



wwPDB X-ray Structure Validation Summary Report ⓘ

May 29, 2020 – 01:43 pm BST

PDB ID : 2C1U
Title : CRYSTAL STRUCTURE OF THE DI-HAEM CYTOCHROME C PEROXIDASE FROM PARACOCCLUS PANTOTROPHUS - OXIDISED FORM
Authors : Echalier, A.; Fulop, V.
Deposited on : 2005-09-21
Resolution : 1.95 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)
Xtriage (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) : 2.11

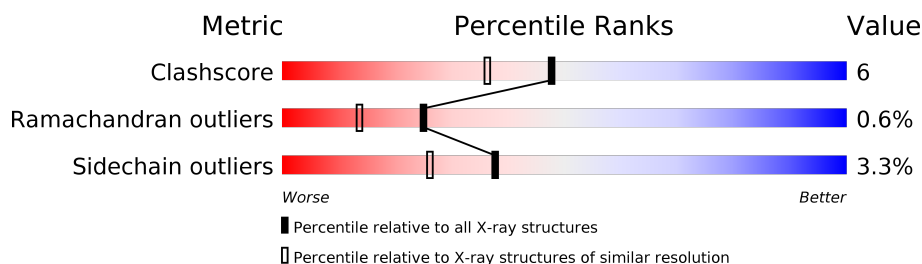
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.95 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	2705 (1.96-1.96)
Ramachandran outliers	138981	2678 (1.96-1.96)
Sidechain outliers	138945	2678 (1.96-1.96)

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 on 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	338	
1	B	338	
1	C	338	
1	D	338	

2 Entry composition [i](#)

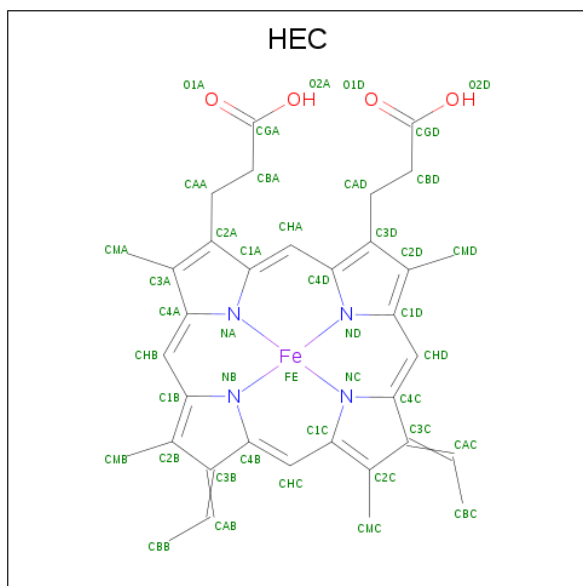
There are 4 unique types of molecules in this entry. The entry contains 11016 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 DI-HAEM CYTOCHROME C PEROXIDASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	329	Total	C	N	O	S	0	0	0
			2479	1566	415	486	12			
1	B	324	Total	C	N	O	S	0	0	0
			2446	1545	410	479	12			
1	C	329	Total	C	N	O	S	0	0	0
			2479	1567	415	485	12			
1	D	322	Total	C	N	O	S	0	0	0
			2431	1537	408	474	12			

- Molecule 2 is HEME C (three-letter code: HEC) (formula: $C_{34}H_{34}FeN_4O_4$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
2	A	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	B	1	Total	C	Fe	N	O	
			43	34	1	4	4	
2	B	1	Total	C	Fe	N	O	
			43	34	1	4	4	
2	C	1	Total	C	Fe	N	O	
			43	34	1	4	4	
2	C	1	Total	C	Fe	N	O	
			43	34	1	4	4	
2	D	1	Total	C	Fe	N	O	
			43	34	1	4	4	
2	D	1	Total	C	Fe	N	O	
			43	34	1	4	4	

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Ca		
			1	1	0	0
3	A	1	Total	Ca		
			1	1	0	0
3	D	1	Total	Ca		
			1	1	0	0
3	C	1	Total	Ca		
			1	1	0	0

- Molecule 4 is water.

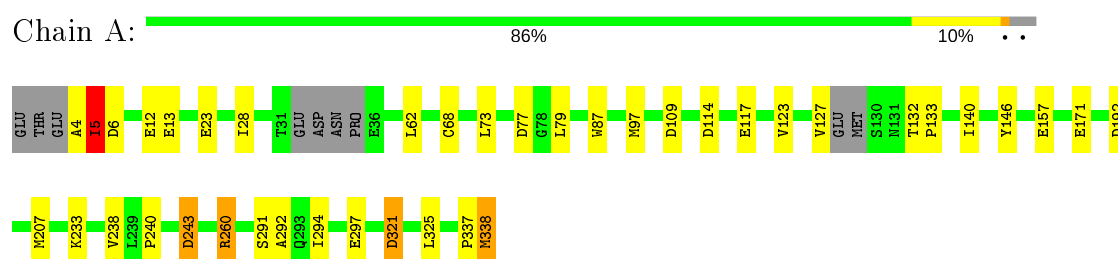
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	208	Total	O		
			208	208	0	0
4	B	199	Total	O		
			199	199	0	0
4	C	221	Total	O		
			221	221	0	0
4	D	205	Total	O		
			205	205	0	0

3 Residue-property plots

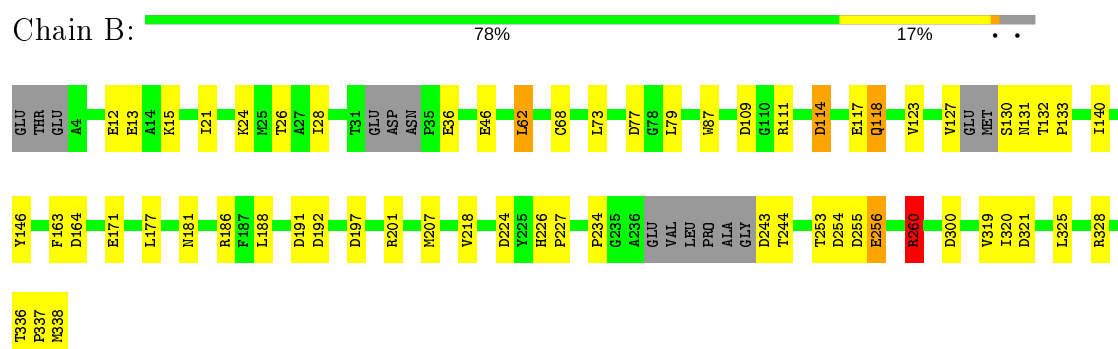
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according 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.

Note EDS was not executed.

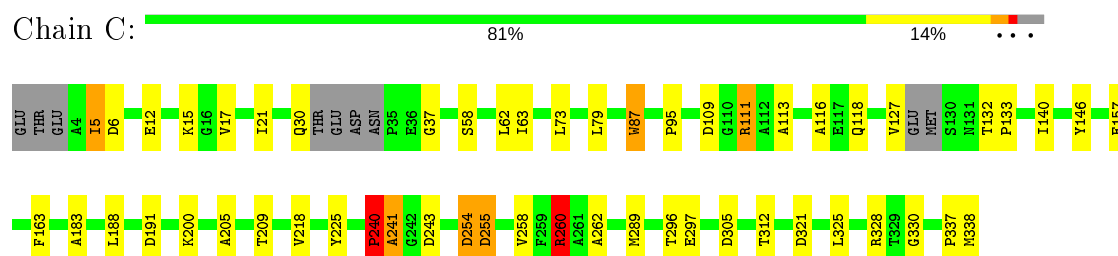
• Molecule 1: DI-HAEM CYTOCHROME C PEROXIDASE



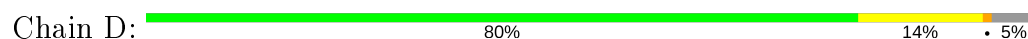
• Molecule 1: DI-HAEM CYTOCHROME C PEROXIDASE



• Molecule 1: DI-HAEM CYTOCHROME C PEROXIDASE



• Molecule 1: DI-HAEM CYTOCHROME C PEROXIDASE



GLU	THR	GLU	A4	I5	V17	F18	I21	P22	E23	K24	K29	Q30	T31	GLU	ASP	ASN	P35	E36	L62	L73	L79	S82	I83	W87	P91	R92	D109	D114	E117	V127	GLU	MET	S130	N131	T132	P133	E171	L177	P180	A183	D191
D192	V218	Y225	G235	A236	GLU	VAL	LEU	PRO	ALA	GLY	ASP	THR	G245	R246	T250	D254	D255	V258	F259	R260	E281	S291	I294	T312	E316	V319	I320	D321	L325	R328	T336	P337	M338								

4 Data and refinement statistics

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

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	78.80 Å 51.10 Å 167.90 Å 90.00° 98.20° 90.00°	Depositor
Resolution (Å)	29.88 – 1.95	Depositor
% Data completeness (in resolution range)	88.6 (29.88-1.95)	Depositor
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	REFMAC 5.1.9999	Depositor
R, R_{free}	0.169 , 0.220	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	11016	wwPDB-VP
Average B, all atoms (Å ²)	34.0	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: CA, HEC

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	0.74	0/2535	0.89	8/3449 (0.2%)
1	B	0.72	0/2501	0.89	10/3400 (0.3%)
1	C	0.80	2/2536 (0.1%)	0.89	10/3450 (0.3%)
1	D	0.71	0/2486	0.87	10/3379 (0.3%)
All	All	0.74	2/10058 (0.0%)	0.88	38/13678 (0.3%)

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	B	0	1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	12	GLU	CD-OE2	11.62	1.38	1.25
1	C	12	GLU	CD-OE1	5.39	1.31	1.25

The worst 5 of 38 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	260	ARG	NE-CZ-NH2	-10.38	115.11	120.30
1	B	328	ARG	NE-CZ-NH1	9.75	125.18	120.30
1	B	328	ARG	NE-CZ-NH2	-8.69	115.95	120.30
1	C	328	ARG	NE-CZ-NH1	7.49	124.05	120.30
1	A	321	ASP	CB-CG-OD2	7.02	124.61	118.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	234	PRO	Peptide

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	2479	0	2392	24	0
1	B	2446	0	2358	33	0
1	C	2479	0	2393	38	0
1	D	2431	0	2347	31	0
2	A	86	0	60	5	0
2	B	86	0	60	6	0
2	C	86	0	60	7	0
2	D	86	0	60	2	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
4	A	208	0	0	2	0
4	B	199	0	0	4	0
4	C	221	0	0	3	0
4	D	205	0	0	3	0
All	All	11016	0	9730	122	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 122 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:C:296:THR:O	1:C:297:GLU:HG2	1.48	1.12
1:D:83:ILE:CD1	1:D:91:PRO:HA	1.96	0.96
1:D:83:ILE:HD13	1:D:91:PRO:HA	1.59	0.83
1:D:83:ILE:HD11	1:D:91:PRO:CA	2.13	0.79
1:C:79:LEU:HD11	1:D:62:LEU:HG	1.69	0.74

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 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	323/338 (96%)	309 (96%)	12 (4%)	2 (1%)	25	14
1	B	316/338 (94%)	299 (95%)	15 (5%)	2 (1%)	25	14
1	C	323/338 (96%)	312 (97%)	8 (2%)	3 (1%)	17	8
1	D	314/338 (93%)	300 (96%)	13 (4%)	1 (0%)	41	30
All	All	1276/1352 (94%)	1220 (96%)	48 (4%)	8 (1%)	25	14

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	5	ILE
1	C	5	ILE
1	C	240	PRO
1	C	241	ALA
1	B	244	THR

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	258/267 (97%)	251 (97%)	7 (3%)	44	34
1	B	255/267 (96%)	242 (95%)	13 (5%)	24	11
1	C	258/267 (97%)	251 (97%)	7 (3%)	44	34

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	D	253/267 (95%)	246 (97%)	7 (3%)	43	33
All	All	1024/1068 (96%)	990 (97%)	34 (3%)	38	26

5 of 34 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	131	ASN
1	B	319	VAL
1	D	250	THR
1	B	256	GLU
1	A	338	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	30	GLN
1	B	106	GLN
1	B	131	ASN
1	C	118	GLN
1	D	181	ASN

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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 12 ligands modelled in this entry, 4 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
2	HEC	B	401	1	26,50,50	2.29	8 (30%)	18,82,82	2.67	7 (38%)
2	HEC	C	401	1	26,50,50	2.65	8 (30%)	18,82,82	1.93	3 (16%)
2	HEC	A	401	1	26,50,50	2.42	8 (30%)	18,82,82	2.35	5 (27%)
2	HEC	D	402	1	26,50,50	2.12	4 (15%)	18,82,82	1.54	3 (16%)
2	HEC	B	402	1	26,50,50	2.24	7 (26%)	18,82,82	1.99	6 (33%)
2	HEC	A	402	1	26,50,50	2.50	5 (19%)	18,82,82	1.61	4 (22%)
2	HEC	C	402	1	26,50,50	2.37	9 (34%)	18,82,82	2.14	6 (33%)
2	HEC	D	401	1	26,50,50	2.45	6 (23%)	18,82,82	2.49	7 (38%)

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
2	HEC	B	401	1	-	0/6/54/54	-
2	HEC	C	401	1	-	0/6/54/54	-
2	HEC	A	401	1	-	0/6/54/54	-
2	HEC	D	402	1	-	0/6/54/54	-
2	HEC	B	402	1	-	0/6/54/54	-
2	HEC	A	402	1	-	0/6/54/54	-
2	HEC	C	402	1	-	0/6/54/54	-
2	HEC	D	401	1	-	0/6/54/54	-

The worst 5 of 55 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	401	HEC	C3B-C2B	-7.95	1.32	1.40
2	A	402	HEC	C3B-C2B	-7.22	1.33	1.40
2	A	401	HEC	C3B-C2B	-7.15	1.33	1.40
2	C	401	HEC	C3C-C2C	-6.51	1.34	1.40
2	C	402	HEC	C3B-C2B	-6.43	1.34	1.40

The worst 5 of 41 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
2	D	401	HEC	CMC-C2C-C1C	-6.84	117.96	128.46
2	A	401	HEC	CMB-C2B-C1B	-5.51	120.00	128.46
2	B	401	HEC	CMC-C2C-C1C	-5.43	120.12	128.46
2	B	401	HEC	CBD-CAD-C3D	-4.89	103.46	112.49
2	B	402	HEC	CBD-CAD-C3D	-4.47	104.25	112.49

There are no chirality outliers.

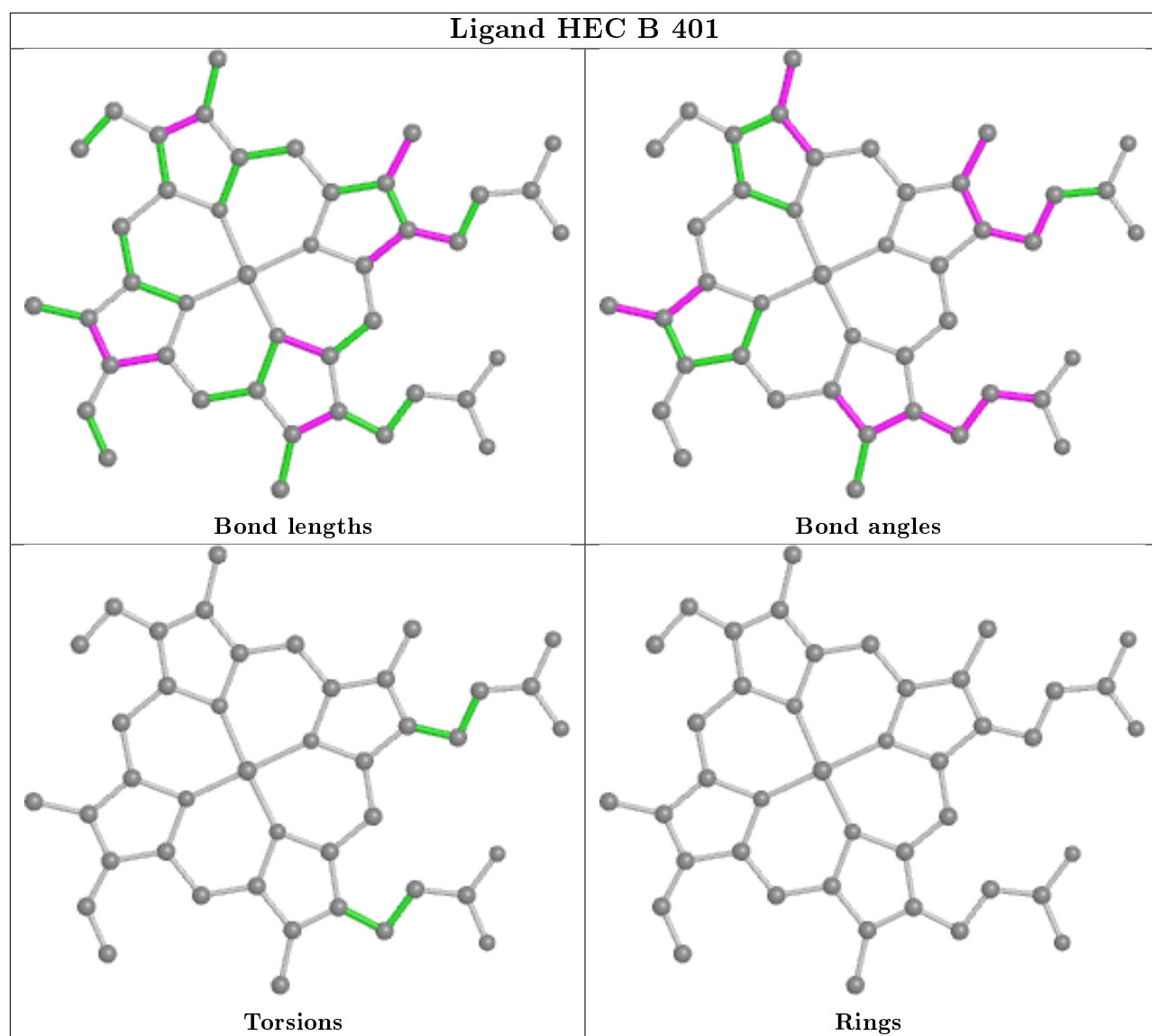
There are no torsion outliers.

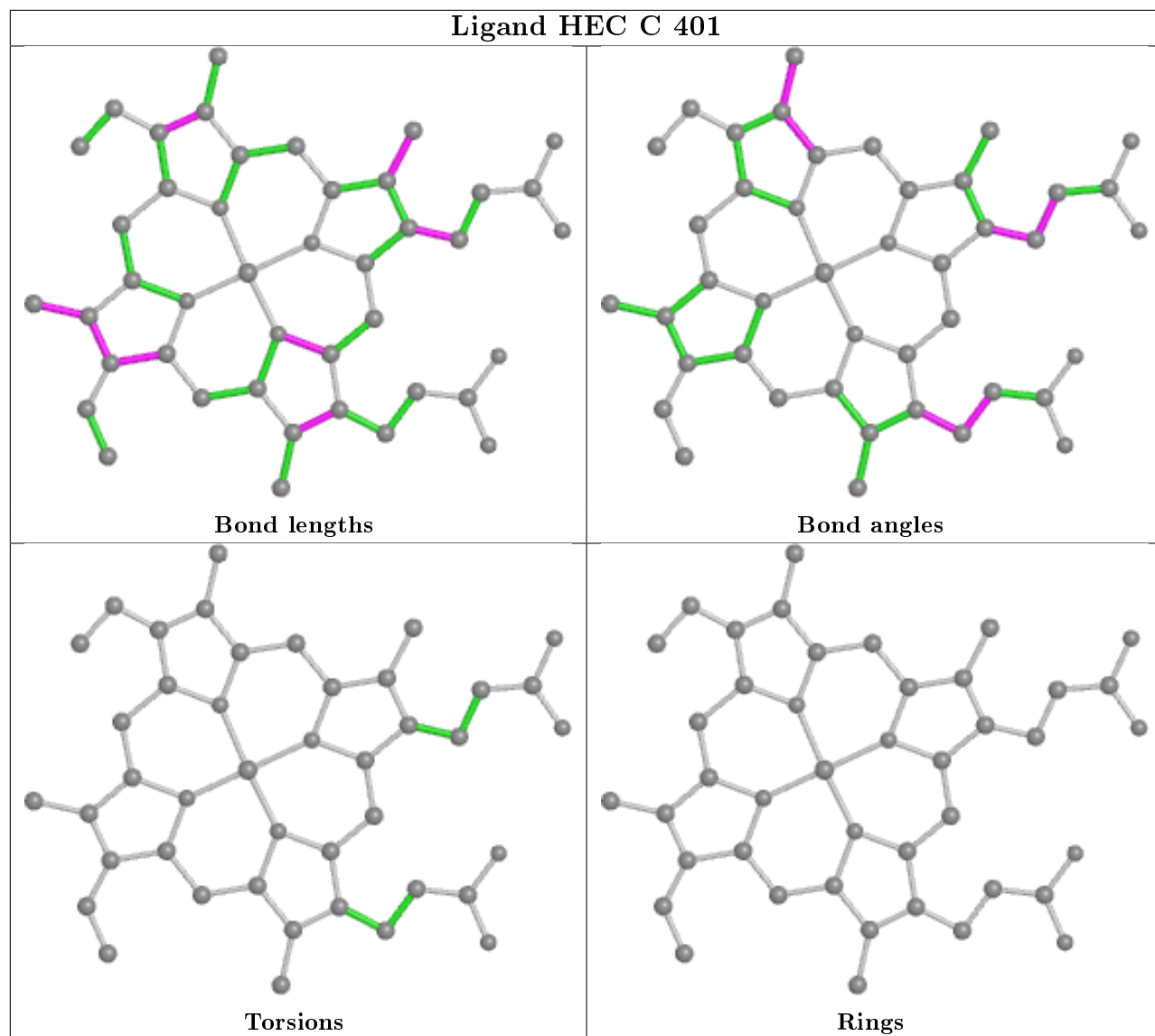
There are no ring outliers.

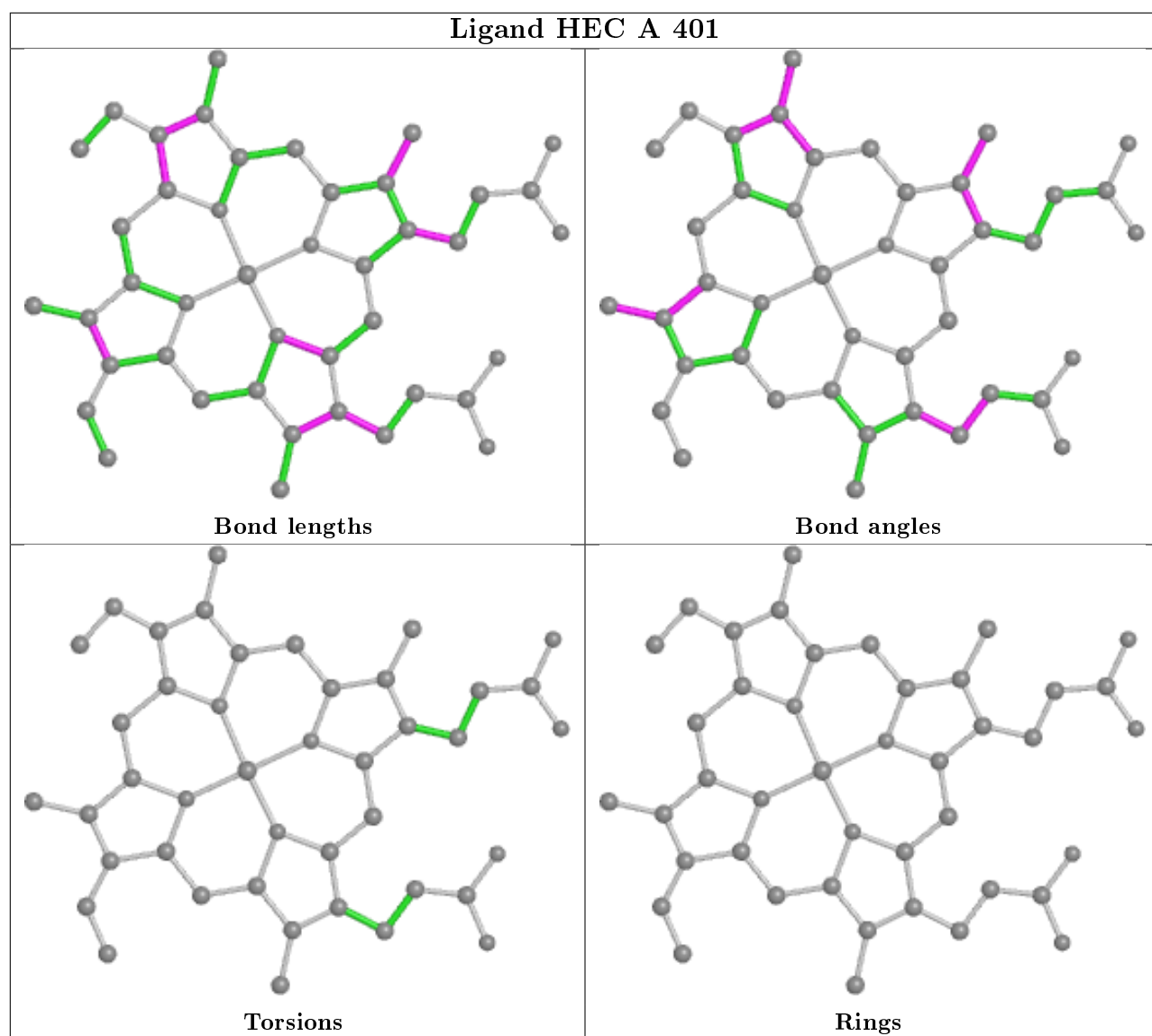
7 monomers are involved in 20 short contacts:

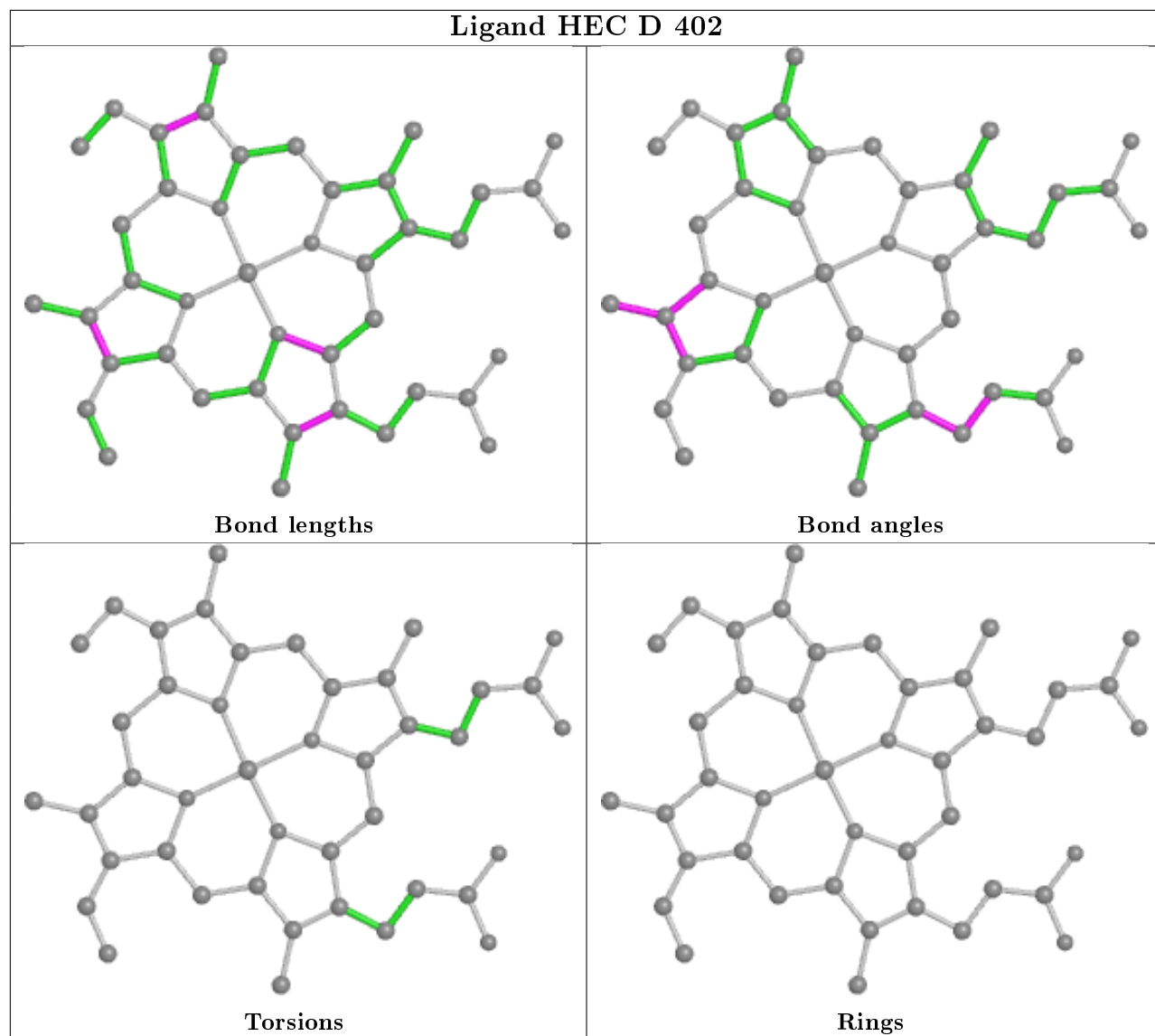
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	401	HEC	5	0
2	C	401	HEC	3	0
2	A	401	HEC	4	0
2	B	402	HEC	1	0
2	A	402	HEC	1	0
2	C	402	HEC	4	0
2	D	401	HEC	2	0

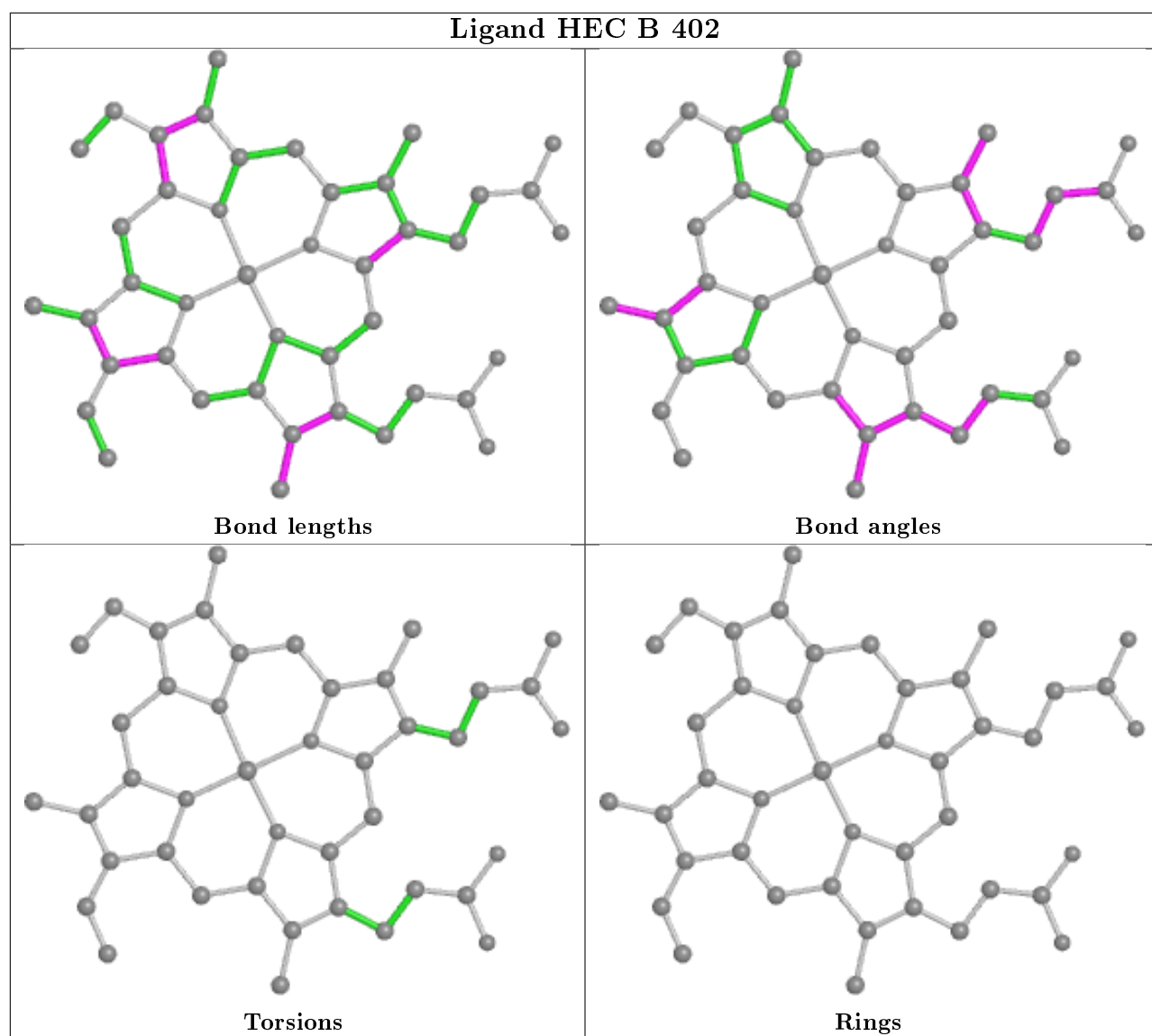
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. 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.

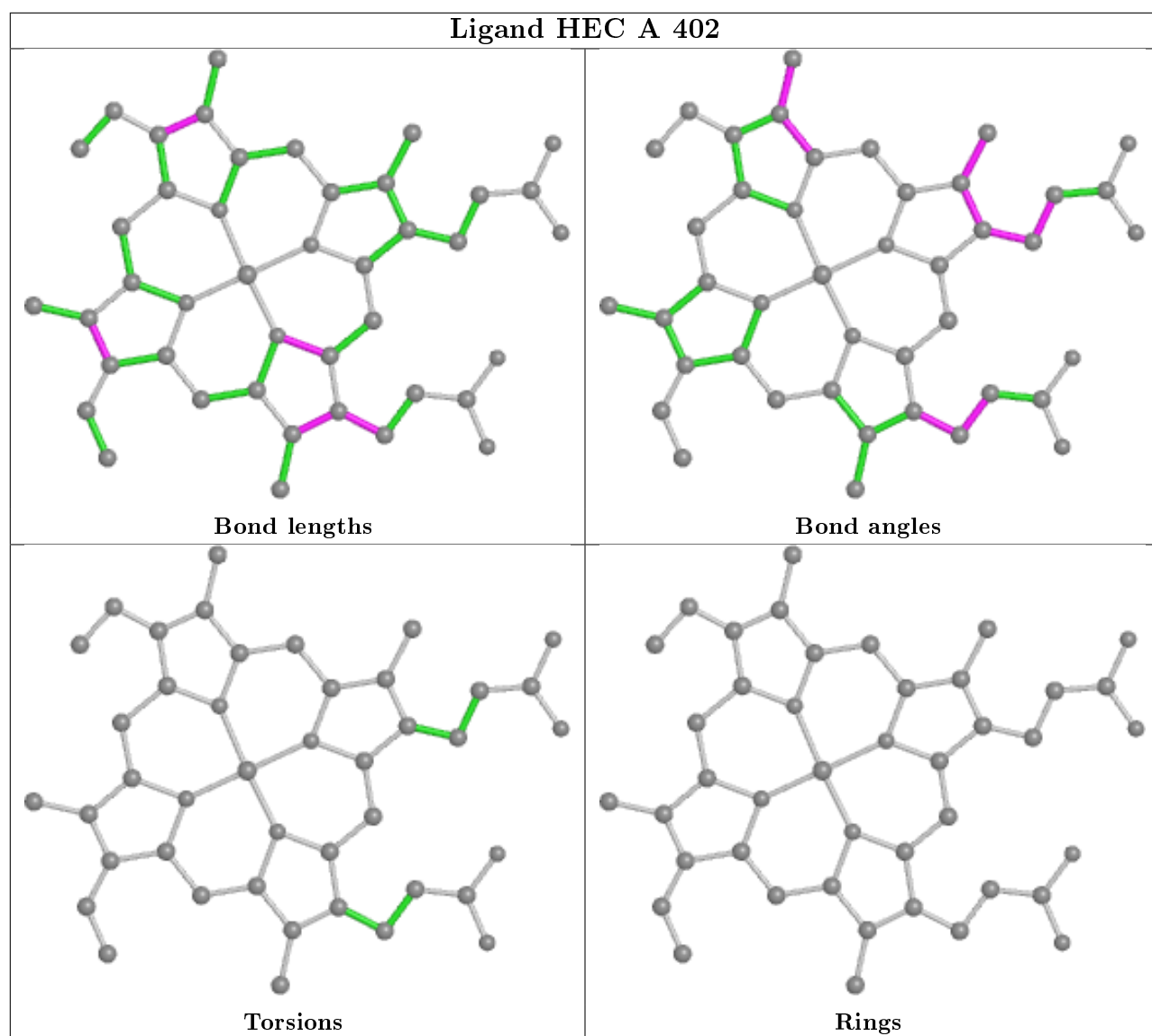


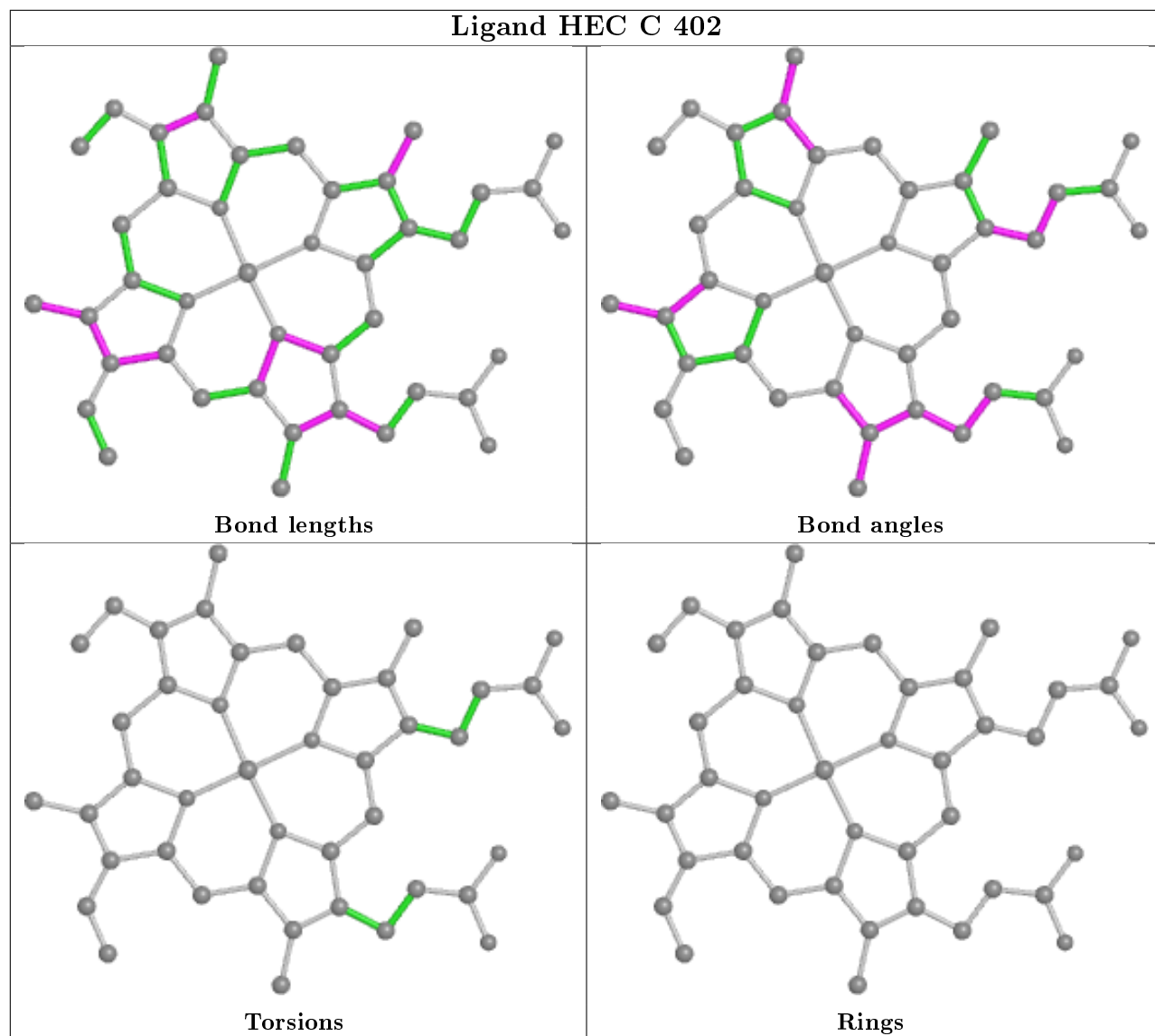


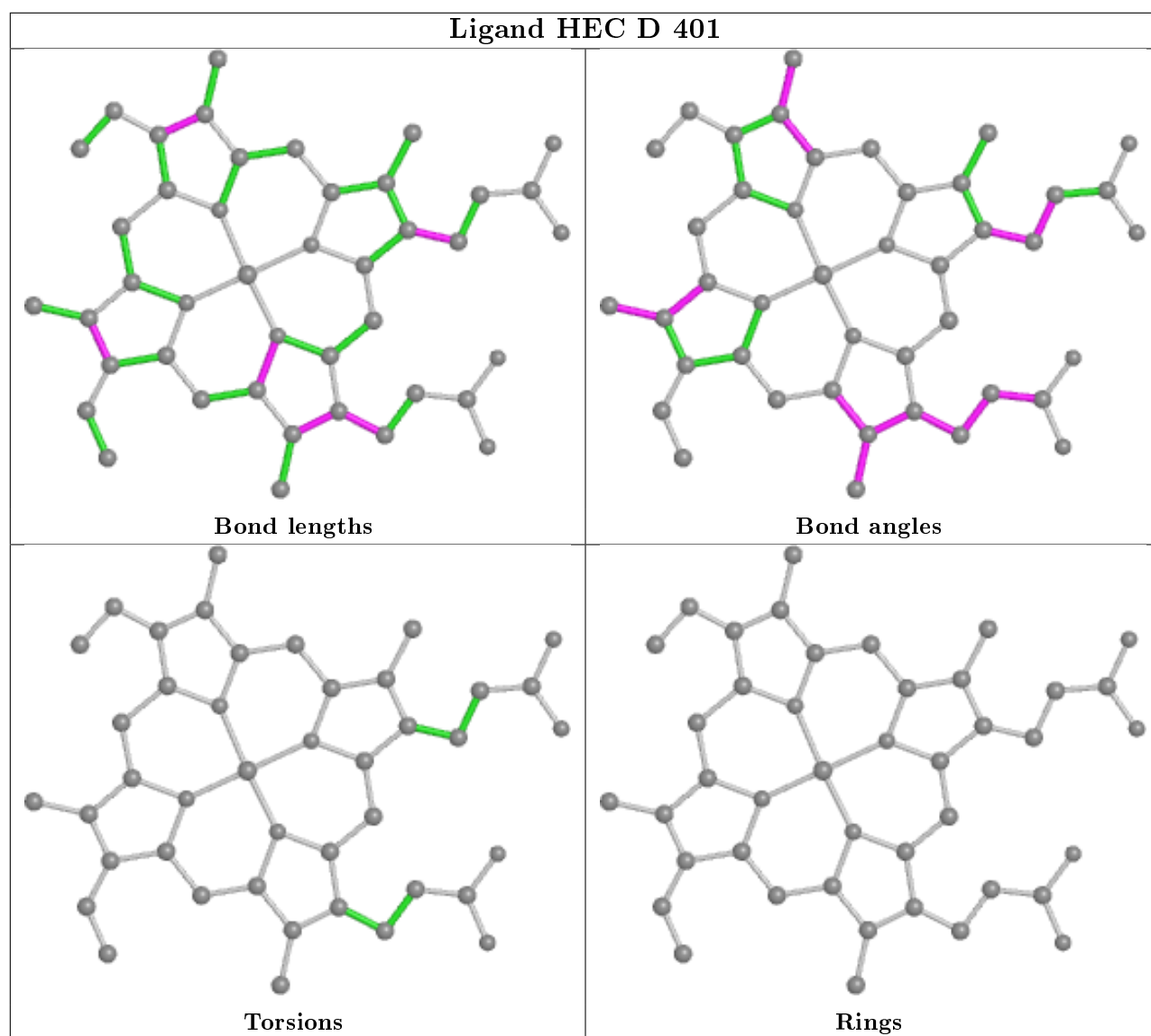












5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

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.