



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 27, 2021 – 08:36 PM EST

PDB ID : 1SP3  
Title : Crystal structure of octaheme cytochrome c from *Shewanella oneidensis*  
Authors : Mowat, C.G.; Rothery, E.; Miles, C.S.; McIver, L.; Doherty, M.K.; Drewette, K.; Taylor, P.; Walkinshaw, M.D.; Chapman, S.K.; Reid, G.A.  
Deposited on : 2004-03-16  
Resolution : 2.20 Å (reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto: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) : 1.13  
EDS : 2.17.1  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.17.1

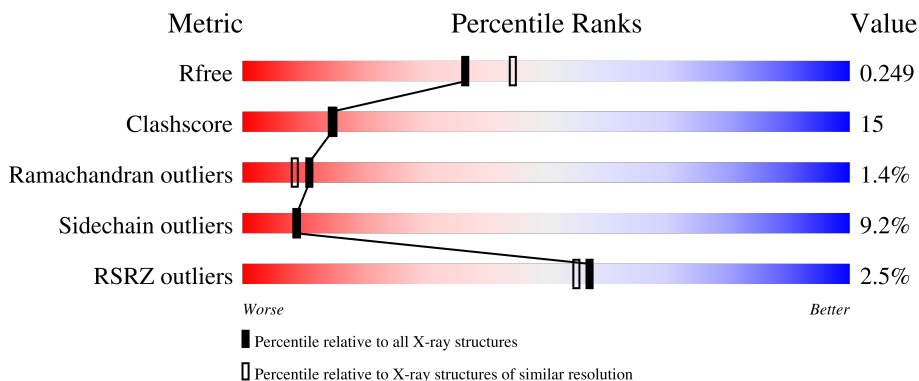
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.20 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)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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  $\geq 3$ , 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  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	443	 2% 71% 21% 6% ..

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 4243 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.

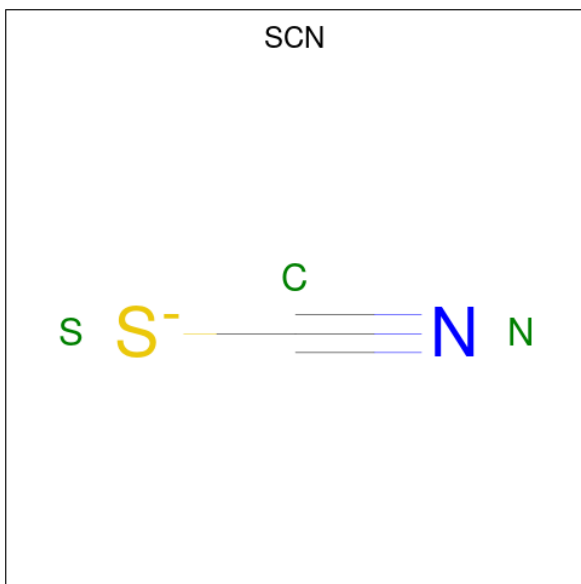
- Molecule 1 is a protein called cytochrome c, putative.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	436	3426	2140	598	655	33	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	211	ASN	GLN	conflict	UNP Q8E9W8

- Molecule 2 is THIOCYANATE ION (three-letter code: SCN) (formula: CNS).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	S		
2	A	1	3	1	1	1	0	0

- Molecule 3 is HEME C (three-letter code: HEC) (formula: C<sub>34</sub>H<sub>34</sub>FeN<sub>4</sub>O<sub>4</sub>).



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

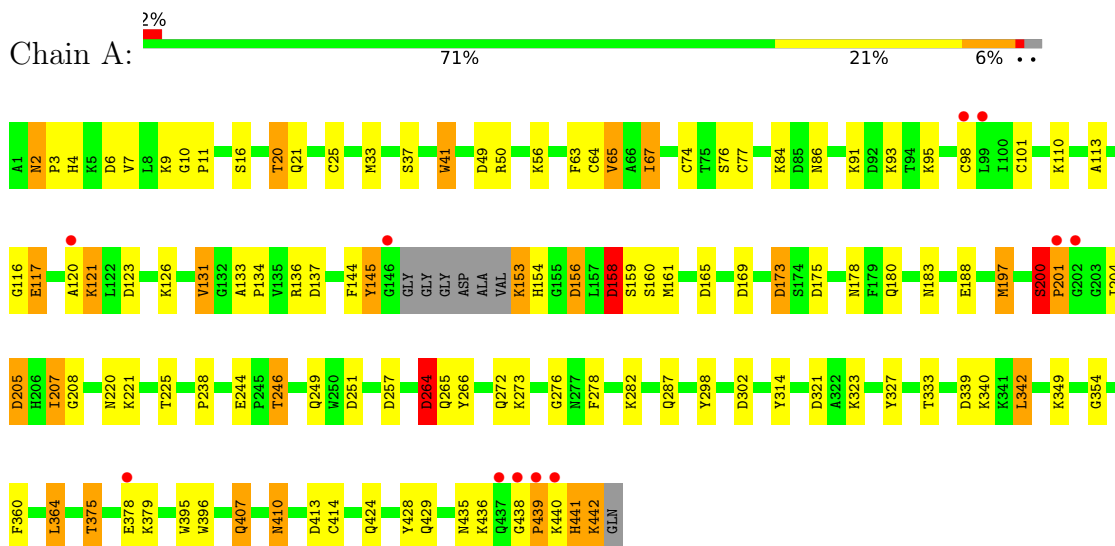
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	470	Total	O	0	0
			470	470		

### 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 and electron density. 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. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). 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: cytochrome c, putative



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	40.59Å 61.77Å 185.24Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	23.77 – 2.20 23.76 – 2.20	Depositor EDS
% Data completeness (in resolution range)	100.0 (23.77-2.20) 96.6 (23.76-2.20)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.33 (at 2.19Å)	Xtrriage
Refinement program	REFMAC 5.1.24	Depositor
R, $R_{free}$	0.155 , 0.252 0.158 , 0.249	Depositor DCC
$R_{free}$ test set	1217 reflections (5.14%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	35.5	Xtrriage
Anisotropy	0.044	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 41.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	4243	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	37.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.08% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 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: HEC, SCN

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	1.20	4/3519 (0.1%)	1.15	22/4760 (0.5%)

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	4

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	314	TYR	CD2-CE2	7.79	1.51	1.39
1	A	327	TYR	CD2-CE2	5.74	1.48	1.39
1	A	37	SER	CB-OG	-5.40	1.35	1.42
1	A	41	TRP	CB-CG	5.23	1.59	1.50

All (22) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	342	LEU	C-N-CA	-8.74	99.85	121.70
1	A	165	ASP	CB-CG-OD2	7.25	124.82	118.30
1	A	207	ILE	C-N-CA	-7.02	107.56	122.30
1	A	156	ASP	CB-CG-OD2	6.89	124.50	118.30
1	A	205	ASP	CB-CG-OD2	6.84	124.45	118.30
1	A	413	ASP	CB-CG-OD2	6.56	124.21	118.30
1	A	208	GLY	N-CA-C	-6.55	96.73	113.10
1	A	50	ARG	NE-CZ-NH2	-6.03	117.29	120.30
1	A	6	ASP	CB-CG-OD2	5.75	123.47	118.30

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	175	ASP	CB-CG-OD2	5.68	123.41	118.30
1	A	342	LEU	O-C-N	-5.64	113.67	122.70
1	A	173	ASP	CB-CG-OD1	5.60	123.34	118.30
1	A	137	ASP	CB-CG-OD2	5.51	123.26	118.30
1	A	246	THR	CB-CA-C	-5.46	96.86	111.60
1	A	158	ASP	CB-CG-OD1	5.32	123.09	118.30
1	A	131	VAL	N-CA-CB	-5.18	100.10	111.50
1	A	264	ASP	CB-CG-OD2	5.18	122.96	118.30
1	A	207	ILE	CB-CA-C	-5.07	101.46	111.60
1	A	257	ASP	CB-CG-OD2	5.06	122.85	118.30
1	A	91	LYS	CD-CE-NZ	5.05	123.32	111.70
1	A	123	ASP	CB-CG-OD2	5.04	122.84	118.30
1	A	169	ASP	CB-CG-OD2	5.00	122.80	118.30

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	10	GLY	Peptide
1	A	153	LYS	Peptide
1	A	264	ASP	Peptide
1	A	441	HIS	Peptide

## 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	3426	0	3243	99	0
2	A	3	0	0	0	0
3	A	344	0	243	28	0
4	A	470	0	0	24	4
All	All	4243	0	3486	106	4

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 15.

All (106) 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:74:CYS:SG	3:A:802:HEC:CAB	2.10	1.38
1:A:25:CYS:SG	3:A:801:HEC:HAC	1.77	1.15
1:A:429:GLN:H	1:A:439:PRO:HG3	1.17	1.07
1:A:101:CYS:SG	3:A:803:HEC:HAC	1.94	1.04
1:A:74:CYS:SG	3:A:802:HEC:HAB	2.00	1.00
1:A:9:LYS:HG3	4:A:1183:HOH:O	1.67	0.94
1:A:410:ASN:HA	4:A:1048:HOH:O	1.71	0.91
1:A:429:GLN:H	1:A:439:PRO:CG	1.90	0.83
1:A:74:CYS:SG	3:A:802:HEC:CBB	2.68	0.82
1:A:9:LYS:CG	4:A:1183:HOH:O	2.26	0.81
1:A:429:GLN:N	1:A:439:PRO:HG3	1.95	0.81
1:A:101:CYS:SG	3:A:803:HEC:C3C	2.69	0.80
1:A:200:SER:HB3	1:A:201:PRO:CA	2.12	0.79
1:A:200:SER:HB3	1:A:201:PRO:HA	1.64	0.79
1:A:439:PRO:HA	4:A:879:HOH:O	1.86	0.75
1:A:33:MET:HG2	4:A:1108:HOH:O	1.88	0.74
1:A:56:LYS:HE2	1:A:67:ILE:HD11	1.69	0.74
1:A:77:CYS:SG	3:A:802:HEC:C3C	2.78	0.72
1:A:101:CYS:SG	3:A:803:HEC:CBC	2.77	0.71
1:A:244:GLU:HG2	4:A:848:HOH:O	1.90	0.71
1:A:25:CYS:HG	3:A:801:HEC:HAC	1.53	0.71
1:A:439:PRO:HG2	4:A:962:HOH:O	1.91	0.71
1:A:221:LYS:HD3	4:A:1110:HOH:O	1.90	0.71
1:A:414:CYS:SG	3:A:808:HEC:C3C	2.79	0.70
1:A:441:HIS:O	1:A:442:LYS:HB2	1.91	0.70
1:A:414:CYS:SG	3:A:808:HEC:CBC	2.80	0.66
1:A:113:ALA:HB3	4:A:1051:HOH:O	1.94	0.66
1:A:25:CYS:SG	3:A:801:HEC:C3C	2.84	0.66
1:A:264:ASP:HB3	1:A:266:TYR:H	1.59	0.66
1:A:197:MET:HE1	1:A:395:TRP:O	1.95	0.66
1:A:49:ASP:CB	4:A:975:HOH:O	2.44	0.65
1:A:74:CYS:SG	3:A:802:HEC:C3B	2.82	0.65
1:A:154:HIS:CE1	1:A:396:TRP:HB3	2.30	0.65
1:A:64:CYS:SG	4:A:1263:HOH:O	2.54	0.65
1:A:287:GLN:NE2	4:A:1184:HOH:O	2.29	0.65
1:A:273:LYS:HE2	1:A:360:PHE:CE1	2.32	0.64
1:A:21:GLN:OE1	4:A:925:HOH:O	2.15	0.64
1:A:200:SER:HB3	1:A:201:PRO:O	1.98	0.64
1:A:84:LYS:HE3	4:A:1086:HOH:O	1.98	0.63
1:A:49:ASP:HB2	4:A:975:HOH:O	1.99	0.62
1:A:56:LYS:HE3	3:A:802:HEC:NA	2.14	0.62
1:A:25:CYS:SG	3:A:801:HEC:CBC	2.81	0.62

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:77:CYS:SG	3:A:802:HEC:CBC	2.85	0.61
1:A:298:TYR:CE1	1:A:302:ASP:HB2	2.36	0.60
1:A:16:SER:O	1:A:20:THR:CG2	2.50	0.59
1:A:200:SER:HB3	1:A:201:PRO:C	2.23	0.59
1:A:56:LYS:CE	1:A:67:ILE:HD11	2.33	0.58
1:A:273:LYS:HE2	1:A:360:PHE:CZ	2.37	0.58
1:A:101:CYS:HG	3:A:803:HEC:HAC	1.69	0.58
1:A:272:GLN:HE21	1:A:278:PHE:HE1	1.51	0.58
1:A:95:LYS:HE3	4:A:905:HOH:O	2.06	0.56
1:A:2:ASN:HD22	1:A:4:HIS:H	1.54	0.55
1:A:428:TYR:HA	1:A:439:PRO:CB	2.36	0.55
1:A:16:SER:O	1:A:20:THR:HG22	2.07	0.55
1:A:438:GLY:H	1:A:439:PRO:HD3	1.70	0.55
1:A:56:LYS:HE3	3:A:802:HEC:C4A	2.37	0.54
1:A:7:VAL:HG12	1:A:7:VAL:O	2.05	0.54
1:A:339:ASP:O	1:A:342:LEU:O	2.24	0.54
3:A:807:HEC:HBB3	3:A:807:HEC:HMB1	1.90	0.53
1:A:41:TRP:NE1	3:A:804:HEC:HMD2	2.24	0.52
1:A:180:GLN:H	1:A:183:ASN:HD22	1.57	0.52
1:A:49:ASP:HB3	4:A:975:HOH:O	2.10	0.51
1:A:354:GLY:O	1:A:364:LEU:HD13	2.11	0.51
1:A:205:ASP:HB3	1:A:207:ILE:O	2.11	0.51
3:A:802:HEC:HMB1	3:A:802:HEC:HBB3	1.92	0.51
1:A:56:LYS:HE2	1:A:67:ILE:CD1	2.40	0.50
1:A:438:GLY:N	1:A:439:PRO:HD3	2.26	0.50
1:A:144:PHE:CZ	1:A:161:MET:HB2	2.46	0.50
1:A:16:SER:O	1:A:20:THR:HG23	2.13	0.49
1:A:180:GLN:H	1:A:183:ASN:ND2	2.10	0.49
1:A:33:MET:CG	4:A:1108:HOH:O	2.55	0.48
1:A:136:ARG:HH12	1:A:178:ASN:ND2	2.11	0.48
1:A:375:THR:OG1	4:A:1237:HOH:O	2.20	0.48
3:A:804:HEC:HBC3	3:A:804:HEC:HMC1	1.96	0.48
1:A:438:GLY:N	4:A:936:HOH:O	2.46	0.47
1:A:98:CYS:HB3	3:A:801:HEC:HMB3	1.95	0.47
1:A:63:PHE:HE1	1:A:396:TRP:CH2	2.33	0.47
1:A:120:ALA:O	1:A:121:LYS:HB2	2.13	0.47
1:A:428:TYR:CD2	1:A:439:PRO:HB2	2.49	0.47
1:A:429:GLN:HB2	1:A:439:PRO:HG3	1.96	0.47
3:A:801:HEC:HMB1	3:A:801:HEC:HBB3	1.98	0.46
1:A:156:ASP:O	1:A:161:MET:HE2	2.17	0.45
3:A:801:HEC:HMC1	3:A:801:HEC:HBC3	1.98	0.45

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:238:PRO:HG3	1:A:407:GLN:O	2.17	0.45
1:A:2:ASN:HB2	1:A:3:PRO:CD	2.48	0.44
1:A:74:CYS:SG	3:A:802:HEC:HBB3	2.57	0.43
1:A:428:TYR:HA	1:A:439:PRO:HB3	2.00	0.43
1:A:158:ASP:H	1:A:161:MET:HG2	1.83	0.43
3:A:806:HEC:HBC3	3:A:806:HEC:HMC1	1.99	0.43
1:A:110:LYS:HE3	1:A:116:GLY:HA2	2.00	0.43
1:A:7:VAL:O	1:A:7:VAL:CG1	2.67	0.42
1:A:145:TYR:CD1	1:A:145:TYR:N	2.87	0.42
1:A:321:ASP:O	1:A:441:HIS:HE1	2.02	0.42
1:A:349:LYS:NZ	4:A:1137:HOH:O	2.18	0.42
3:A:803:HEC:HHA	3:A:803:HEC:HBA1	2.02	0.41
1:A:117:GLU:OE1	4:A:845:HOH:O	2.22	0.41
1:A:220:ASN:C	1:A:220:ASN:OD1	2.58	0.41
1:A:249:GLN:HA	1:A:333:THR:O	2.21	0.41
1:A:64:CYS:O	1:A:65:VAL:HB	2.21	0.41
1:A:197:MET:CE	4:A:1047:HOH:O	2.68	0.41
1:A:251:ASP:O	1:A:276:GLY:HA3	2.21	0.41
1:A:180:GLN:O	1:A:183:ASN:HB2	2.22	0.40
1:A:133:ALA:HA	1:A:134:PRO:HD3	1.91	0.40
1:A:204:ILE:HD11	1:A:323:LYS:O	2.22	0.40
1:A:117:GLU:HG3	4:A:845:HOH:O	2.21	0.40
1:A:375:THR:O	1:A:379:LYS:HG3	2.22	0.40

All (4) 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)
4:A:975:HOH:O	4:A:1274:HOH:O[1_455]	1.49	0.71
4:A:971:HOH:O	4:A:1276:HOH:O[4_455]	1.53	0.67
4:A:1075:HOH:O	4:A:1123:HOH:O[3_545]	2.04	0.16
4:A:1044:HOH:O	4:A:1123:HOH:O[3_545]	2.07	0.13

## 5.3 Torsion angles

### 5.3.1 Protein backbone

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	432/443 (98%)	406 (94%)	20 (5%)	6 (1%)	11 8

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	200	SER
1	A	201	PRO
1	A	11	PRO
1	A	173	ASP
1	A	439	PRO
1	A	65	VAL

### 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	370/373 (99%)	336 (91%)	34 (9%)	9 9

All (34) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	2	ASN
1	A	20	THR
1	A	67	ILE
1	A	76	SER
1	A	86	ASN
1	A	93	LYS
1	A	117	GLU
1	A	121	LYS
1	A	126	LYS
1	A	131	VAL
1	A	145	TYR
1	A	153	LYS
1	A	158	ASP

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	159	SER
1	A	160	SER
1	A	188	GLU
1	A	197	MET
1	A	200	SER
1	A	225	THR
1	A	246	THR
1	A	264	ASP
1	A	265	GLN
1	A	282	LYS
1	A	340	LYS
1	A	364	LEU
1	A	375	THR
1	A	378	GLU
1	A	407	GLN
1	A	410	ASN
1	A	424	GLN
1	A	435	ASN
1	A	436	LYS
1	A	440	LYS
1	A	442	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (12) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	2	ASN
1	A	154	HIS
1	A	177	ASN
1	A	178	ASN
1	A	183	ASN
1	A	249	GLN
1	A	272	GLN
1	A	277	ASN
1	A	363	ASN
1	A	407	GLN
1	A	429	GLN
1	A	441	HIS

### 5.3.3 RNA

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](#)

9 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	HEC	A	808	1	26,50,50	2.53	8 (30%)	18,82,82	1.93	7 (38%)
3	HEC	A	806	1	26,50,50	2.16	6 (23%)	18,82,82	2.68	7 (38%)
3	HEC	A	803	1	26,50,50	2.92	13 (50%)	18,82,82	1.91	5 (27%)
3	HEC	A	801	1	26,50,50	2.36	8 (30%)	18,82,82	2.73	7 (38%)
3	HEC	A	804	1	26,50,50	2.68	9 (34%)	18,82,82	3.02	7 (38%)
2	SCN	A	810	3	1,2,2	0.53	0	0,1,1	0.00	-
3	HEC	A	805	1	26,50,50	2.31	10 (38%)	18,82,82	2.99	9 (50%)
3	HEC	A	807	1	26,50,50	2.69	8 (30%)	18,82,82	2.88	11 (61%)
3	HEC	A	802	1,2	26,50,50	2.50	14 (53%)	18,82,82	2.37	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
3	HEC	A	808	1	-	0/6/54/54	-
3	HEC	A	806	1	-	0/6/54/54	-
3	HEC	A	803	1	-	0/6/54/54	-
3	HEC	A	801	1	-	0/6/54/54	-

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	HEC	A	804	1	-	1/6/54/54	-
3	HEC	A	805	1	-	0/6/54/54	-
3	HEC	A	807	1	-	1/6/54/54	-
3	HEC	A	802	1,2	-	0/6/54/54	-

All (76) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	807	HEC	C3B-C2B	-8.83	1.31	1.40
3	A	803	HEC	C3B-C2B	-7.58	1.32	1.40
3	A	808	HEC	C3C-C2C	-7.01	1.33	1.40
3	A	804	HEC	C3B-C2B	-6.98	1.33	1.40
3	A	803	HEC	C3C-C2C	-6.72	1.33	1.40
3	A	804	HEC	C3C-C2C	-6.64	1.33	1.40
3	A	801	HEC	C3B-C2B	-6.51	1.34	1.40
3	A	808	HEC	C3B-C2B	-6.07	1.34	1.40
3	A	806	HEC	C3B-C2B	-5.95	1.34	1.40
3	A	802	HEC	C3D-C2D	5.46	1.53	1.37
3	A	803	HEC	C3D-C2D	5.38	1.53	1.37
3	A	805	HEC	C3B-C2B	-5.28	1.35	1.40
3	A	806	HEC	C3D-C2D	5.24	1.53	1.37
3	A	807	HEC	CAD-C3D	5.18	1.59	1.52
3	A	801	HEC	C3C-C2C	-4.97	1.35	1.40
3	A	804	HEC	CAA-C2A	4.66	1.60	1.52
3	A	805	HEC	C3C-C2C	-4.54	1.36	1.40
3	A	802	HEC	CAD-C3D	4.38	1.58	1.52
3	A	802	HEC	C3B-C2B	-4.37	1.36	1.40
3	A	805	HEC	CMD-C2D	4.35	1.60	1.51
3	A	808	HEC	C3D-C2D	4.06	1.49	1.37
3	A	803	HEC	CBC-CAC	-3.83	1.35	1.49
3	A	804	HEC	C3D-C2D	3.81	1.48	1.37
3	A	805	HEC	CAD-C3D	3.76	1.57	1.52
3	A	801	HEC	CBB-CAB	-3.75	1.35	1.49
3	A	806	HEC	CBB-CAB	-3.71	1.35	1.49
3	A	801	HEC	C3D-C2D	3.64	1.48	1.37
3	A	801	HEC	CBC-CAC	-3.62	1.35	1.49
3	A	807	HEC	CBB-CAB	-3.56	1.36	1.49
3	A	802	HEC	CAA-C2A	3.39	1.58	1.52
3	A	807	HEC	CMD-C2D	3.37	1.58	1.51
3	A	808	HEC	CMB-C2B	3.35	1.59	1.51
3	A	802	HEC	CMA-C3A	3.33	1.59	1.51
3	A	803	HEC	CMD-C2D	3.31	1.58	1.51

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	807	HEC	C3C-C2C	-3.29	1.37	1.40
3	A	806	HEC	CBC-CAC	-3.29	1.37	1.49
3	A	808	HEC	CBC-CAC	-3.23	1.37	1.49
3	A	803	HEC	CBB-CAB	-3.19	1.37	1.49
3	A	807	HEC	C3D-C2D	3.15	1.47	1.37
3	A	808	HEC	CBB-CAB	-3.15	1.37	1.49
3	A	802	HEC	C4D-ND	3.04	1.42	1.36
3	A	802	HEC	CBB-CAB	-2.99	1.38	1.49
3	A	803	HEC	C4A-C3A	2.98	1.49	1.42
3	A	802	HEC	CBC-CAC	-2.92	1.38	1.49
3	A	807	HEC	CBC-CAC	-2.91	1.38	1.49
3	A	805	HEC	CBB-CAB	-2.85	1.38	1.49
3	A	805	HEC	C3D-C2D	2.83	1.46	1.37
3	A	806	HEC	C3C-C2C	-2.79	1.37	1.40
3	A	802	HEC	C1B-NB	2.79	1.41	1.36
3	A	805	HEC	CBC-CAC	-2.75	1.39	1.49
3	A	804	HEC	CAD-C3D	2.74	1.56	1.52
3	A	805	HEC	CMC-C2C	2.72	1.58	1.51
3	A	804	HEC	CBB-CAB	-2.69	1.39	1.49
3	A	803	HEC	C3C-C4C	2.65	1.47	1.43
3	A	801	HEC	CMC-C2C	2.53	1.57	1.51
3	A	801	HEC	CAA-C2A	2.51	1.56	1.52
3	A	804	HEC	CBC-CAC	-2.41	1.40	1.49
3	A	803	HEC	CMB-C2B	2.41	1.57	1.51
3	A	807	HEC	C1D-CHD	-2.36	1.34	1.41
3	A	802	HEC	C3C-C2C	-2.33	1.38	1.40
3	A	804	HEC	C1C-NC	2.32	1.40	1.36
3	A	803	HEC	CMA-C3A	2.30	1.57	1.51
3	A	806	HEC	CAA-C2A	2.27	1.56	1.52
3	A	808	HEC	C1A-C2A	2.27	1.47	1.42
3	A	802	HEC	CMB-C2B	2.26	1.57	1.51
3	A	802	HEC	C3B-C4B	2.24	1.47	1.43
3	A	801	HEC	CMD-C2D	2.21	1.56	1.51
3	A	803	HEC	C1A-C2A	2.19	1.47	1.42
3	A	802	HEC	C4A-C3A	2.18	1.47	1.42
3	A	803	HEC	C1D-ND	2.17	1.40	1.36
3	A	808	HEC	CMA-C3A	2.14	1.56	1.51
3	A	803	HEC	CAA-C2A	2.12	1.55	1.52
3	A	804	HEC	CMD-C2D	2.11	1.56	1.51
3	A	805	HEC	C1C-CHC	-2.09	1.35	1.41
3	A	805	HEC	CMB-C2B	2.08	1.56	1.51
3	A	802	HEC	CMC-C2C	2.04	1.56	1.51



All (60) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	806	HEC	CBD-CAD-C3D	-6.91	99.74	112.49
3	A	801	HEC	CBA-CAA-C2A	-6.50	100.50	112.48
3	A	804	HEC	CMB-C2B-C3B	6.25	133.17	125.82
3	A	804	HEC	CMB-C2B-C1B	-6.02	119.22	128.46
3	A	804	HEC	CBD-CAD-C3D	-5.99	101.44	112.49
3	A	805	HEC	CMC-C2C-C3C	5.69	132.52	125.82
3	A	805	HEC	CBD-CAD-C3D	-5.34	102.63	112.49
3	A	801	HEC	CBD-CAD-C3D	-5.14	103.01	112.49
3	A	805	HEC	CMC-C2C-C1C	-4.95	120.86	128.46
3	A	806	HEC	C1D-C2D-C3D	-4.78	103.67	107.00
3	A	801	HEC	C1D-C2D-C3D	-4.76	103.69	107.00
3	A	807	HEC	CAD-CBD-CGD	-4.63	104.90	112.67
3	A	802	HEC	CMC-C2C-C3C	4.48	131.09	125.82
3	A	805	HEC	CBA-CAA-C2A	-4.43	104.32	112.48
3	A	807	HEC	CBD-CAD-C3D	-4.40	104.37	112.49
3	A	802	HEC	CMB-C2B-C3B	4.30	130.88	125.82
3	A	805	HEC	CAA-CBA-CGA	-4.29	105.47	112.67
3	A	802	HEC	CMC-C2C-C1C	-4.28	121.89	128.46
3	A	807	HEC	C1D-C2D-C3D	-4.27	104.03	107.00
3	A	807	HEC	CMC-C2C-C1C	-4.20	122.01	128.46
3	A	807	HEC	CBA-CAA-C2A	-4.04	105.04	112.48
3	A	803	HEC	CBD-CAD-C3D	-3.98	105.15	112.49
3	A	807	HEC	CMC-C2C-C3C	3.79	130.28	125.82
3	A	806	HEC	CMC-C2C-C1C	-3.76	122.68	128.46
3	A	808	HEC	CAD-CBD-CGD	-3.76	106.36	112.67
3	A	804	HEC	C4C-C3C-C2C	3.76	110.41	106.35
3	A	807	HEC	CMB-C2B-C1B	-3.43	123.19	128.46
3	A	803	HEC	CMC-C2C-C1C	-3.41	123.22	128.46
3	A	804	HEC	C1D-C2D-C3D	-3.34	104.67	107.00
3	A	801	HEC	CMC-C2C-C1C	-3.27	123.44	128.46
3	A	806	HEC	CMC-C2C-C3C	3.21	129.59	125.82
3	A	802	HEC	CMB-C2B-C1B	-3.17	123.59	128.46
3	A	802	HEC	CBD-CAD-C3D	-3.16	106.66	112.49
3	A	805	HEC	C1D-C2D-C3D	-3.11	104.83	107.00
3	A	806	HEC	CAD-CBD-CGD	-2.82	107.94	112.67
3	A	803	HEC	C1D-C2D-C3D	-2.81	105.04	107.00
3	A	808	HEC	CBD-CAD-C3D	-2.81	107.31	112.49
3	A	808	HEC	CBA-CAA-C2A	-2.81	107.31	112.48
3	A	804	HEC	C3C-C4C-NC	-2.70	105.85	110.94
3	A	802	HEC	C4B-C3B-C2B	2.59	109.15	106.35
3	A	808	HEC	CMC-C2C-C1C	-2.58	124.50	128.46
3	A	804	HEC	CMD-C2D-C1D	2.55	132.39	128.46

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	806	HEC	CMB-C2B-C1B	-2.55	124.55	128.46
3	A	803	HEC	CMB-C2B-C1B	-2.51	124.61	128.46
3	A	808	HEC	CMC-C2C-C3C	2.49	128.75	125.82
3	A	801	HEC	C4B-C3B-C2B	2.49	109.04	106.35
3	A	801	HEC	CMC-C2C-C3C	2.45	128.71	125.82
3	A	801	HEC	CMB-C2B-C1B	-2.37	124.83	128.46
3	A	803	HEC	CMC-C2C-C3C	2.35	128.58	125.82
3	A	807	HEC	CMB-C2B-C3B	2.31	128.53	125.82
3	A	807	HEC	C3C-C4C-NC	-2.28	106.65	110.94
3	A	805	HEC	CMB-C2B-C1B	-2.27	124.98	128.46
3	A	808	HEC	C4C-C3C-C2C	2.25	108.78	106.35
3	A	807	HEC	C4B-C3B-C2B	2.23	108.76	106.35
3	A	805	HEC	C4B-C3B-C2B	2.22	108.75	106.35
3	A	807	HEC	CMD-C2D-C3D	2.22	129.12	124.94
3	A	805	HEC	C4C-C3C-C2C	2.20	108.72	106.35
3	A	802	HEC	C3C-C4C-NC	-2.19	106.80	110.94
3	A	806	HEC	CAA-CBA-CGA	-2.08	109.18	112.67
3	A	808	HEC	CMB-C2B-C1B	-2.05	125.31	128.46

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	804	HEC	C2A-CAA-CBA-CGA
3	A	807	HEC	C3D-CAD-CBD-CGD

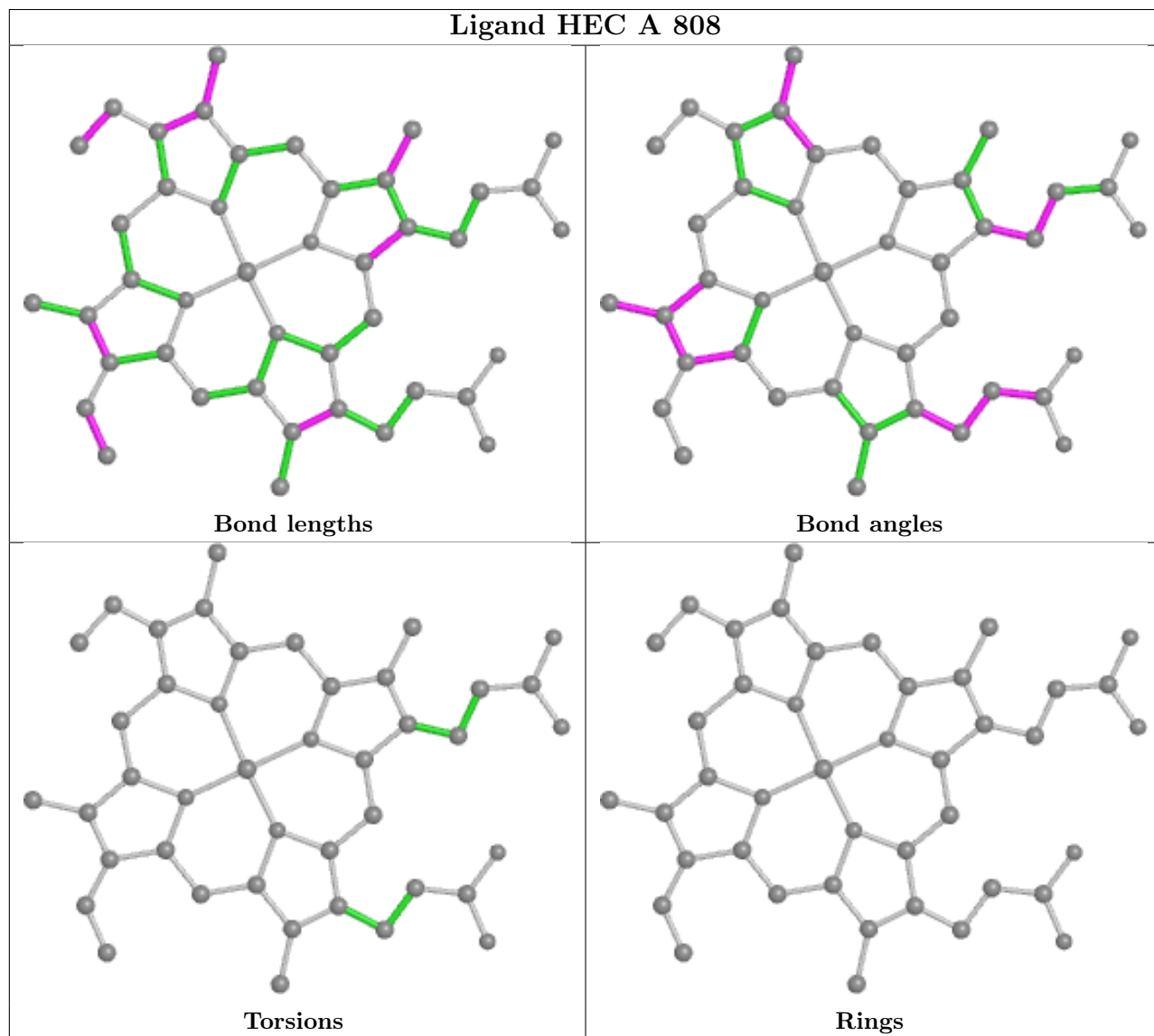
There are no ring outliers.

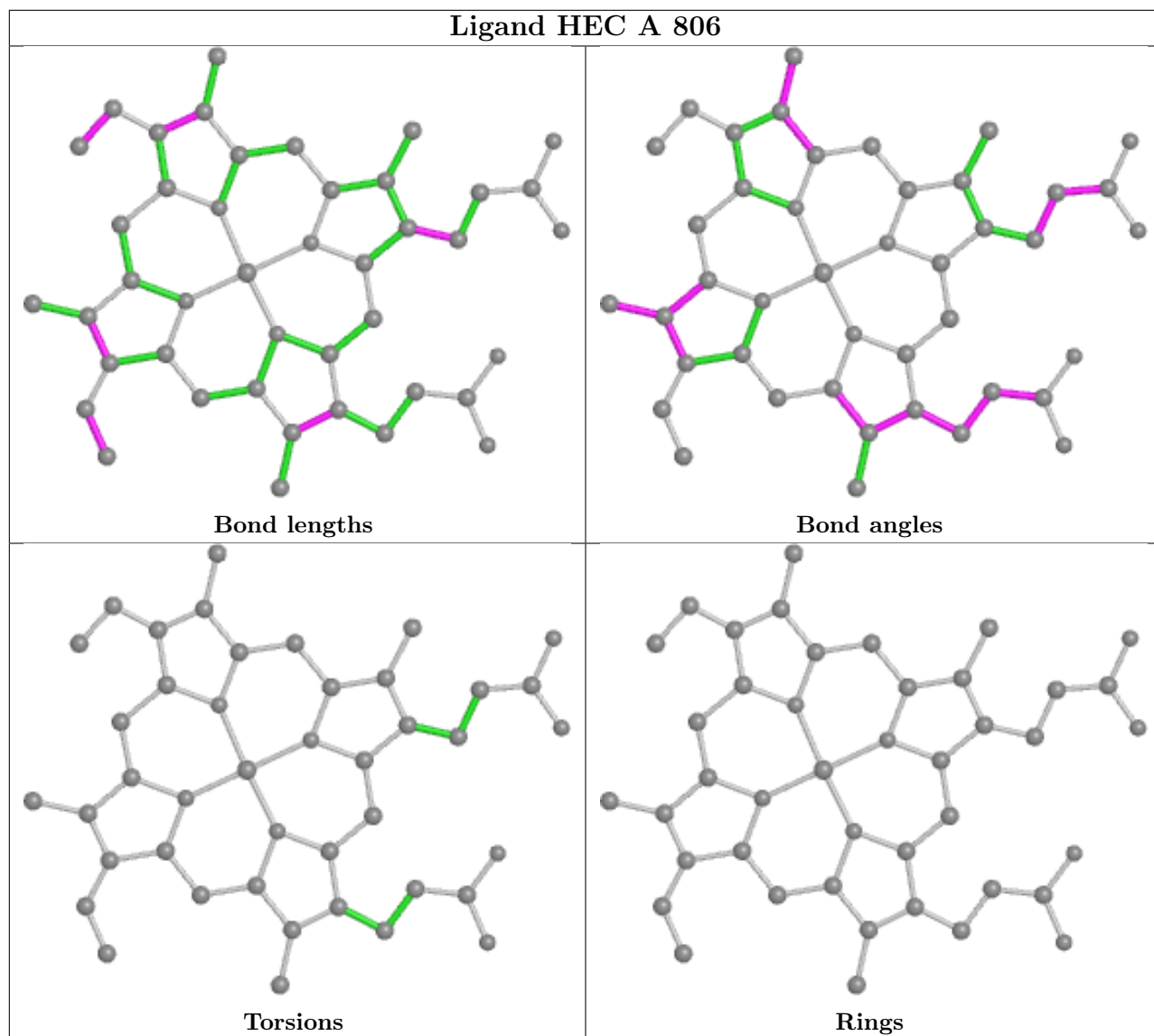
7 monomers are involved in 28 short contacts:

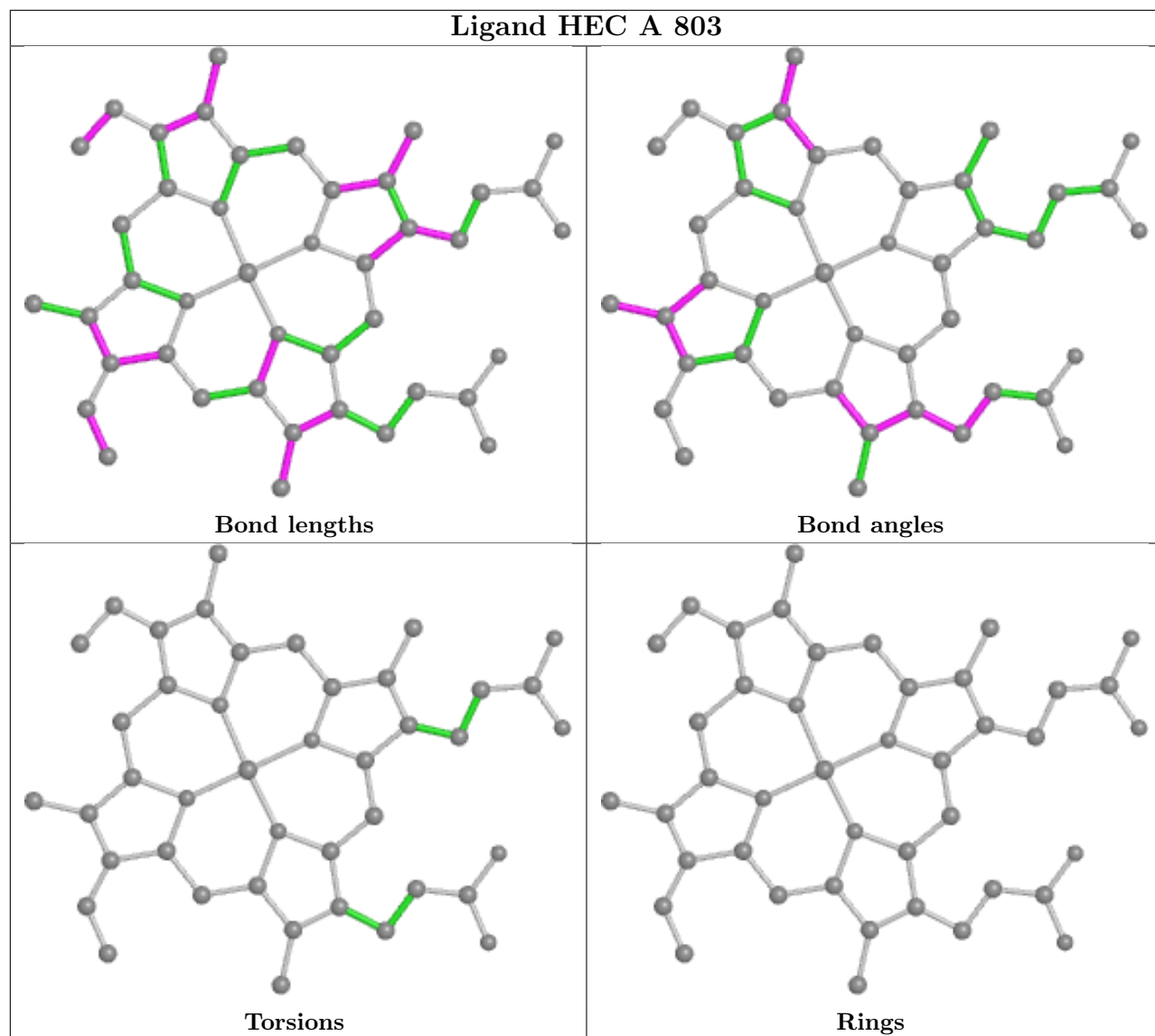
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	808	HEC	2	0
3	A	806	HEC	1	0
3	A	803	HEC	5	0
3	A	801	HEC	7	0
3	A	804	HEC	2	0
3	A	807	HEC	1	0
3	A	802	HEC	10	0

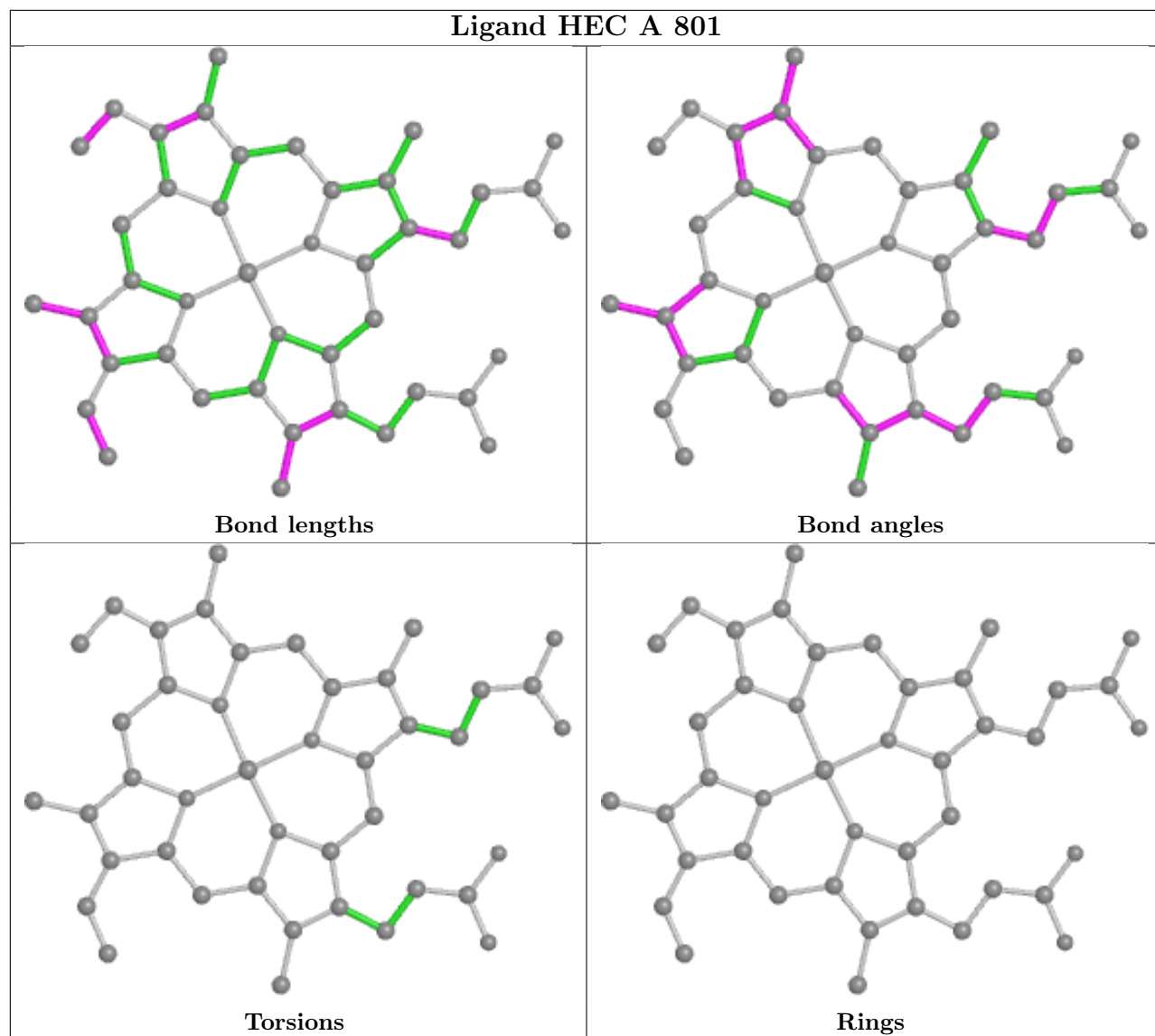
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will

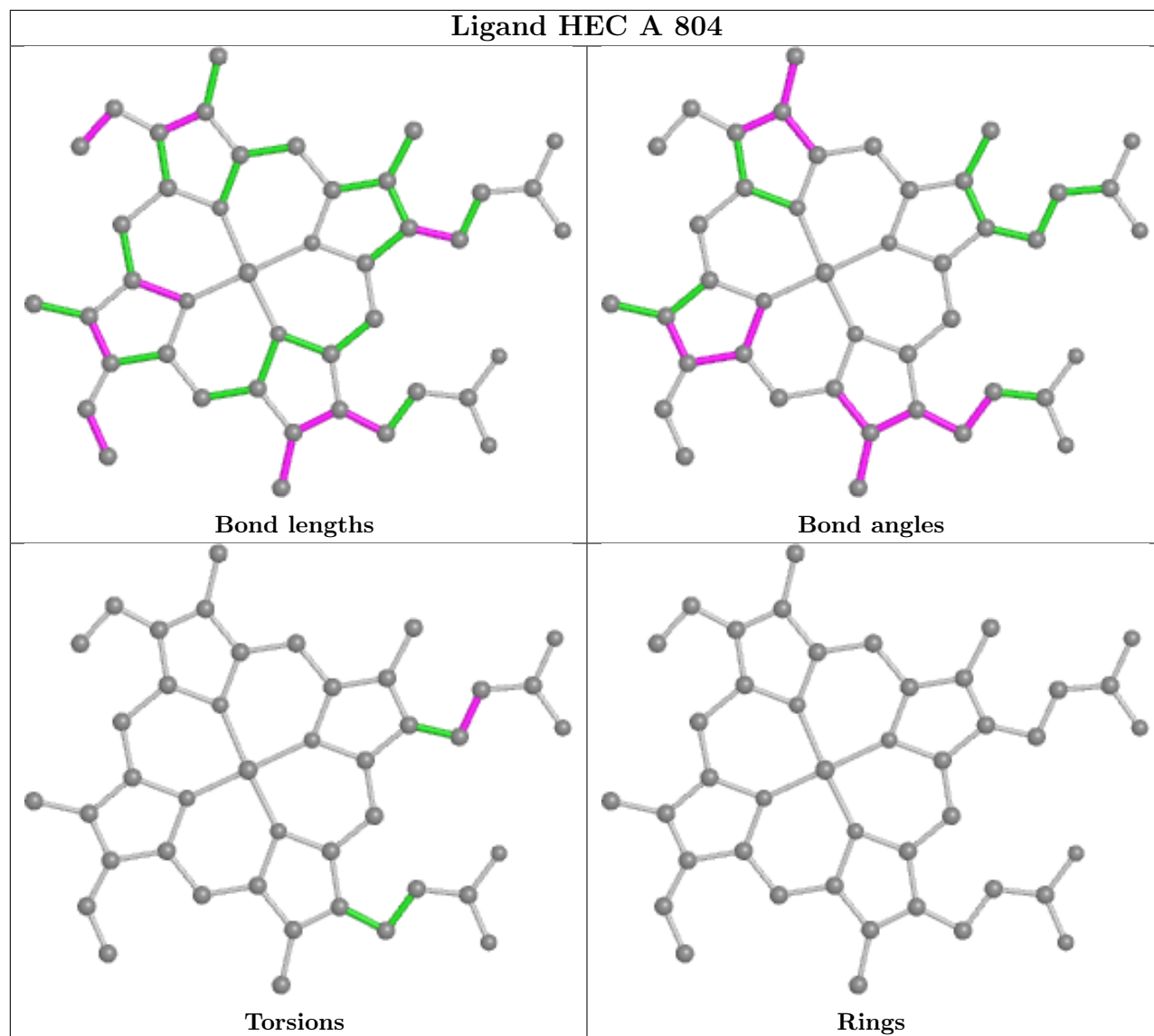
also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted 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.

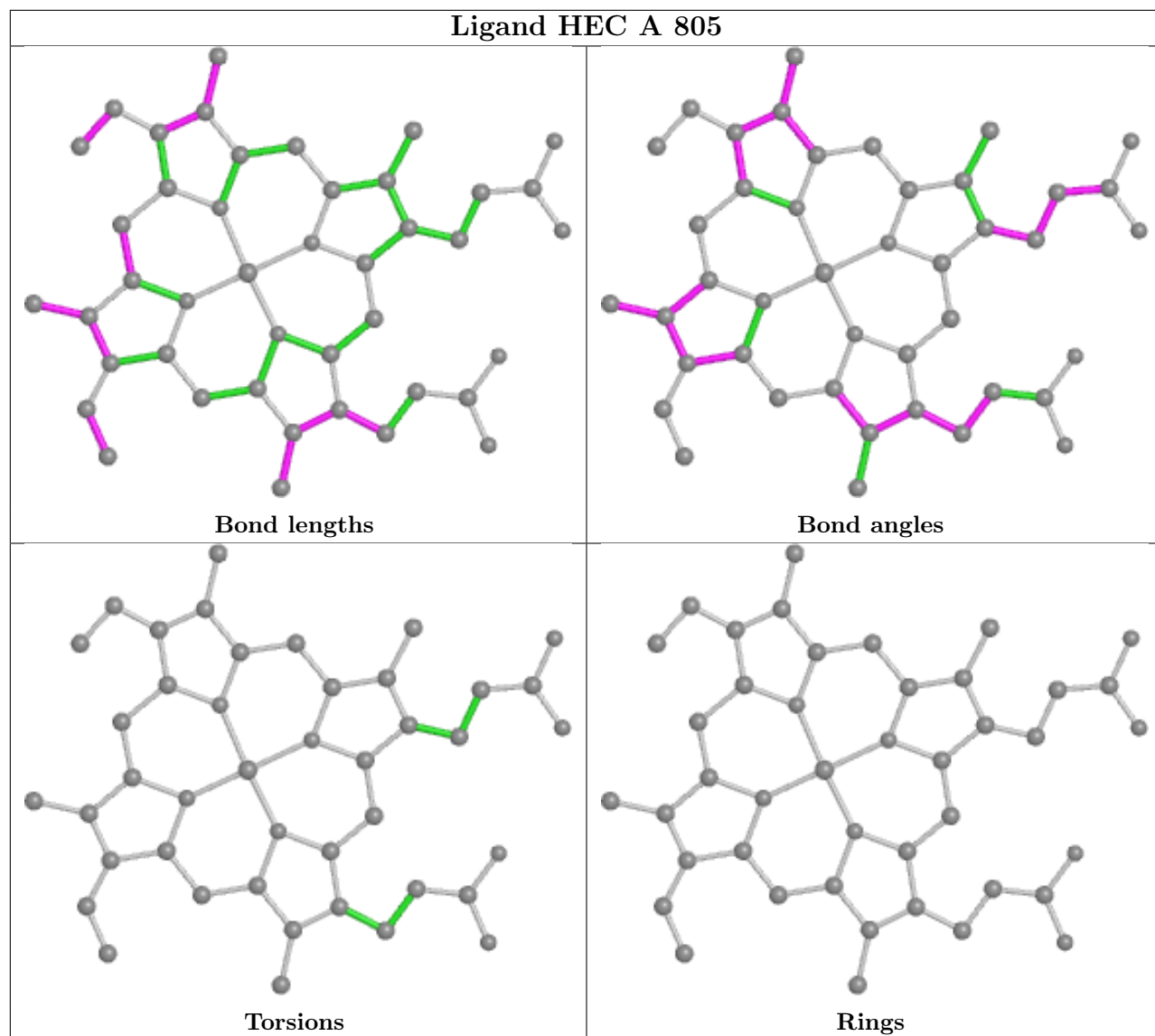




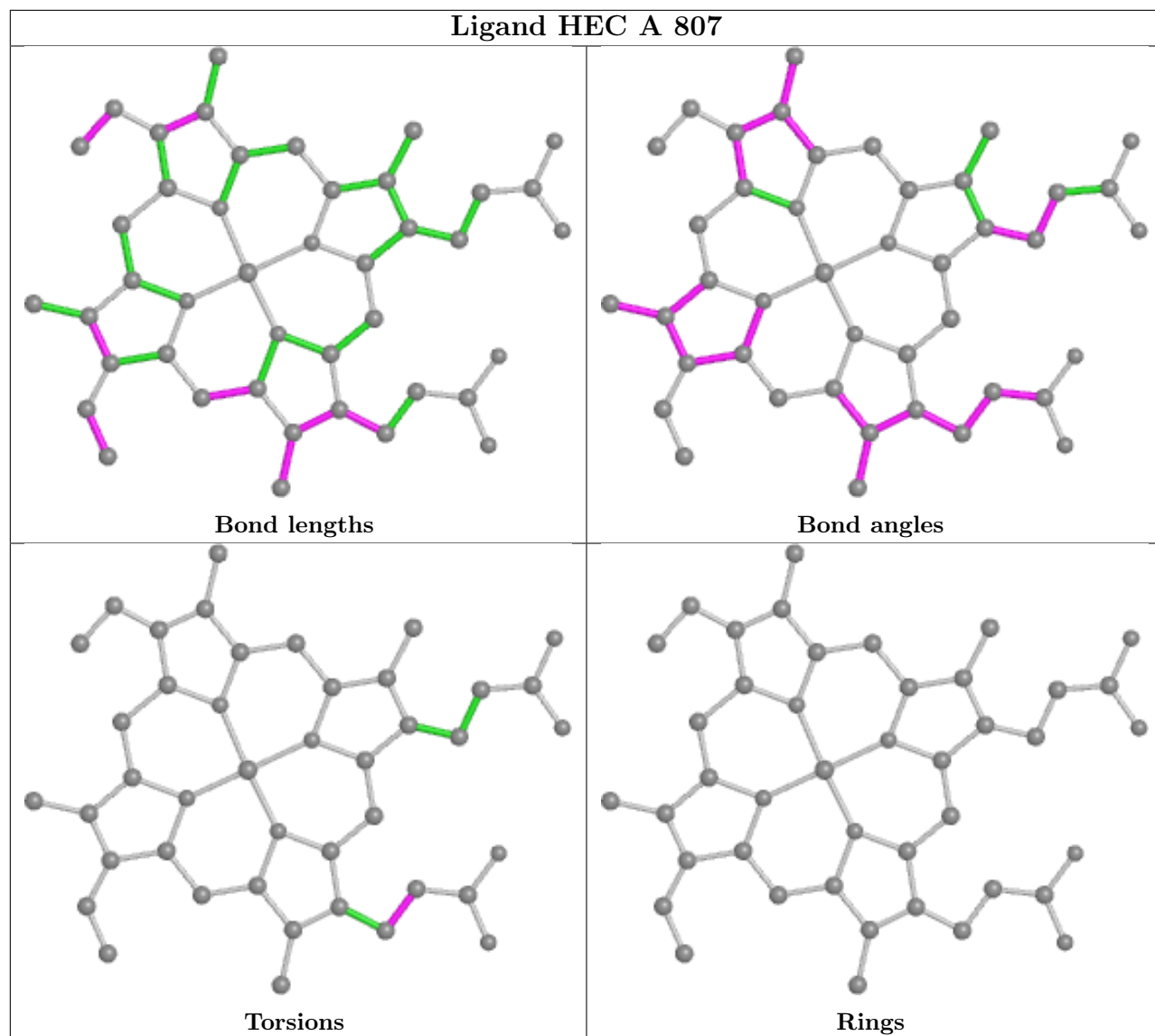


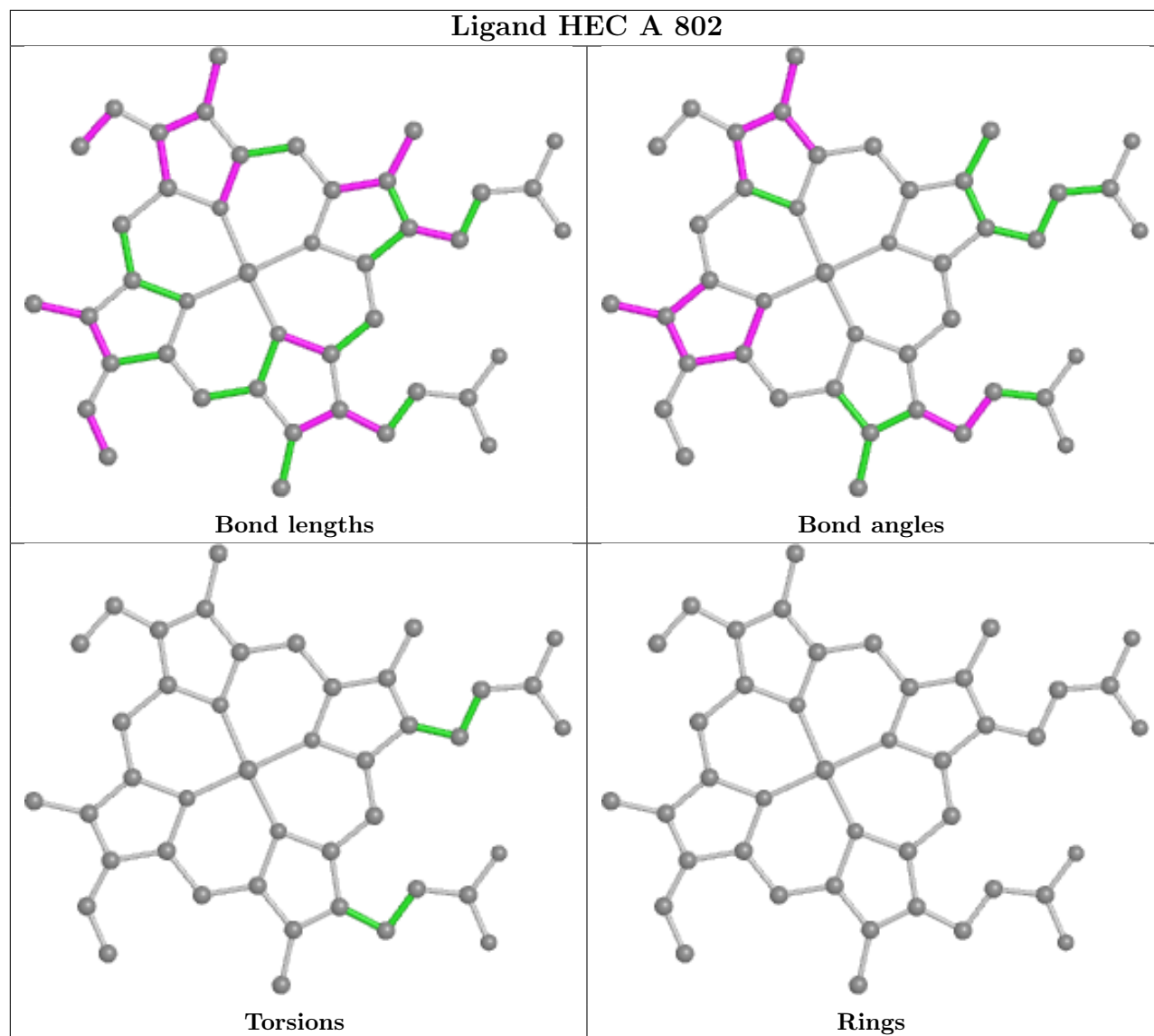












## 5.7 Other polymers [\(i\)](#)

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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ > 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q < 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	436/443 (98%)	-0.17	11 (2%) 57 55	23, 35, 55, 79	0

All (11) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	202	GLY	10.9
1	A	439	PRO	4.2
1	A	437	GLN	4.1
1	A	146	GLY	3.8
1	A	99	LEU	3.6
1	A	201	PRO	3.5
1	A	120	ALA	2.7
1	A	440	LYS	2.6
1	A	438	GLY	2.5
1	A	378	GLU	2.4
1	A	98	CYS	2.2

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

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

### 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

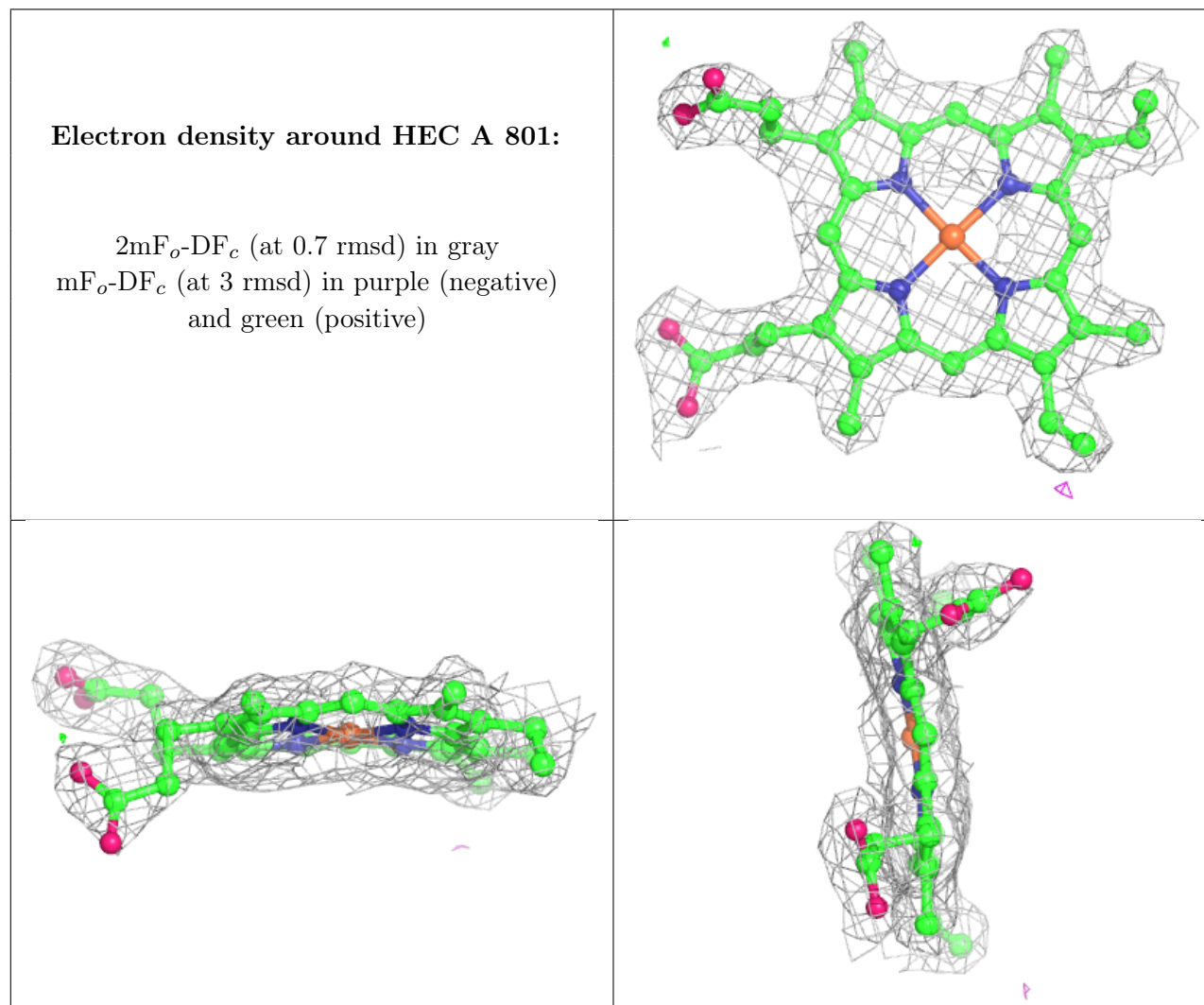
### 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum,

median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

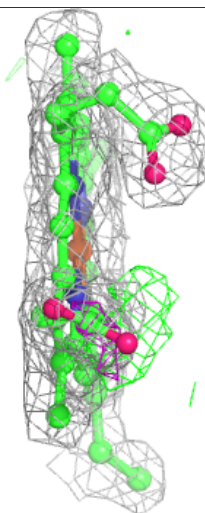
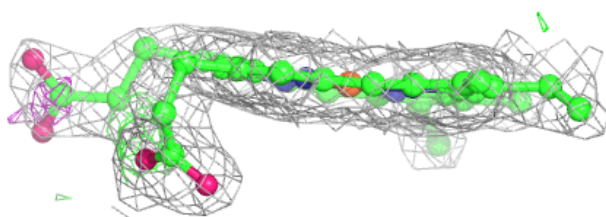
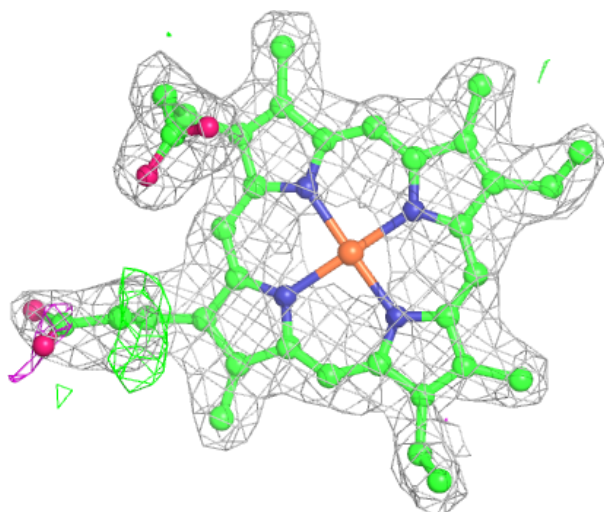
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
3	HEC	A	801	43/43	0.94	0.13	24,32,46,59	0
3	HEC	A	802	43/43	0.94	0.12	24,32,47,61	0
3	HEC	A	803	43/43	0.94	0.15	22,29,45,49	0
3	HEC	A	806	43/43	0.94	0.14	17,29,45,56	0
3	HEC	A	807	43/43	0.94	0.13	17,27,43,45	0
2	SCN	A	810	3/3	0.95	0.15	39,39,41,42	0
3	HEC	A	808	43/43	0.96	0.12	21,28,44,53	0
3	HEC	A	805	43/43	0.98	0.11	17,25,35,40	0
3	HEC	A	804	43/43	0.98	0.13	20,24,33,34	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.



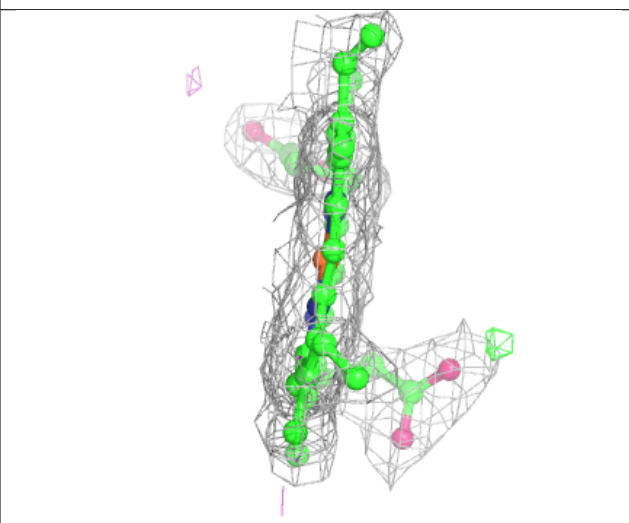
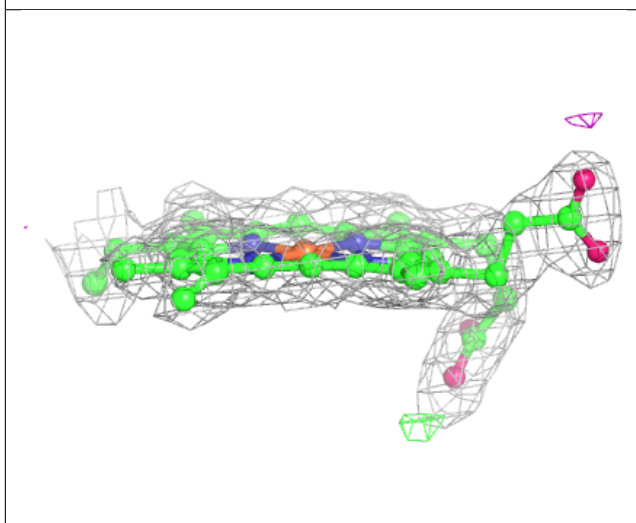
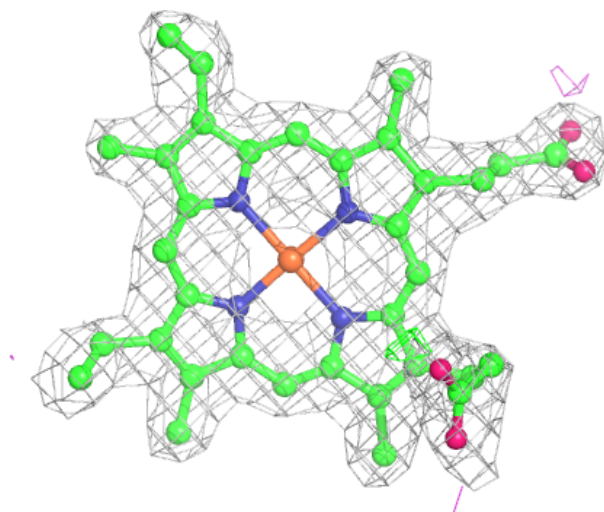
**Electron density around HEC A 802:**

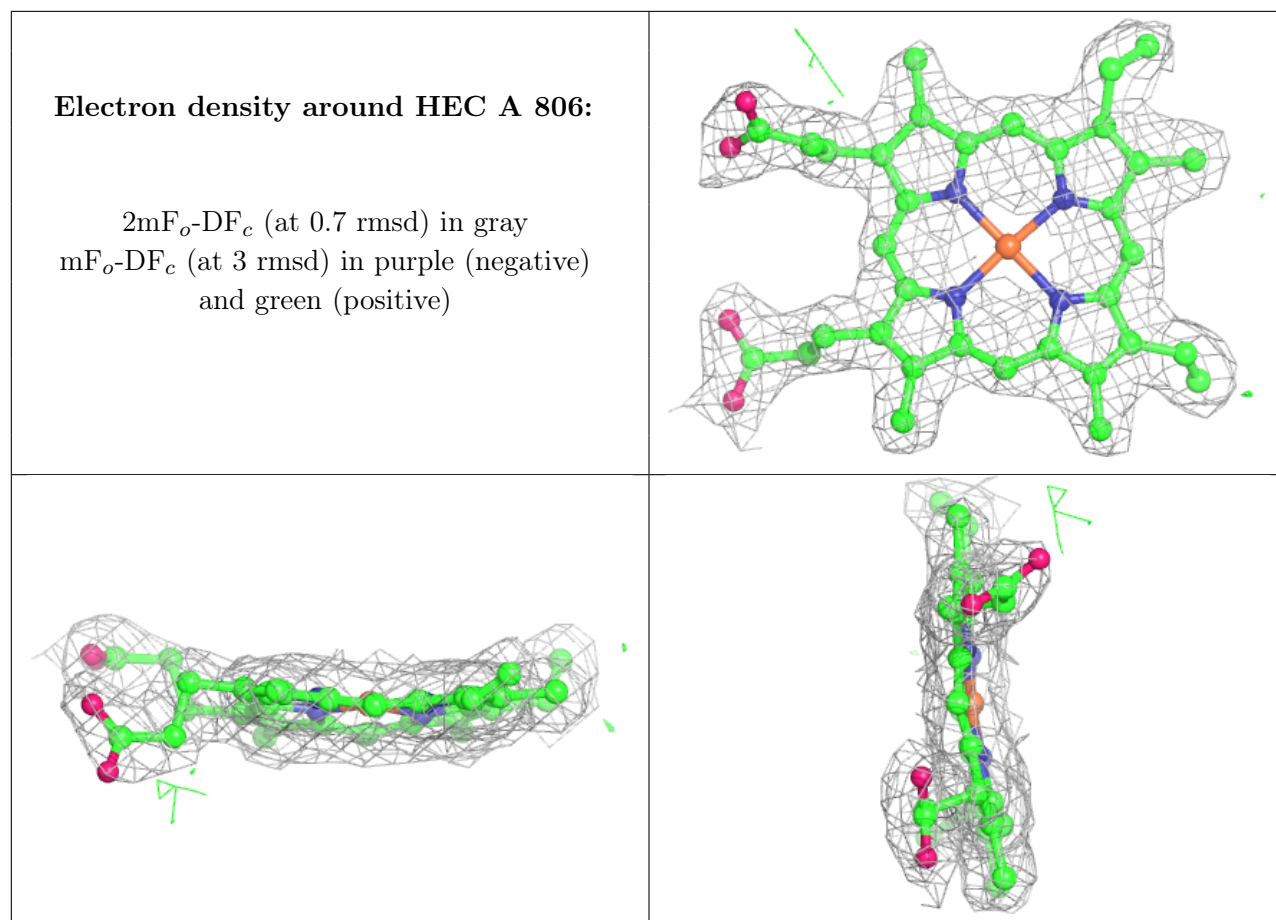
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



**Electron density around HEC A 803:**

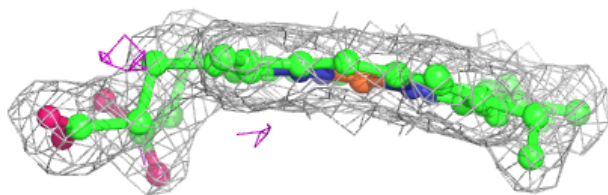
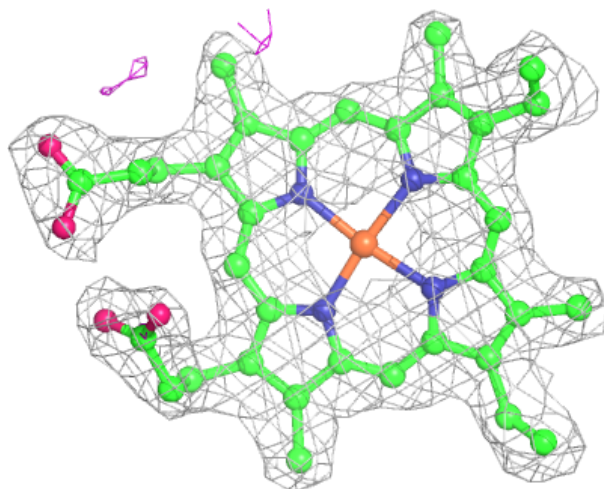
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)





**Electron density around HEC A 807:**

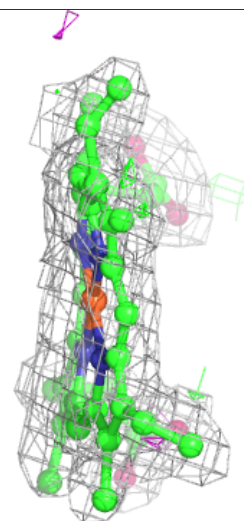
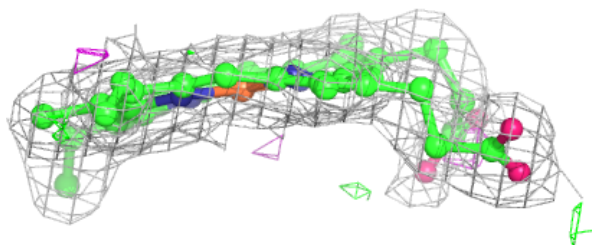
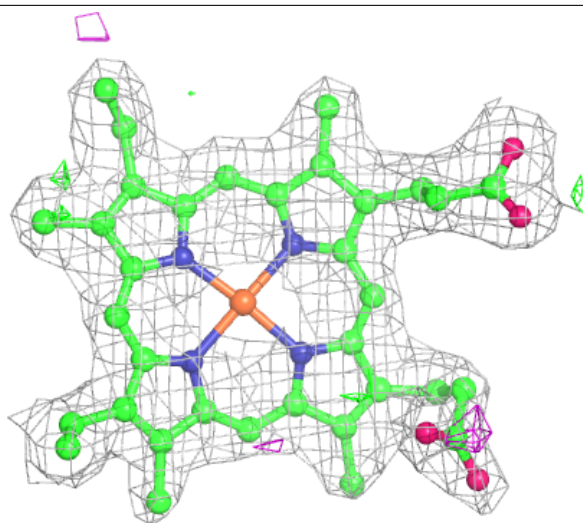
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)





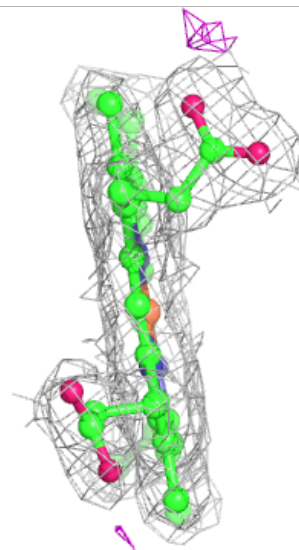
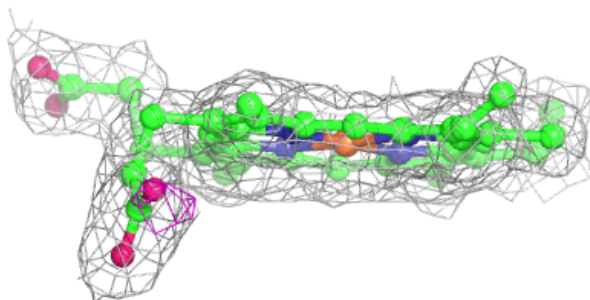
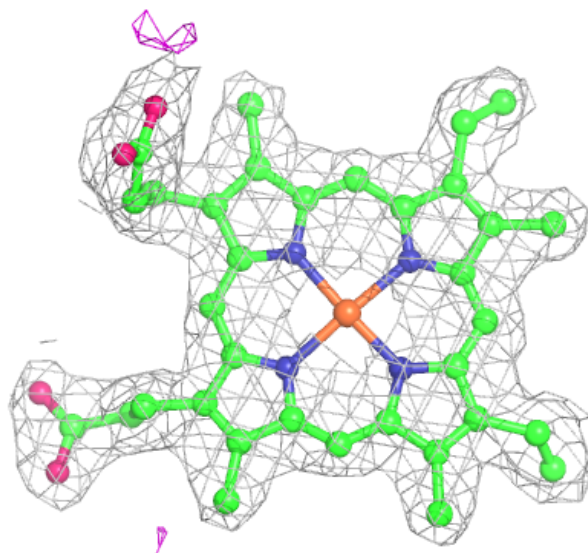
**Electron density around HEC A 808:**

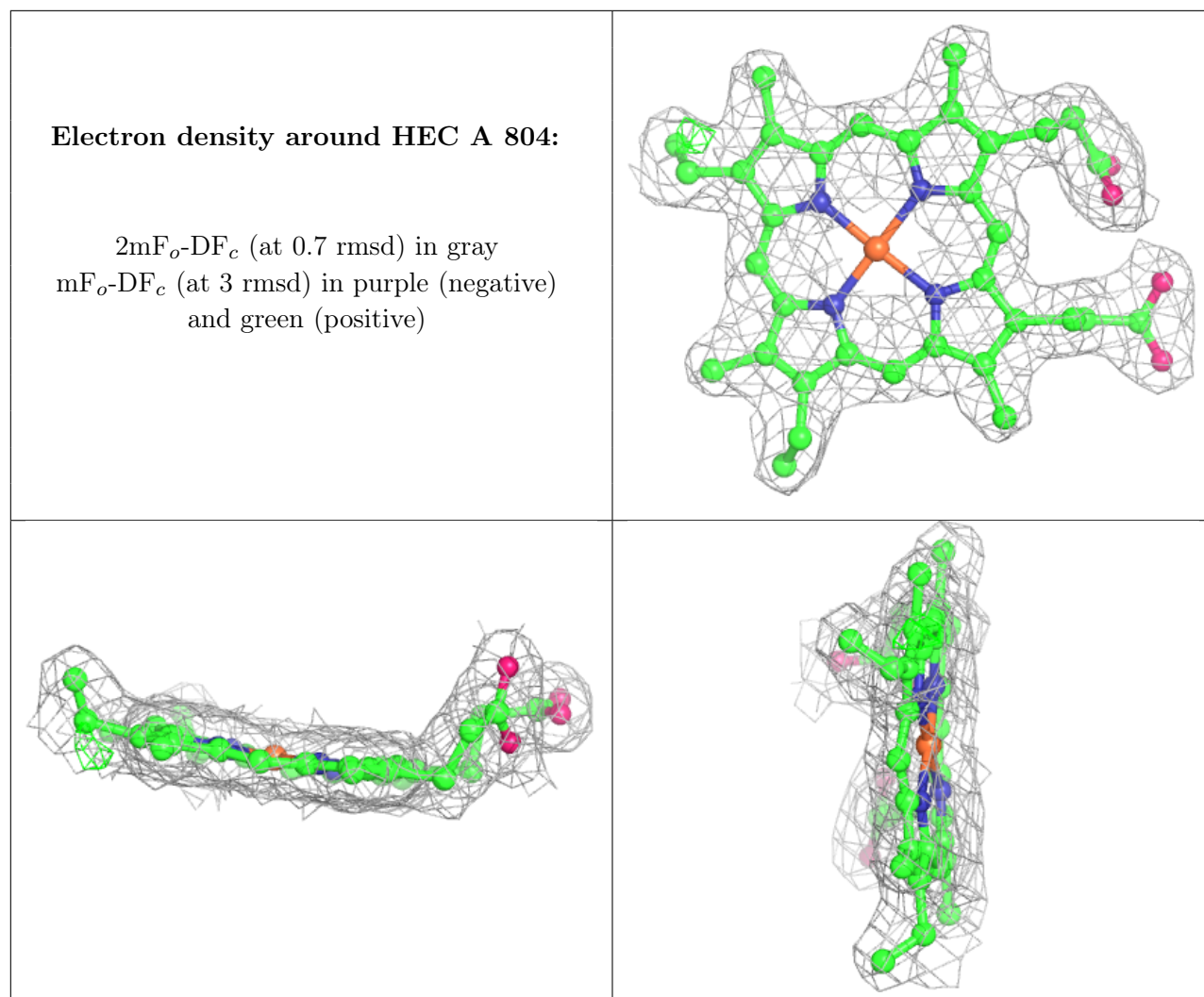
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



**Electron density around HEC A 805:**

$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)





## 6.5 Other polymers [i](#)

There are no such residues in this entry.