q168 P169

4.2.19 Score per residue for model 19

• Molecule 1: PROTEIN (COLLAGENASE)

Chain A: 82% 8% 11%

V1 L2 T3 E4 G5 N6 P7 F38 F49 H68 R69 D70 N80 D98 E99 D100 R114 H118 T141 F142 S143 G144 D145 Q157 G161 Q164 N165 P166 V167 Q168 P169

4.2.20 Score per residue for model 20

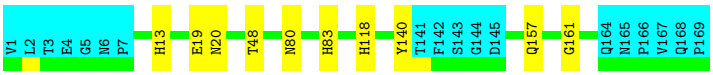
• Molecule 1: PROTEIN (COLLAGENASE)

Chain A: 75% 13% • 11%

4.2.21 Score per residue for model 21

• Molecule 1: PROTEIN (COLLAGENASE)

Chain A:



4.2.22 Score per residue for model 22

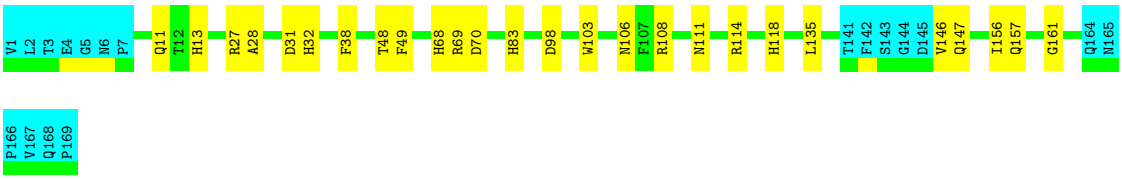
• Molecule 1: PROTEIN (COLLAGENASE)

Chain A:

4.2.23 Score per residue for model 23

• Molecule 1: PROTEIN (COLLAGENASE)

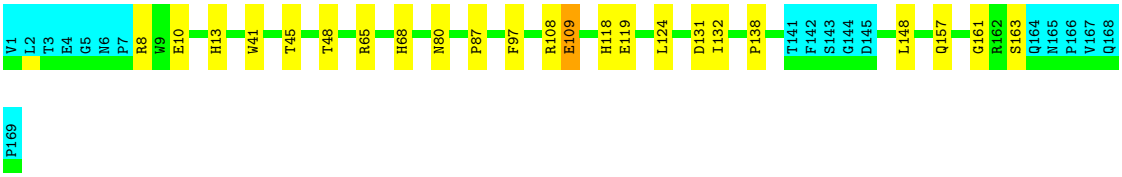
Chain A:



4.2.24 Score per residue for model 24

• Molecule 1: PROTEIN (COLLAGENASE)

Chain A:



4.2.25 Score per residue for model 25

- Molecule 1: PROTEIN (COLLAGENASE)

Chain A: 76% 12% • 11%

4.2.26 Score per residue for model 26

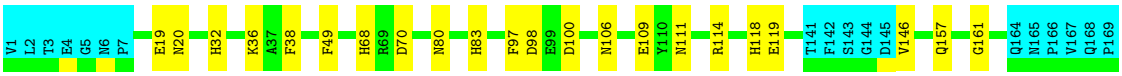
- Molecule 1: PROTEIN (COLLAGENASE)

Chain A: 76% 12% • 11%

4.2.27 Score per residue for model 27

- Molecule 1: PROTEIN (COLLAGENASE)

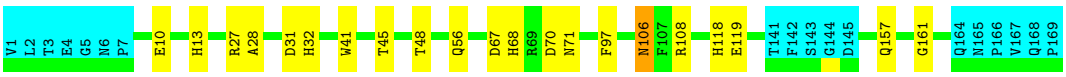
Chain A: 76% 13% 11%



4.2.28 Score per residue for model 28

- Molecule 1: PROTEIN (COLLAGENASE)

Chain A: 77% 12% • 11%



4.2.29 Score per residue for model 29

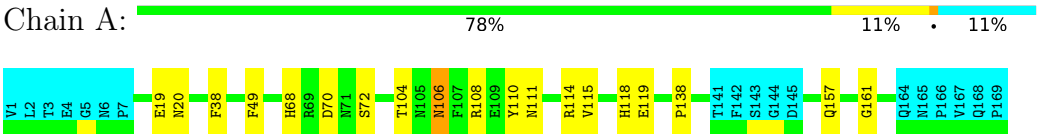
- Molecule 1: PROTEIN (COLLAGENASE)

Chain A: 77% 11% • 11%

4.2.30 Score per residue for model 30 (medoid)

- Molecule 1: PROTEIN (COLLAGENASE)

Chain A:



5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *HYBRID DISTANCE GEOMETRY-DYNAMICAL SIMULATED ANNEALING METHOD*.

Of the 100 calculated structures, 30 were deposited, based on the following criterion: *LOWEST ENERGY, LOWEST VIOLATIONS, CONSISTENT FOLD*.

The following table shows the software used for structure solution, optimisation and refinement.

Software name	Classification	Version
X-PLOR	refinement	3.84
X-PLOR	structure solution	

No chemical shift data was provided.

6 Model quality (i)

6.1 Standard geometry (i)

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

There are no covalent bond-length or bond-angle outliers.

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

There are no planarity outliers.

6.2 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 each 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 averaged over the ensemble.

Mol	Chain	Non-H	H(model)	H(added)	Clashes
1	A	1201	1106	1106	17±5
4	A	27	23	22	3±2
All	All	36930	33870	33840	523

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.

All unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:45:THR:HG22	1:A:157:GLN:NE2	0.84	1.87	15	1
1:A:80:ASN:ND2	4:A:173:CGS:HG12	0.71	2.00	3	8
1:A:32:HIS:CE1	1:A:36:LYS:NZ	0.62	2.68	26	3
1:A:106:ASN:HD21	1:A:108:ARG:HE	0.62	1.36	15	2
1:A:45:THR:HG22	1:A:157:GLN:HE21	0.62	1.52	15	1
1:A:80:ASN:CG	4:A:173:CGS:HG12	0.62	2.14	17	12
1:A:106:ASN:HD22	1:A:106:ASN:N	0.62	1.92	5	4
1:A:68:HIS:N	1:A:68:HIS:ND1	0.62	2.48	16	4
1:A:106:ASN:ND2	1:A:108:ARG:HE	0.62	1.93	15	2
1:A:80:ASN:OD1	4:A:173:CGS:HG12	0.61	1.94	18	4
1:A:80:ASN:HD22	1:A:80:ASN:N	0.60	1.94	8	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:114:ARG:NH1	4:A:173:CGS:HOM2	0.60	2.11	8	1
1:A:114:ARG:NH1	4:A:173:CGS:HOM1	0.60	2.11	29	2
1:A:118:HIS:ND1	1:A:136:MET:SD	0.59	2.76	7	1
1:A:38:PHE:CD1	1:A:49:PHE:CD2	0.59	2.91	30	4
1:A:20:ASN:HD22	1:A:21:TYR:N	0.58	1.96	14	1
1:A:68:HIS:ND1	1:A:68:HIS:N	0.58	2.49	1	13
1:A:38:PHE:CD1	1:A:49:PHE:CD1	0.58	2.91	23	8
1:A:106:ASN:HD22	1:A:106:ASN:H	0.58	1.41	5	9
1:A:157:GLN:O	1:A:161:GLY:N	0.58	2.36	12	30
1:A:118:HIS:ND1	4:A:173:CGS:HOM2	0.58	2.14	20	2
1:A:13:HIS:HD1	1:A:48:THR:HG1	0.57	1.38	24	1
1:A:106:ASN:ND2	1:A:106:ASN:N	0.56	2.52	4	12
1:A:136:MET:SD	1:A:136:MET:N	0.56	2.78	29	1
1:A:80:ASN:ND2	4:A:173:CGS:HA	0.56	2.16	6	2
1:A:20:ASN:ND2	1:A:21:TYR:N	0.56	2.54	14	1
1:A:107:PHE:CD2	1:A:107:PHE:O	0.56	2.58	16	4
1:A:106:ASN:N	1:A:106:ASN:HD22	0.56	1.98	20	7
1:A:97:PHE:CE1	1:A:119:GLU:OE1	0.56	2.59	6	12
1:A:106:ASN:O	1:A:111:ASN:ND2	0.55	2.39	25	4
1:A:13:HIS:ND1	1:A:48:THR:OG1	0.55	2.40	13	10
1:A:32:HIS:CE1	1:A:36:LYS:HZ1	0.55	2.20	13	1
1:A:80:ASN:ND2	4:A:173:CGS:HG23	0.55	2.17	6	2
1:A:70:ASP:OD1	1:A:71:ASN:N	0.55	2.40	14	7
1:A:68:HIS:CD2	1:A:70:ASP:OD2	0.55	2.60	18	11
1:A:13:HIS:CD2	1:A:48:THR:OG1	0.54	2.61	28	8
1:A:106:ASN:H	1:A:106:ASN:HD22	0.54	1.45	15	2
1:A:17:ARG:NH2	1:A:55:GLY:O	0.54	2.41	14	2
1:A:68:HIS:CE1	1:A:70:ASP:OD2	0.54	2.60	29	6
1:A:32:HIS:CE1	1:A:36:LYS:HZ3	0.54	2.19	26	2
1:A:107:PHE:O	1:A:107:PHE:CG	0.54	2.59	16	2
1:A:106:ASN:ND2	1:A:108:ARG:CZ	0.54	2.70	11	1
1:A:97:PHE:CZ	1:A:119:GLU:OE1	0.54	2.61	18	14
1:A:106:ASN:N	1:A:106:ASN:ND2	0.54	2.56	5	3
1:A:10:GLU:N	1:A:10:GLU:OE1	0.54	2.40	28	3
1:A:31:ASP:OD1	1:A:32:HIS:N	0.54	2.41	23	2
1:A:8:ARG:NH1	1:A:124:LEU:O	0.53	2.41	24	1
1:A:56:GLN:H	1:A:56:GLN:CD	0.53	2.06	15	2
1:A:11:GLN:NE2	1:A:58:ASP:OD2	0.53	2.41	4	1
1:A:20:ASN:HD22	1:A:21:TYR:H	0.53	1.47	14	1
1:A:68:HIS:NE2	1:A:70:ASP:OD2	0.53	2.42	19	3
1:A:108:ARG:O	1:A:109:GLU:C	0.52	2.47	17	11

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:82:ALA:O	4:A:173:CGS:HG23	0.52	2.04	20	1
1:A:106:ASN:HD21	1:A:108:ARG:NE	0.52	2.02	15	2
1:A:101:GLU:N	1:A:101:GLU:OE1	0.52	2.43	18	3
1:A:118:HIS:CD2	4:A:173:CGS:O48	0.52	2.62	16	25
1:A:38:PHE:CD1	1:A:49:PHE:CG	0.52	2.98	27	6
1:A:97:PHE:CE2	1:A:119:GLU:OE2	0.52	2.62	22	3
1:A:45:THR:N	1:A:163:SER:OG	0.52	2.43	18	3
1:A:68:HIS:N	1:A:68:HIS:HD1	0.51	2.03	18	4
1:A:97:PHE:CE2	1:A:119:GLU:OE1	0.51	2.63	4	1
1:A:106:ASN:HD21	1:A:108:ARG:CZ	0.51	2.18	11	1
1:A:106:ASN:ND2	1:A:108:ARG:CG	0.51	2.74	4	5
1:A:80:ASN:ND2	1:A:82:ALA:O	0.51	2.41	26	3
1:A:17:ARG:NE	1:A:52:VAL:O	0.51	2.40	13	2
1:A:38:PHE:CE1	1:A:49:PHE:CD1	0.50	2.99	27	5
1:A:32:HIS:NE2	1:A:36:LYS:CE	0.50	2.73	14	2
1:A:162:ARG:NH1	1:A:162:ARG:CG	0.50	2.74	18	1
1:A:122:HIS:CE1	1:A:128:HIS:CE1	0.50	3.00	29	1
1:A:32:HIS:NE2	1:A:36:LYS:NZ	0.50	2.60	13	1
1:A:127:SER:O	1:A:136:MET:SD	0.50	2.69	17	1
1:A:17:ARG:NH2	1:A:53:SER:O	0.49	2.45	25	1
1:A:118:HIS:ND1	1:A:135:LEU:O	0.49	2.45	16	1
1:A:106:ASN:O	1:A:111:ASN:OD1	0.49	2.31	12	6
4:A:173:CGS:HG13	4:A:173:CGS:CY	0.49	2.38	4	4
1:A:17:ARG:NH2	1:A:54:GLU:C	0.49	2.65	13	2
1:A:146:VAL:O	1:A:147:GLN:NE2	0.49	2.45	23	3
1:A:90:GLY:O	1:A:92:GLY:N	0.49	2.46	6	1
1:A:119:GLU:N	1:A:119:GLU:OE1	0.49	2.45	30	1
1:A:67:ASP:OD1	1:A:68:HIS:N	0.49	2.46	11	4
1:A:28:ALA:O	1:A:31:ASP:OD1	0.48	2.30	23	3
1:A:109:GLU:H	1:A:109:GLU:CD	0.48	2.12	27	2
1:A:45:THR:CA	1:A:163:SER:OG	0.48	2.62	18	3
1:A:39:GLN:NE2	1:A:39:GLN:N	0.48	2.62	3	1
1:A:102:ARG:NH2	1:A:108:ARG:NH2	0.48	2.62	6	1
1:A:25:LEU:HD13	1:A:105:ASN:CG	0.48	2.27	13	1
1:A:137:TYR:O	1:A:139:SER:N	0.47	2.47	22	4
1:A:115:VAL:O	1:A:119:GLU:OE1	0.47	2.31	30	3
1:A:106:ASN:OD1	1:A:108:ARG:CG	0.47	2.63	8	2
1:A:19:GLU:OE2	1:A:60:MET:SD	0.47	2.73	14	1
1:A:157:GLN:OE1	1:A:163:SER:N	0.47	2.47	15	1
1:A:68:HIS:CD2	1:A:70:ASP:CG	0.47	2.88	3	1
1:A:80:ASN:N	1:A:80:ASN:ND2	0.47	2.62	8	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:70:ASP:OD1	1:A:72:SER:N	0.47	2.48	22	2
1:A:98:ASP:O	1:A:103:TRP:NE1	0.47	2.48	23	3
1:A:80:ASN:CB	4:A:173:CGS:HG12	0.47	2.40	24	1
1:A:17:ARG:NH2	1:A:54:GLU:O	0.47	2.48	14	1
1:A:162:ARG:CG	1:A:162:ARG:HH11	0.47	2.23	18	1
1:A:80:ASN:ND2	1:A:82:ALA:H	0.46	2.08	20	1
1:A:80:ASN:HD22	4:A:173:CGS:HG12	0.46	1.68	26	1
1:A:86:GLN:CD	1:A:86:GLN:N	0.46	2.69	1	1
1:A:104:THR:OG1	1:A:110:TYR:O	0.46	2.32	10	7
1:A:106:ASN:HD21	1:A:108:ARG:CG	0.46	2.23	20	6
1:A:118:HIS:HD1	4:A:173:CGS:HOM2	0.46	1.69	20	1
1:A:131:ASP:OD1	1:A:132:ILE:N	0.46	2.49	24	4
4:A:173:CGS:HG13	4:A:173:CGS:CD	0.46	2.41	24	1
1:A:44:VAL:HG23	1:A:45:THR:HG23	0.46	1.87	5	2
1:A:69:ARG:O	1:A:70:ASP:OD1	0.46	2.34	19	7
1:A:19:GLU:O	1:A:20:ASN:OD1	0.46	2.34	21	7
1:A:68:HIS:CE1	1:A:70:ASP:CG	0.46	2.89	27	2
1:A:80:ASN:ND2	4:A:173:CGS:O33	0.46	2.48	26	3
1:A:115:VAL:HG22	4:A:173:CGS:HOM3	0.46	1.88	9	1
1:A:98:ASP:OD1	1:A:100:ASP:N	0.46	2.42	19	2
1:A:97:PHE:CZ	1:A:119:GLU:CD	0.45	2.89	6	3
1:A:90:GLY:C	1:A:92:GLY:N	0.45	2.70	6	1
1:A:114:ARG:NH2	1:A:140:TYR:CE1	0.45	2.84	14	1
1:A:18:ILE:H	1:A:51:LYS:NZ	0.45	2.09	4	1
1:A:113:HIS:ND1	1:A:113:HIS:C	0.45	2.69	17	1
1:A:118:HIS:CE1	1:A:135:LEU:O	0.45	2.70	16	1
1:A:114:ARG:HH12	4:A:173:CGS:HOM1	0.45	1.71	29	2
1:A:114:ARG:NH1	1:A:135:LEU:CD2	0.45	2.80	23	1
1:A:11:GLN:CD	1:A:11:GLN:N	0.45	2.70	23	1
1:A:65:ARG:NH1	1:A:65:ARG:CB	0.44	2.80	24	1
1:A:106:ASN:ND2	1:A:108:ARG:NE	0.44	2.65	11	1
1:A:27:ARG:O	1:A:31:ASP:OD1	0.44	2.36	28	3
1:A:17:ARG:HH22	1:A:54:GLU:C	0.44	2.16	14	2
1:A:74:PHE:CD2	1:A:98:ASP:OD2	0.44	2.71	17	1
1:A:80:ASN:O	1:A:80:ASN:OD1	0.44	2.36	24	1
1:A:131:ASP:OD1	1:A:133:GLY:N	0.44	2.45	15	2
1:A:106:ASN:ND2	1:A:106:ASN:H	0.44	2.11	4	1
1:A:38:PHE:CE1	1:A:49:PHE:CD2	0.44	3.06	17	3
1:A:109:GLU:O	1:A:140:TYR:OH	0.43	2.36	9	1
1:A:80:ASN:ND2	1:A:80:ASN:O	0.43	2.51	8	1
4:A:173:CGS:O32	4:A:173:CGS:HD	0.43	2.14	19	2

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:17:ARG:CZ	1:A:53:SER:O	0.43	2.66	25	1
1:A:41:TRP:O	1:A:45:THR:OG1	0.43	2.36	28	6
1:A:35:GLU:O	1:A:39:GLN:OE1	0.43	2.35	3	2
1:A:83:HIS:O	1:A:83:HIS:CG	0.43	2.70	7	1
1:A:39:GLN:NE2	1:A:39:GLN:CA	0.43	2.81	3	1
1:A:151:ASP:OD1	1:A:151:ASP:C	0.42	2.58	3	1
1:A:156:ILE:HG23	1:A:157:GLN:N	0.42	2.30	12	3
1:A:151:ASP:O	1:A:151:ASP:OD1	0.42	2.38	17	1
1:A:135:LEU:O	1:A:137:TYR:N	0.42	2.47	15	4
1:A:106:ASN:C	1:A:108:ARG:H	0.42	2.17	5	1
1:A:118:HIS:CD2	4:A:173:CGS:HOM2	0.42	2.49	7	1
1:A:114:ARG:HH12	4:A:173:CGS:HOM2	0.42	1.74	8	1
1:A:102:ARG:O	1:A:110:TYR:CD2	0.42	2.73	10	2
1:A:69:ARG:NH1	1:A:69:ARG:CG	0.41	2.82	26	2
1:A:106:ASN:C	1:A:106:ASN:HD22	0.41	2.17	22	1
1:A:90:GLY:C	1:A:92:GLY:H	0.41	2.19	6	1
1:A:83:HIS:CD2	1:A:83:HIS:N	0.41	2.88	18	1
1:A:80:ASN:HD22	1:A:81:LEU:N	0.41	2.14	20	1
1:A:98:ASP:OD2	1:A:101:GLU:OE2	0.41	2.39	8	1
1:A:111:ASN:HD22	1:A:112:LEU:N	0.41	2.14	2	1
1:A:65:ARG:NH1	1:A:99:GLU:O	0.41	2.52	6	1
1:A:31:ASP:C	1:A:31:ASP:OD1	0.41	2.59	7	1
4:A:173:CGS:HG13	4:A:173:CGS:C5	0.41	2.46	26	2
1:A:114:ARG:NH1	1:A:114:ARG:CG	0.41	2.83	27	1
1:A:39:GLN:O	1:A:43:ASN:CG	0.41	2.59	4	2
1:A:80:ASN:HD21	4:A:173:CGS:HA	0.41	1.76	6	1
1:A:151:ASP:C	1:A:151:ASP:OD1	0.41	2.59	6	1
1:A:65:ARG:O	1:A:66:GLY:C	0.41	2.59	16	2
1:A:108:ARG:CB	1:A:108:ARG:NH1	0.41	2.84	18	1
1:A:151:ASP:OD1	1:A:151:ASP:O	0.41	2.38	3	1
1:A:139:SER:OG	1:A:140:TYR:N	0.41	2.54	4	1
1:A:137:TYR:C	1:A:139:SER:H	0.41	2.18	8	1
1:A:39:GLN:O	1:A:43:ASN:OD1	0.41	2.39	14	1
1:A:31:ASP:OD1	1:A:31:ASP:C	0.41	2.59	15	1
1:A:99:GLU:O	1:A:99:GLU:CG	0.41	2.69	19	1
1:A:52:VAL:O	1:A:52:VAL:HG23	0.41	2.16	25	1
1:A:80:ASN:OD1	1:A:80:ASN:O	0.41	2.39	11	1
1:A:86:GLN:H	1:A:86:GLN:CD	0.41	2.19	20	1
1:A:138:PRO:O	4:A:173:CGS:HC2	0.41	2.16	24	1
1:A:14:LEU:HD12	1:A:49:PHE:CE2	0.41	2.50	13	1
1:A:77:PRO:O	1:A:101:GLU:OE2	0.41	2.38	20	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:32:HIS:O	1:A:36:LYS:N	0.41	2.43	27	1
1:A:98:ASP:OD1	1:A:98:ASP:C	0.40	2.59	6	2
1:A:132:ILE:HG23	1:A:133:GLY:N	0.40	2.30	14	1
1:A:41:TRP:CH2	1:A:148:LEU:HD22	0.40	2.51	24	1
1:A:27:ARG:NH2	1:A:31:ASP:OD2	0.40	2.54	1	1
1:A:75:ASP:O	1:A:75:ASP:OD1	0.40	2.39	26	1
1:A:25:LEU:HD13	1:A:105:ASN:ND2	0.40	2.31	13	1
1:A:80:ASN:HB3	4:A:173:CGS:HG12	0.40	1.94	14	1
1:A:109:GLU:CD	1:A:109:GLU:N	0.40	2.75	27	1
1:A:118:HIS:ND1	4:A:173:CGS:O27	0.40	2.55	14	1
1:A:98:ASP:OD1	1:A:99:GLU:N	0.40	2.54	18	1

6.3 Torsion angles (i)

6.3.1 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 NMR 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	151/169 (89%)	139±1 (92±1%)	11±2 (7±1%)	1±1 (1±1%)	29	74
All	All	4530/5070 (89%)	4183 (92%)	322 (7%)	25 (1%)	29	74

All 8 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	138	PRO	8
1	A	109	GLU	7
1	A	26	PRO	3
1	A	89	PRO	2
1	A	107	PHE	2
1	A	91	ILE	1
1	A	136	MET	1
1	A	87	PRO	1

6.3.2 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 PDB entries followed by that with respect to all NMR 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	125/141 (89%)	124±1 (99±1%)	1±1 (1±1%)	74 96
All	All	3750/4230 (89%)	3708 (99%)	42 (1%)	74 96

All 13 unique residues with a non-rotameric sidechain are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	106	ASN	15
1	A	83	HIS	13
1	A	111	ASN	3
1	A	80	ASN	2
1	A	39	GLN	1
1	A	26	PRO	1
1	A	27	ARG	1
1	A	119	GLU	1
1	A	20	ASN	1
1	A	162	ARG	1
1	A	114	ARG	1
1	A	101	GLU	1
1	A	136	MET	1

6.3.3 RNA (i)

There are no RNA molecules in this entry.

6.4 Non-standard residues in protein, DNA, RNA chains (i)

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

6.5 Carbohydrates (i)

There are no monosaccharides in this entry.

6.6 Ligand geometry (i)

Of 4 ligands modelled in this entry, 3 are monoatomic - leaving 1 for Mogul analysis.

In the following table, the Counts columns list the number of bonds for which Mogul statistics could be retrieved, the number of bonds that are observed in the model and the number of bonds 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 is the number of standard deviations the observed value is removed from the expected value. A bond length with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond lengths.

Mol	Type	Chain	Res	Link	Bond lengths		
					Counts	RMSZ	#Z>2
4	CGS	A	173	2	28,28,28	3.00±0.03	13±1 (45±2%)

In the following table, the Counts columns list the number of angles for which Mogul statistics could be retrieved, the number of angles that are observed in the model and the number of 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 angle is the number of standard deviations the observed value is removed from the expected value. A bond angle with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond angles.

Mol	Type	Chain	Res	Link	Bond angles		
					Counts	RMSZ	#Z>2
4	CGS	A	173	2	37,39,39	1.46±0.04	3±1 (8±1%)

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
4	CGS	A	173	2	-	0±0,32,32,32	0±0,2,2,2

All unique bond outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)	Models	
								Worst	Total
4	A	173	CGS	S4-N1	10.00	1.78	1.63	19	30
4	A	173	CGS	CE2-C17	4.57	1.46	1.38	4	30
4	A	173	CGS	CD2-C20	4.45	1.47	1.38	13	30
4	A	173	CGS	CD1-C20	4.39	1.47	1.38	10	30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)	Models	
								Worst	Total
4	A	173	CGS	CE1-C17	4.22	1.45	1.38	29	30
4	A	173	CGS	CD2-CE2	4.19	1.46	1.38	14	30
4	A	173	CGS	CE1-CD1	3.89	1.45	1.38	12	30
4	A	173	CGS	CD-C5	3.39	1.46	1.38	16	30
4	A	173	CGS	CE-CD	3.11	1.45	1.38	24	30
4	A	173	CGS	CY-N11	2.98	1.40	1.34	17	30
4	A	173	CGS	CY-C5	2.61	1.44	1.38	17	30
4	A	173	CGS	CC-N1	2.59	1.52	1.48	25	24
4	A	173	CGS	C17-S4	2.21	1.79	1.76	29	4
4	A	173	CGS	O27-C20	2.15	1.42	1.37	7	5
4	A	173	CGS	CE-CZ	2.06	1.43	1.37	10	17

All unique angle outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
4	A	173	CGS	CB-CA-N1	5.43	122.27	112.99	25	30
4	A	173	CGS	C5-CC-N1	4.06	120.45	113.48	24	30
4	A	173	CGS	O33-S4-N1	2.54	102.33	106.97	4	23
4	A	173	CGS	O32-S4-C17	2.23	110.86	108.05	6	2
4	A	173	CGS	O32-S4-N1	2.07	103.19	106.97	17	2
4	A	173	CGS	C17-S4-N1	2.03	110.98	107.36	12	1
4	A	173	CGS	CA-N1-S4	2.00	114.21	118.30	13	1

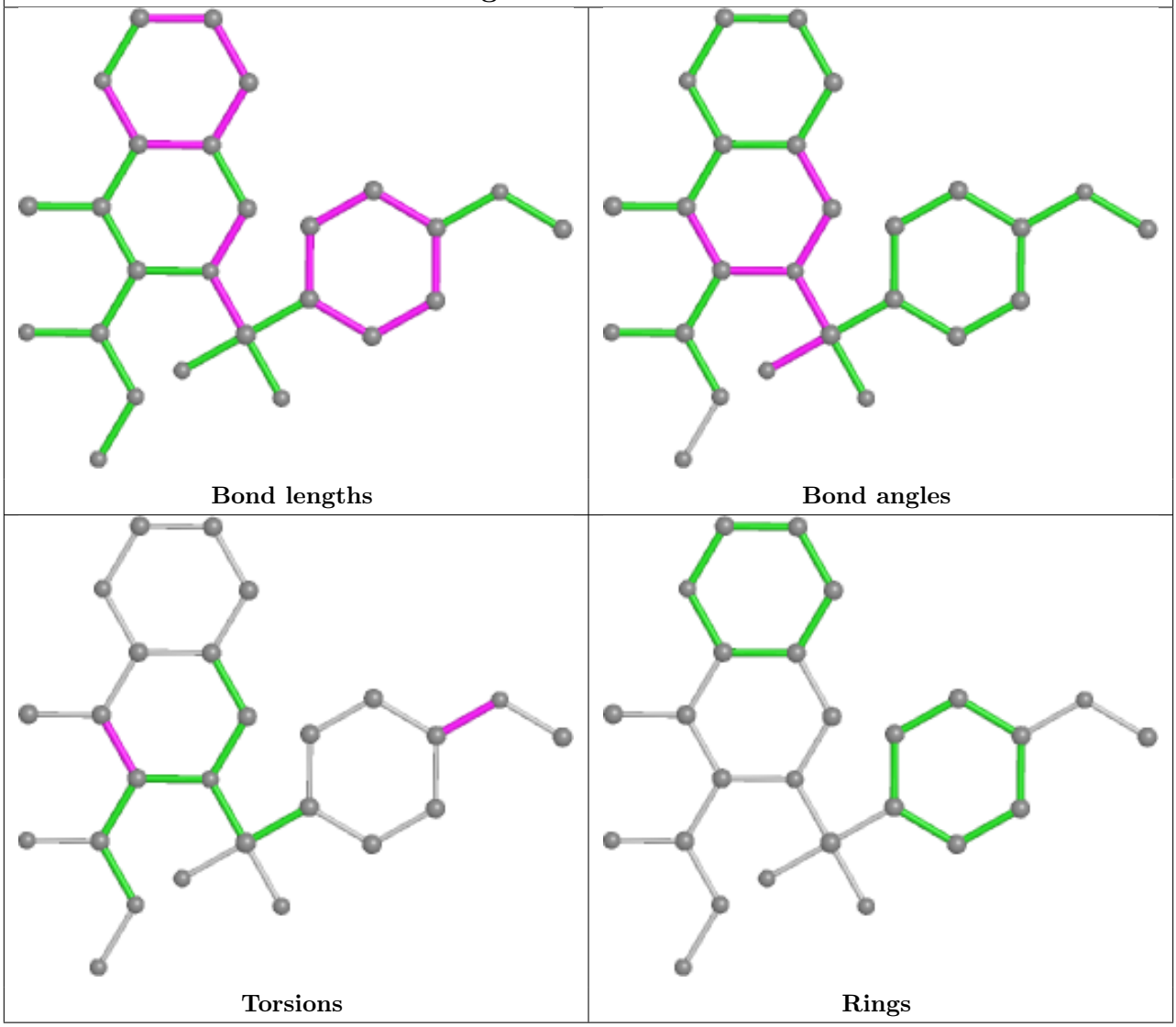
There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

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 CGS A 173



6.7 Other polymers (i)

There are no such molecules in this entry.

6.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

7 Chemical shift validation (i)

No chemical shift data were provided