



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 17, 2020 – 02:03 am BST

PDB ID : 4D6U  
Title : Cytochrome bc1 bound to the 4(1H)-pyridone GSK932121  
Authors : Capper, M.J.; O'Neill, P.M.; Fisher, N.; Strange, R.W.; Moss, D.; Ward, S.A.;  
Berry, N.G.; Lawrenson, A.S.; Hasnain, S.S.; Biagini, G.A.; Antonyuk, S.V.  
Deposited on : 2014-11-14  
Resolution : 4.09 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](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 : 2.11  
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) : 2.11

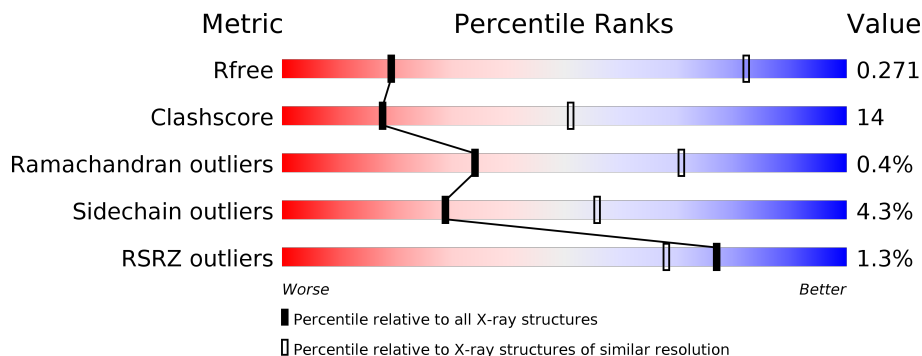
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 4.09 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	1193 (4.50-3.70)
Clashscore	141614	1003 (4.44-3.76)
Ramachandran outliers	138981	1005 (4.48-3.72)
Sidechain outliers	138945	1199 (4.50-3.70)
RSRZ outliers	127900	1034 (4.50-3.70)

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  $\geq 3$ , 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	480	
2	B	453	
3	C	379	
3	P	379	
4	D	325	
4	Q	325	

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	E	274	
5	I	274	
5	R	274	
6	F	111	
6	S	111	
7	G	82	
7	T	82	
8	H	91	
8	U	91	
9	J	64	
9	W	64	
10	N	480	
11	O	453	
12	V	274	

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
14	G8U	C	503	-	-	X	X
14	G8U	P	503	-	-	X	X
15	PO4	D	503	-	-	-	X
15	PO4	D	504	-	-	-	X
15	PO4	F	501	-	-	-	X
15	PO4	N	501	-	-	-	X
15	PO4	S	501	-	-	-	X
16	PEE	C	505	X	-	-	-
16	PEE	D	506	X	-	-	-
16	PEE	P	505	X	-	-	-
16	PEE	Q	506	X	-	-	X
19	FES	R	501	-	-	X	-
20	GOL	R	502	-	-	-	X

## 2 Entry composition [i](#)

There are 20 unique types of molecules in this entry. The entry contains 31080 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 B-C1 COMPLEX SUBUNIT 1, MITOCHONDRIAL.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	444	3439	2148	607	664	20	0	0	0

- Molecule 2 is a protein called CYTOCHROME B-C1 COMPLEX SUBUNIT 2, MITOCHONDRIAL.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	422	3164	1988	561	608	7	0	0	0

- Molecule 3 is a protein called CYTOCHROME B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	374	2968	1993	463	494	18	0	0	0
3	P	370	2936	1973	456	489	18	0	0	0

- Molecule 4 is a protein called CYTOCHROME C1, HEME PROTEIN, MITOCHONDRIAL.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	D	240	1912	1222	329	346	15	0	0	0
4	Q	241	1918	1225	330	348	15	0	0	0

- Molecule 5 is a protein called CYTOCHROME B-C1 COMPLEX SUBUNIT RIESKE, MITOCHONDRIAL.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
5	E	73	549	341	92	114	2	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	I	27	Total	C	N	O	S	0	0	0
			196	121	38	36	1			
5	R	196	Total	C	N	O	S	0	0	0
			1518	957	263	290	8			

- Molecule 6 is a protein called CYTOCHROME B-C1 COMPLEX SUBUNIT 7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	F	98	Total	C	N	O	S	0	0	0
			860	547	154	157	2			
6	S	99	Total	C	N	O	S	0	0	0
			869	553	156	158	2			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	56	ASP	ASN	conflict	UNP P00129
S	56	ASP	ASN	conflict	UNP P00129

- Molecule 7 is a protein called CYTOCHROME B-C1 COMPLEX SUBUNIT 8.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	G	80	Total	C	N	O	S	0	0	0
			677	439	127	110	1			
7	T	74	Total	C	N	O	S	0	0	0
			624	408	117	98	1			

- Molecule 8 is a protein called CYTOCHROME B-C1 COMPLEX SUBUNIT 6, MITOCHONDRIAL.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	H	65	Total	C	N	O	S	0	0	0
			529	321	96	107	5			
8	U	66	Total	C	N	O	S	0	0	0
			538	327	98	108	5			

- Molecule 9 is a protein called CYTOCHROME B-C1 COMPLEX SUBUNIT 9.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
9	J	58	Total	C	N	O	0	0	0
			482	317	83	82			

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
9	W	59	487	320	84	83	0	0	0

- Molecule 10 is a protein called CYTOCHROME B-C1 COMPLEX SUBUNIT 1, MITOCHONDRIAL.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
10	N	444	3430	2142	605	663	20	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
N	445	LYS	ARG	conflict	UNP P31800

- Molecule 11 is a protein called CYTOCHROME B-C1 COMPLEX SUBUNIT 2, MITOCHONDRIAL.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
11	O	419	3140	1972	555	606	7	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
O	305	GLU	GLN	conflict	UNP P23004

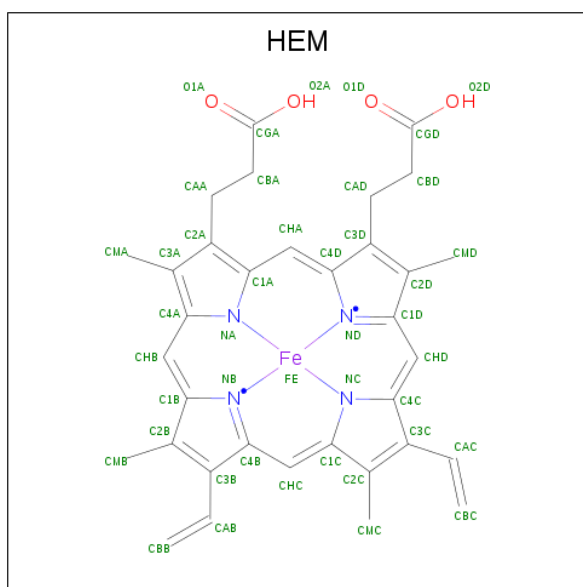
- Molecule 12 is a protein called CYTOCHROME B-C1 COMPLEX SUBUNIT RIESKE, MITOCHONDRIAL.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
12	V	17	127	81	24	22	0	0	0

There is a discrepancy between the modelled and reference sequences:

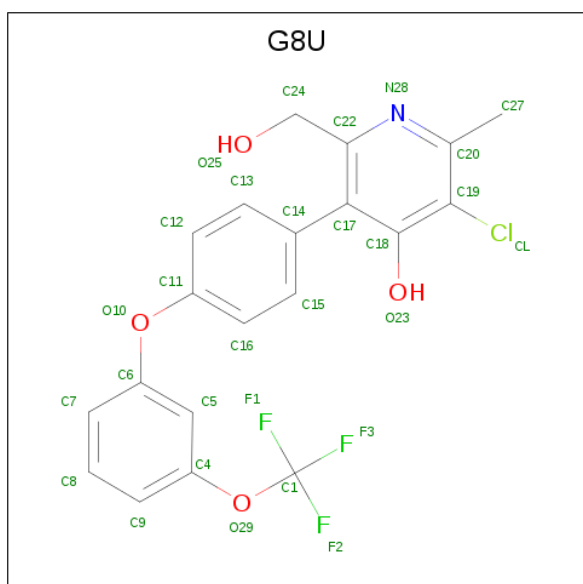
Chain	Residue	Modelled	Actual	Comment	Reference
V	64	VAL	LEU	conflict	UNP P13272

- Molecule 13 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (formula: C<sub>34</sub>H<sub>32</sub>FeN<sub>4</sub>O<sub>4</sub>).



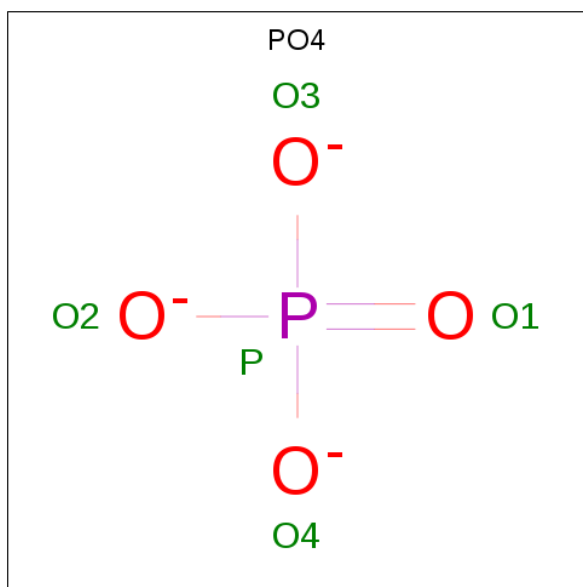
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	
			Total	C	Fe	N			O
13	C	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
13	C	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
13	P	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
13	P	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		

- Molecule 14 is 3-chloro-6-(hydroxymethyl)-2-methyl-5-{4-[3-(trifluoromethoxy)phenoxy]phenyl}pyridin-4-ol (three-letter code: G8U) (formula:  $C_{20}H_{15}ClF_3NO_4$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	
14	C	1	Total	C	Cl	F	N	O	0	0
			29	20	1	3	1	4		
14	P	1	Total	C	Cl	F	N	O	0	0
			29	20	1	3	1	4		

- Molecule 15 is PHOSPHATE ION (three-letter code: PO4) (formula: O<sub>4</sub>P).

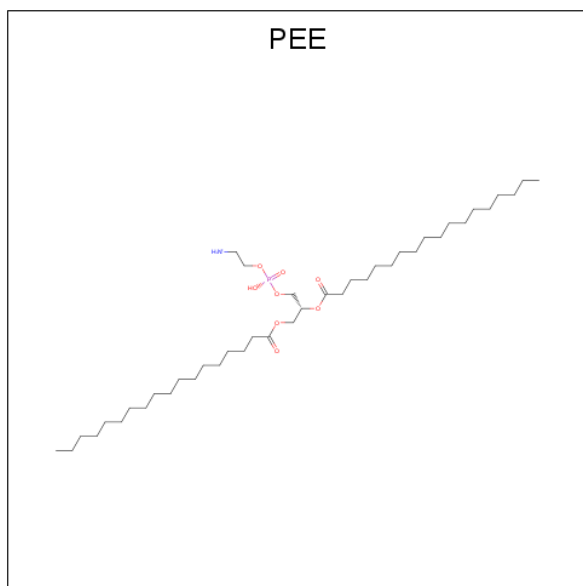


Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
15	C	1	Total	O P	0	0
			5	4 1		
15	D	1	Total	O P	0	0
			5	4 1		
15	D	1	Total	O P	0	0
			5	4 1		
15	D	1	Total	O P	0	0
			5	4 1		
15	E	1	Total	O P	0	0
			5	4 1		
15	F	1	Total	O P	0	0
			5	4 1		
15	N	1	Total	O P	0	0
			5	4 1		
15	P	1	Total	O P	0	0
			5	4 1		
15	S	1	Total	O P	0	0
			5	4 1		

- Molecule 16 is 1,2-Dioleoyl-sn-glycero-3-phosphoethanolamine (three-letter code: PEE)

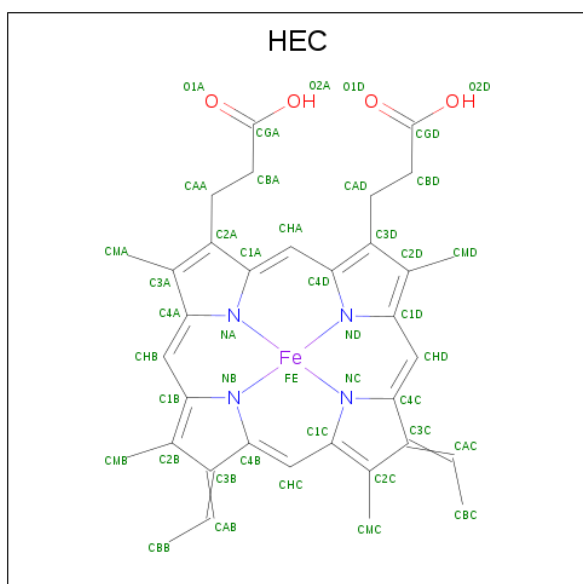


(formula: C<sub>41</sub>H<sub>83</sub>NO<sub>8</sub>P).



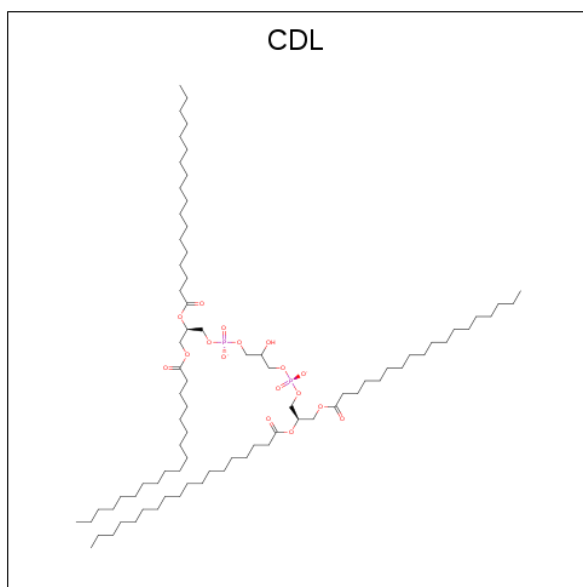
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
16	C	1	Total	C	N	O	P	0	0
			49	39	1	8	1		
16	D	1	Total	C	N	O	P	0	0
			26	16	1	8	1		
16	P	1	Total	C	N	O	P	0	0
			49	39	1	8	1		
16	Q	1	Total	C	N	O	P	0	0
			51	41	1	8	1		

- Molecule 17 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	
			Total	C	Fe	N			O
17	D	1	43	34	1	4	4	0	0
17	Q	1	43	34	1	4	4	0	0

- Molecule 18 is CARDIOLIPIN (three-letter code: CDL) (formula:  $C_{81}H_{156}O_{17}P_2$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	O	P		
18	D	1	39	24	13	2	0	0
18	G	1	44	25	17	2	0	0
18	Q	1	39	24	13	2	0	0
18	T	1	49	30	17	2	0	0

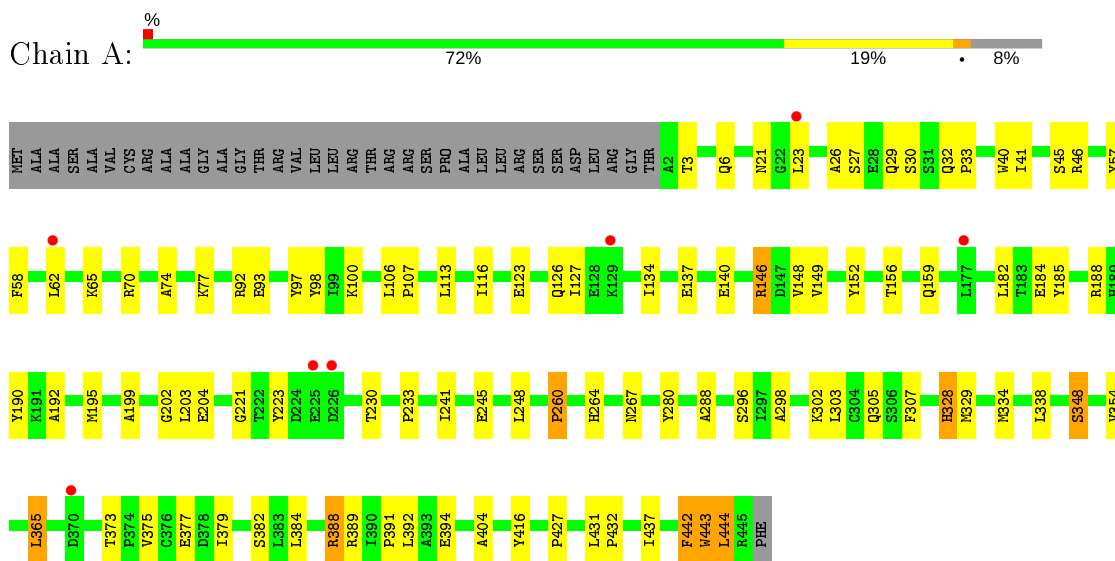
- Molecule 19 is FE2/S2 (INORGANIC) CLUSTER (three-letter code: FES) (formula:  $Fe_2S_2$ ).



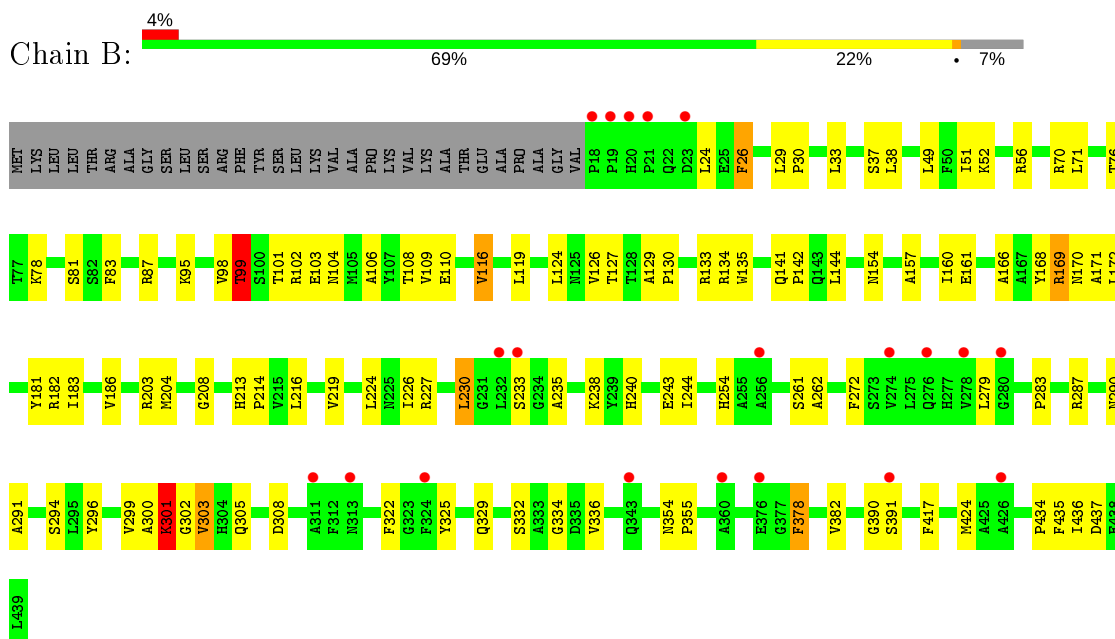
### 3 Residue-property plots [i](#)

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 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 B-C1 COMPLEX SUBUNIT 1, MITOCHONDRIAL

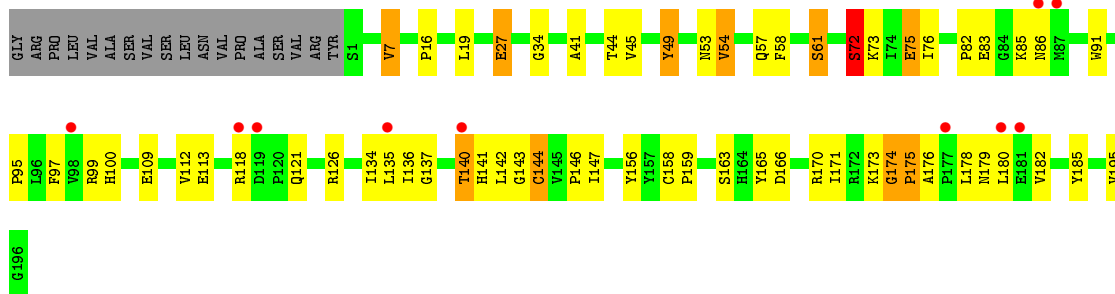


- Molecule 2: CYTOCHROME B-C1 COMPLEX SUBUNIT 2, MITOCHONDRIAL

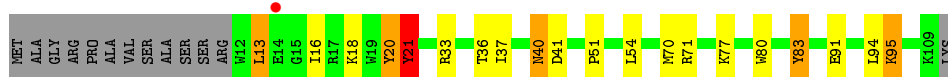








● Molecule 6: CYTOCHROME B-C1 COMPLEX SUBUNIT 7



● Molecule 6: CYTOCHROME B-C1 COMPLEX SUBUNIT 7



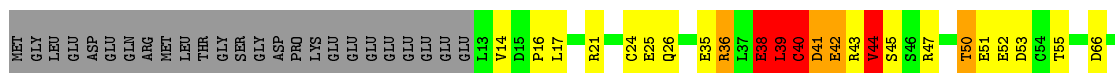
● Molecule 7: CYTOCHROME B-C1 COMPLEX SUBUNIT 8



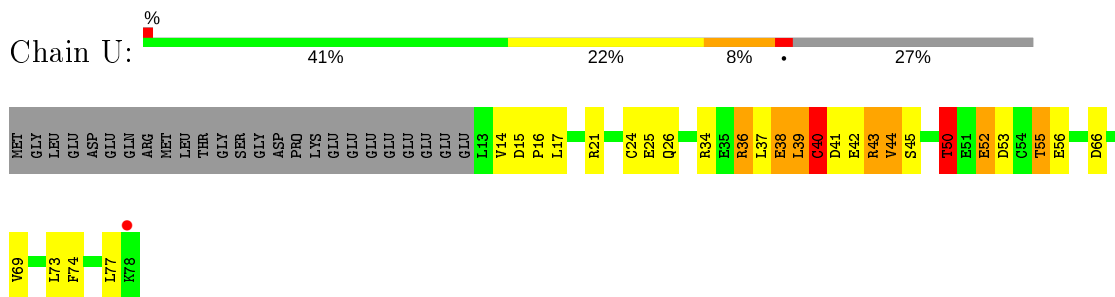
● Molecule 7: CYTOCHROME B-C1 COMPLEX SUBUNIT 8



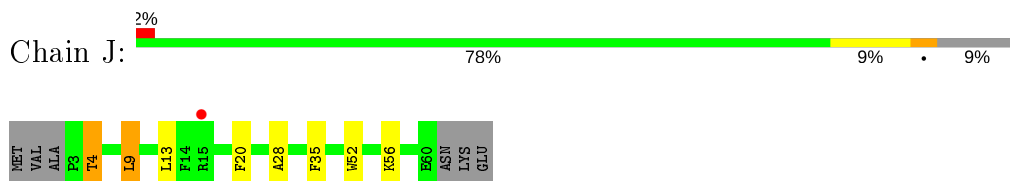
● Molecule 8: CYTOCHROME B-C1 COMPLEX SUBUNIT 6, MITOCHONDRIAL



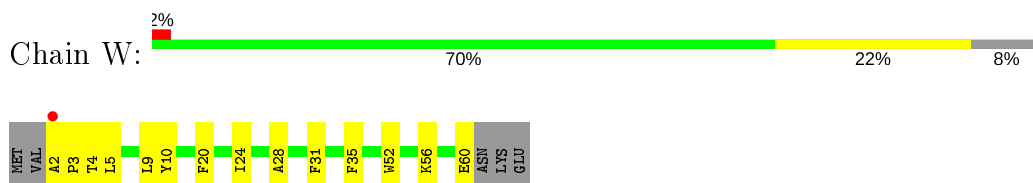
● Molecule 8: CYTOCHROME B-C1 COMPLEX SUBUNIT 6, MITOCHONDRIAL



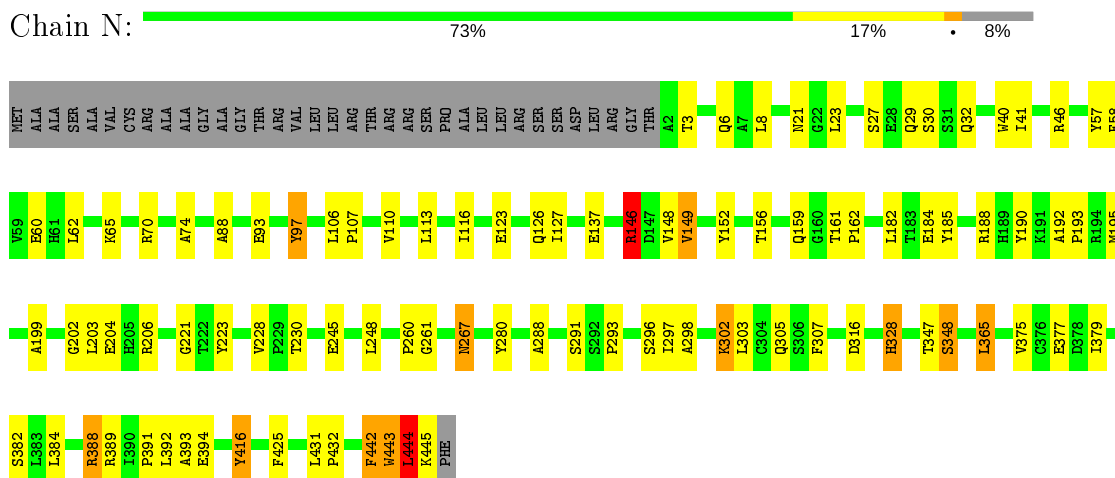
- Molecule 9: CYTOCHROME B-C1 COMPLEX SUBUNIT 9



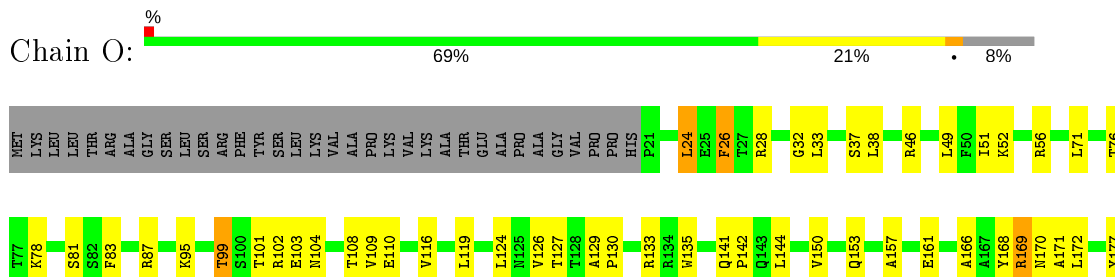
- Molecule 9: CYTOCHROME B-C1 COMPLEX SUBUNIT 9



- Molecule 10: CYTOCHROME B-C1 COMPLEX SUBUNIT 1, MITOCHONDRIAL



- Molecule 11: CYTOCHROME B-C1 COMPLEX SUBUNIT 2, MITOCHONDRIAL







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 65	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	129.49Å 129.49Å 719.94Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	50.00 – 4.09 49.65 – 4.09	Depositor EDS
% Data completeness (in resolution range)	80.6 (50.00-4.09) 80.7 (49.65-4.09)	Depositor EDS
$R_{merge}$	0.15	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.57 (at 4.14Å)	Xtriage
Refinement program	REFMAC 5.8.0073	Depositor
R, $R_{free}$	0.223 , 0.271 0.222 , 0.271	Depositor DCC
$R_{free}$ test set	2147 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	123.3	Xtriage
Anisotropy	0.395	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 135.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	0.086 for h,-h-k,-l	Xtriage
$F_o, F_c$ correlation	0.89	EDS
Total number of atoms	31080	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	163.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.63% 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: GOL, CDL, PO4, G8U, FES, HEC, PEE, HEM

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	0.53	6/3511 (0.2%)	0.74	9/4766 (0.2%)
2	B	0.75	18/3224 (0.6%)	0.78	10/4375 (0.2%)
3	C	0.59	7/3065 (0.2%)	0.64	2/4196 (0.0%)
3	P	0.57	5/3031 (0.2%)	0.64	3/4150 (0.1%)
4	D	0.68	6/1971 (0.3%)	0.62	1/2676 (0.0%)
4	Q	0.62	5/1977 (0.3%)	0.62	3/2684 (0.1%)
5	E	1.41	7/557 (1.3%)	0.70	1/752 (0.1%)
5	I	0.64	0/196	0.96	1/263 (0.4%)
5	R	0.72	8/1552 (0.5%)	0.74	5/2100 (0.2%)
6	F	0.95	8/879 (0.9%)	0.69	1/1180 (0.1%)
6	S	0.91	8/888 (0.9%)	0.73	2/1191 (0.2%)
7	G	0.84	6/699 (0.9%)	1.26	4/946 (0.4%)
7	T	0.78	5/645 (0.8%)	0.70	0/873
8	H	1.04	4/534 (0.7%)	1.20	6/718 (0.8%)
8	U	1.00	5/543 (0.9%)	1.27	10/729 (1.4%)
9	J	0.55	1/495 (0.2%)	0.55	0/667
9	W	0.51	0/500	0.59	0/675
10	N	0.62	8/3501 (0.2%)	0.72	8/4752 (0.2%)
11	O	0.69	15/3197 (0.5%)	0.69	5/4336 (0.1%)
12	V	0.70	0/129	1.04	2/177 (1.1%)
All	All	0.70	122/31094 (0.4%)	0.74	73/42206 (0.2%)

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	1
2	B	1	1
5	R	0	1

*Continued on next page...*

Continued from previous page...

Mol	Chain	#Chirality outliers	#Planarity outliers
8	H	0	1
10	N	0	1
12	V	0	1
All	All	1	6

The worst 5 of 122 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
10	N	444	LEU	C-N	-20.33	0.87	1.34
5	E	27	GLU	CD-OE1	20.21	1.47	1.25
5	E	27	GLU	CD-OE2	15.40	1.42	1.25
2	B	169	ARG	CZ-NH1	15.31	1.52	1.33
4	D	93	LYS	CE-NZ	14.76	1.85	1.49

The worst 5 of 73 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	G	75	ALA	CB-CA-C	-21.60	77.69	110.10
7	G	76	ALA	N-CA-CB	-17.49	85.61	110.10
1	A	146	ARG	NE-CZ-NH1	17.36	128.98	120.30
2	B	300	ALA	CB-CA-C	-15.03	87.56	110.10
2	B	224	LEU	CB-CG-CD1	-13.39	88.23	111.00

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	B	301	LYS	CA

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	146	ARG	Sidechain
2	B	301	LYS	Peptide
8	H	42	GLU	Sidechain
10	N	444	LEU	Mainchain
5	R	27	GLU	Sidechain

## 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	3439	0	3337	78	0
2	B	3164	0	3144	83	0
3	C	2968	0	3028	122	0
3	P	2936	0	2996	103	0
4	D	1912	0	1862	45	0
4	Q	1918	0	1870	51	0
5	E	549	0	547	18	0
5	I	196	0	204	32	0
5	R	1518	0	1504	57	0
6	F	860	0	849	24	0
6	S	869	0	862	21	0
7	G	677	0	673	55	0
7	T	624	0	630	18	0
8	H	529	0	511	34	0
8	U	538	0	524	29	0
9	J	482	0	483	5	0
9	W	487	0	487	12	0
10	N	3430	0	3329	76	0
11	O	3140	0	3121	72	0
12	V	127	0	135	9	0
13	C	86	0	60	8	0
13	P	86	0	60	12	0
14	C	29	0	13	13	0
14	P	29	0	13	9	0
15	C	5	0	0	0	0
15	D	15	0	0	0	0
15	E	5	0	0	0	0
15	F	5	0	0	0	0
15	N	5	0	0	0	0
15	P	5	0	0	0	0
15	S	5	0	0	0	0
16	C	49	0	72	0	0
16	D	26	0	26	1	0
16	P	49	0	72	4	0
16	Q	51	0	82	3	0
17	D	43	0	32	8	0
17	Q	43	0	32	8	0
18	D	39	0	39	0	0
18	G	44	0	32	0	0
18	Q	39	0	39	0	0
18	T	49	0	42	5	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
19	R	4	0	0	3	0
20	R	6	0	8	0	0
All	All	31080	0	30718	841	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 14.

The worst 5 of 841 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)
8:U:39:LEU:CG	8:U:39:LEU:CD2	1.74	1.56
2:B:301:LYS:NZ	2:B:301:LYS:CE	1.69	1.51
3:C:194:MET:SD	3:C:194:MET:CE	2.06	1.43
3:P:194:MET:SD	3:P:194:MET:CE	2.07	1.42
4:D:93:LYS:NZ	4:D:93:LYS:CE	1.85	1.39

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	442/480 (92%)	401 (91%)	38 (9%)	3 (1%)	22	60
2	B	420/453 (93%)	377 (90%)	43 (10%)	0	100	100
3	C	372/379 (98%)	334 (90%)	37 (10%)	1 (0%)	41	75
3	P	368/379 (97%)	331 (90%)	36 (10%)	1 (0%)	41	75
4	D	238/325 (73%)	219 (92%)	18 (8%)	1 (0%)	34	71
4	Q	239/325 (74%)	221 (92%)	16 (7%)	2 (1%)	19	58
5	E	71/274 (26%)	62 (87%)	8 (11%)	1 (1%)	11	45
5	I	23/274 (8%)	20 (87%)	3 (13%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
5	R	194/274 (71%)	168 (87%)	25 (13%)	1 (0%)	29	67
6	F	96/111 (86%)	88 (92%)	8 (8%)	0	100	100
6	S	97/111 (87%)	90 (93%)	7 (7%)	0	100	100
7	G	78/82 (95%)	71 (91%)	7 (9%)	0	100	100
7	T	72/82 (88%)	66 (92%)	6 (8%)	0	100	100
8	H	63/91 (69%)	53 (84%)	9 (14%)	1 (2%)	9	43
8	U	64/91 (70%)	54 (84%)	9 (14%)	1 (2%)	9	43
9	J	56/64 (88%)	47 (84%)	9 (16%)	0	100	100
9	W	57/64 (89%)	49 (86%)	8 (14%)	0	100	100
10	N	442/480 (92%)	402 (91%)	38 (9%)	2 (0%)	29	67
11	O	417/453 (92%)	378 (91%)	38 (9%)	1 (0%)	47	80
12	V	15/274 (6%)	12 (80%)	3 (20%)	0	100	100
All	All	3824/5066 (76%)	3443 (90%)	366 (10%)	15 (0%)	34	71

5 of 15 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	443	TRP
8	H	25	GLU
8	U	25	GLU
5	E	61	SER
1	A	267	ASN

### 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	368/394 (93%)	351 (95%)	17 (5%)	27	54
2	B	331/355 (93%)	324 (98%)	7 (2%)	53	72
3	C	322/327 (98%)	310 (96%)	12 (4%)	34	59
3	P	318/327 (97%)	305 (96%)	13 (4%)	30	57

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
4	D	205/257 (80%)	200 (98%)	5 (2%)	49	69
4	Q	206/257 (80%)	202 (98%)	4 (2%)	57	75
5	E	63/228 (28%)	59 (94%)	4 (6%)	18	46
5	I	21/228 (9%)	20 (95%)	1 (5%)	25	53
5	R	168/228 (74%)	161 (96%)	7 (4%)	30	56
6	F	90/99 (91%)	83 (92%)	7 (8%)	12	39
6	S	91/99 (92%)	84 (92%)	7 (8%)	13	40
7	G	71/72 (99%)	62 (87%)	9 (13%)	4	22
7	T	66/72 (92%)	59 (89%)	7 (11%)	6	27
8	H	62/85 (73%)	55 (89%)	7 (11%)	6	25
8	U	63/85 (74%)	58 (92%)	5 (8%)	12	39
9	J	49/54 (91%)	47 (96%)	2 (4%)	30	57
9	W	49/54 (91%)	48 (98%)	1 (2%)	55	73
10	N	367/394 (93%)	352 (96%)	15 (4%)	30	57
11	O	328/355 (92%)	318 (97%)	10 (3%)	41	64
12	V	15/228 (7%)	14 (93%)	1 (7%)	16	44
All	All	3253/4198 (78%)	3112 (96%)	141 (4%)	29	56

5 of 141 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
8	H	38	GLU
10	N	328	HIS
7	T	18	LEU
8	H	40	CYS
10	N	146	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 38 such sidechains are listed below:

Mol	Chain	Res	Type
4	D	225	HIS
10	N	61	HIS
5	R	57	GLN
6	F	79	GLN
10	N	118	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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

27 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
15	PO4	D	502	-	4,4,4	0.83	0	6,6,6	0.44	0
15	PO4	F	501	-	4,4,4	0.92	0	6,6,6	0.51	0
17	HEC	Q	501	4	26,50,50	2.51	12 (46%)	18,82,82	2.68	7 (38%)
15	PO4	D	503	-	4,4,4	0.93	0	6,6,6	0.46	0
15	PO4	S	501	-	4,4,4	0.88	0	6,6,6	0.36	0
18	CDL	Q	505	-	38,38,99	1.19	2 (5%)	43,47,111	1.14	3 (6%)
16	PEE	C	505	-	48,48,50	0.99	2 (4%)	51,53,55	0.84	2 (3%)
17	HEC	D	501	4	26,50,50	2.46	12 (46%)	18,82,82	2.59	5 (27%)
15	PO4	E	501	-	4,4,4	0.81	0	6,6,6	0.70	0
19	FES	R	501	5	0,4,4	0.00	-	-	-	-
13	HEM	P	501	3	27,50,50	0.90	2 (7%)	17,82,82	1.08	1 (5%)
20	GOL	R	502	-	5,5,5	0.30	0	5,5,5	0.30	0
16	PEE	P	505	-	48,48,50	0.98	2 (4%)	51,53,55	0.88	2 (3%)
15	PO4	C	504	-	4,4,4	1.01	0	6,6,6	0.21	0
13	HEM	C	502	3	27,50,50	0.96	0	17,82,82	1.56	3 (17%)
14	G8U	C	503	3	31,31,31	2.25	6 (19%)	41,45,45	1.66	8 (19%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
16	PEE	Q	506	-	50,50,50	1.02	2 (4%)	53,55,55	0.89	3 (5%)
13	HEM	C	501	3	27,50,50	0.99	2 (7%)	17,82,82	1.39	1 (5%)
14	G8U	P	503	3	31,31,31	2.16	6 (19%)	41,45,45	1.80	6 (14%)
13	HEM	P	502	3	27,50,50	0.89	2 (7%)	17,82,82	1.54	2 (11%)
15	PO4	D	504	-	4,4,4	0.92	0	6,6,6	0.41	0
18	CDL	T	501	-	48,48,99	1.32	4 (8%)	54,60,111	1.25	4 (7%)
18	CDL	G	501	-	43,43,99	1.50	4 (9%)	49,55,111	1.44	6 (12%)
15	PO4	P	504	-	4,4,4	0.86	0	6,6,6	0.68	0
18	CDL	D	505	-	38,38,99	1.20	2 (5%)	43,47,111	1.08	4 (9%)
16	PEE	D	506	-	25,25,50	1.42	2 (8%)	28,30,55	1.22	3 (10%)
15	PO4	N	501	-	4,4,4	0.93	0	6,6,6	0.39	0

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
17	HEC	D	501	4	-	0/6/54/54	-
14	G8U	C	503	3	-	5/15/15/15	0/3/3/3
17	HEC	Q	501	4	-	1/6/54/54	-
18	CDL	T	501	-	-	29/57/57/110	-
16	PEE	Q	506	-	1/1/4/4	28/54/54/54	-
13	HEM	P	501	3	-	4/6/54/54	-
18	CDL	G	501	-	-	32/52/52/110	-
20	GOL	R	502	-	-	3/4/4/4	-
16	PEE	P	505	-	1/1/4/4	31/52/52/54	-
13	HEM	C	501	3	-	1/6/54/54	-
13	HEM	C	502	3	-	1/6/54/54	-
13	HEM	P	502	3	-	2/6/54/54	-
14	G8U	P	503	3	-	4/15/15/15	0/3/3/3
16	PEE	D	506	-	1/1/4/4	13/29/29/54	-
19	FES	R	501	5	-	-	0/1/1/1
18	CDL	Q	505	-	-	13/43/43/110	-
18	CDL	D	505	-	-	24/43/43/110	-
16	PEE	C	505	-	1/1/4/4	29/52/52/54	-

The worst 5 of 62 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
14	C	503	G8U	C19-C20	7.20	1.48	1.38
17	Q	501	HEC	C3C-C2C	6.77	1.47	1.40
17	D	501	HEC	C3C-C2C	6.71	1.47	1.40
14	P	503	G8U	C19-C20	6.55	1.47	1.38
17	Q	501	HEC	C3B-C2B	5.84	1.46	1.40

The worst 5 of 60 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
17	Q	501	HEC	C1D-C2D-C3D	-7.34	101.89	107.00
14	P	503	G8U	C27-C20-C19	-6.01	119.56	123.01
17	D	501	HEC	C1D-C2D-C3D	-5.77	102.98	107.00
14	P	503	G8U	C18-C19-C20	-5.71	119.16	122.79
18	G	501	CDL	OA6-CA5-C11	5.69	121.56	111.09

All (4) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
16	C	505	PEE	C2
16	P	505	PEE	C2
16	Q	506	PEE	C2
16	D	506	PEE	C2

5 of 220 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
18	Q	505	CDL	CA2-OA2-PA1-OA3
13	P	501	HEM	C1A-C2A-CAA-CBA
13	P	501	HEM	C3A-C2A-CAA-CBA
13	P	501	HEM	C3D-CAD-CBD-CGD
20	R	502	GOL	C1-C2-C3-O3

There are no ring outliers.

13 monomers are involved in 72 short contacts:

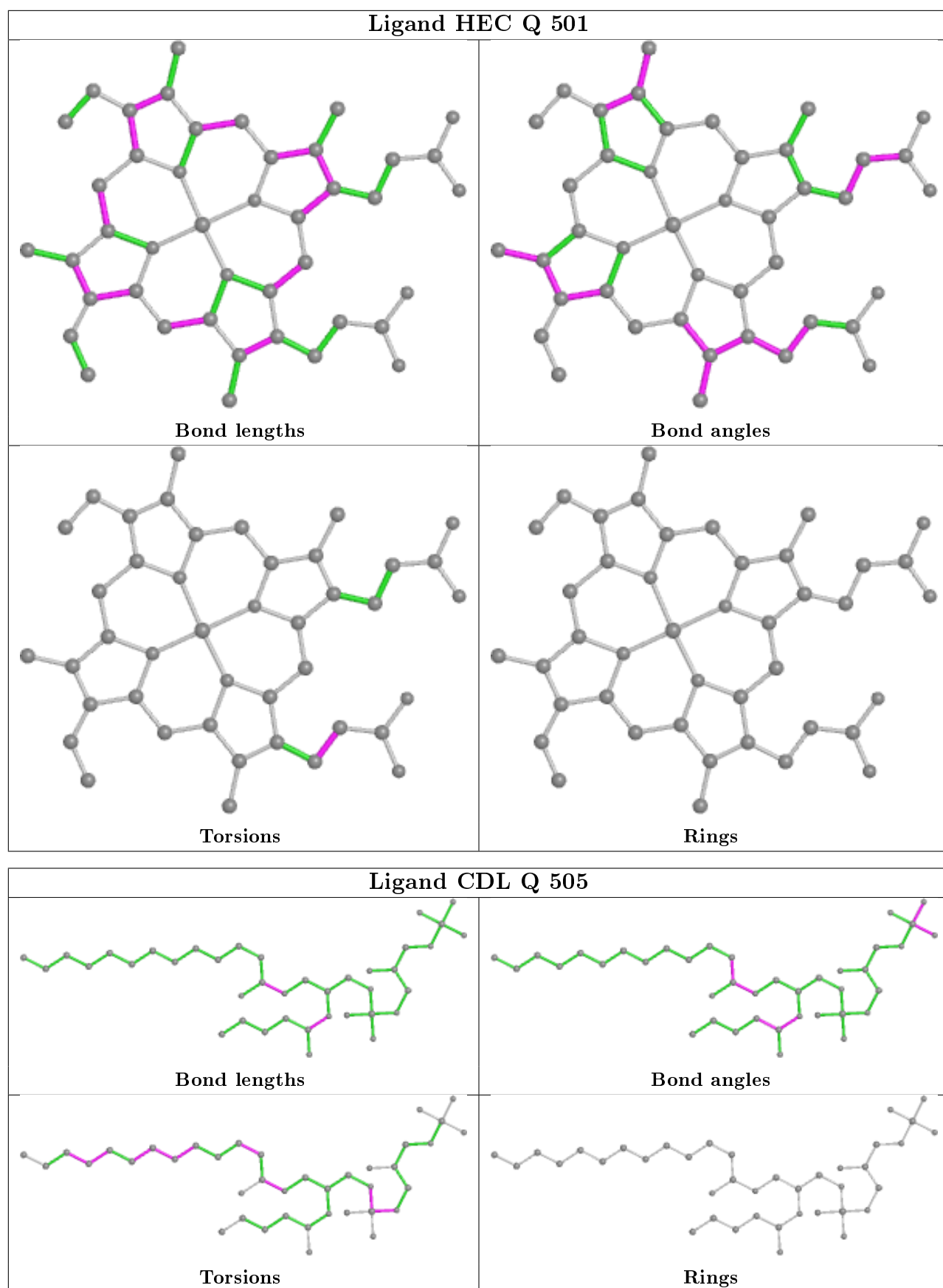
Mol	Chain	Res	Type	Clashes	Symm-Clashes
17	Q	501	HEC	8	0
17	D	501	HEC	8	0
19	R	501	FES	3	0
13	P	501	HEM	10	0
16	P	505	PEE	4	0
13	C	502	HEM	3	0

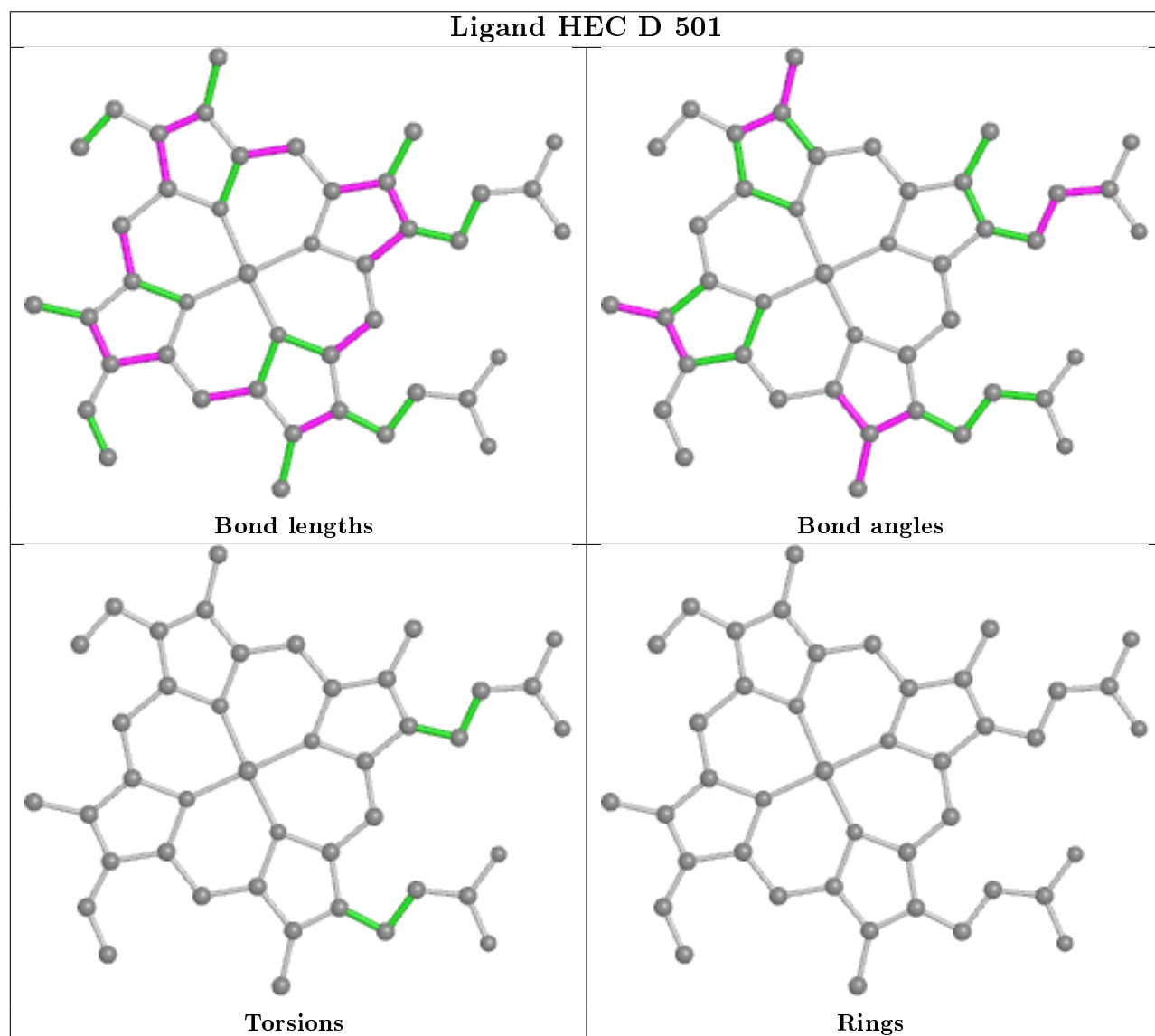
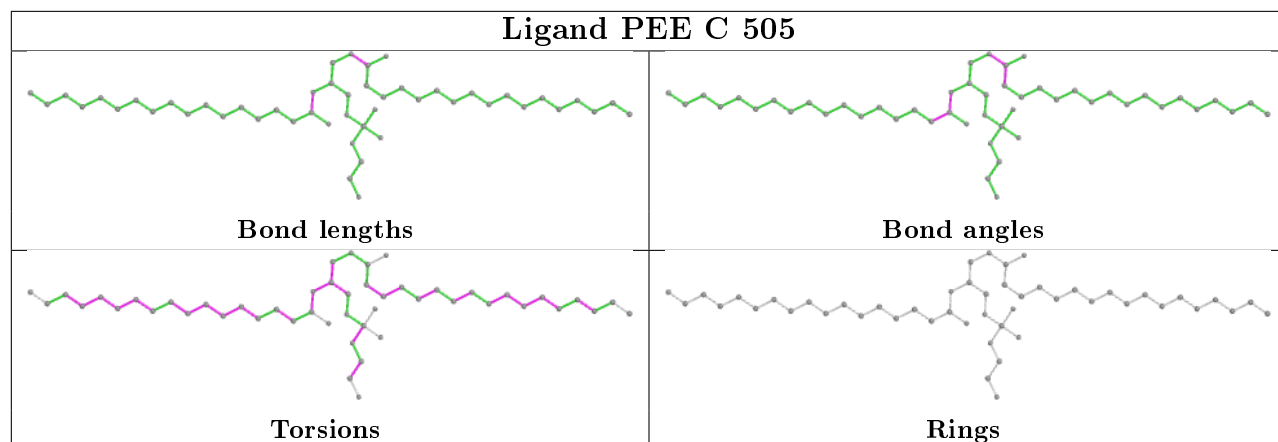
*Continued on next page...*

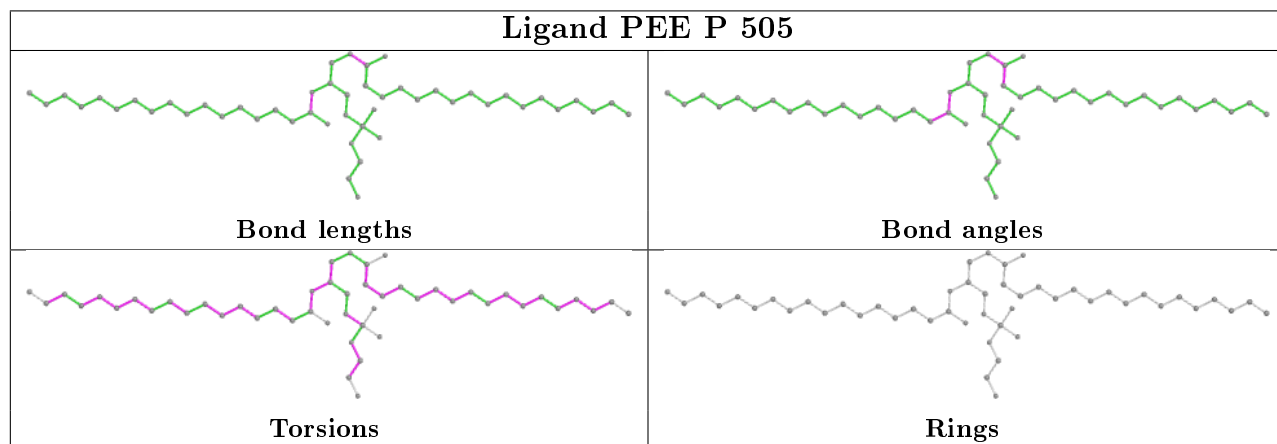
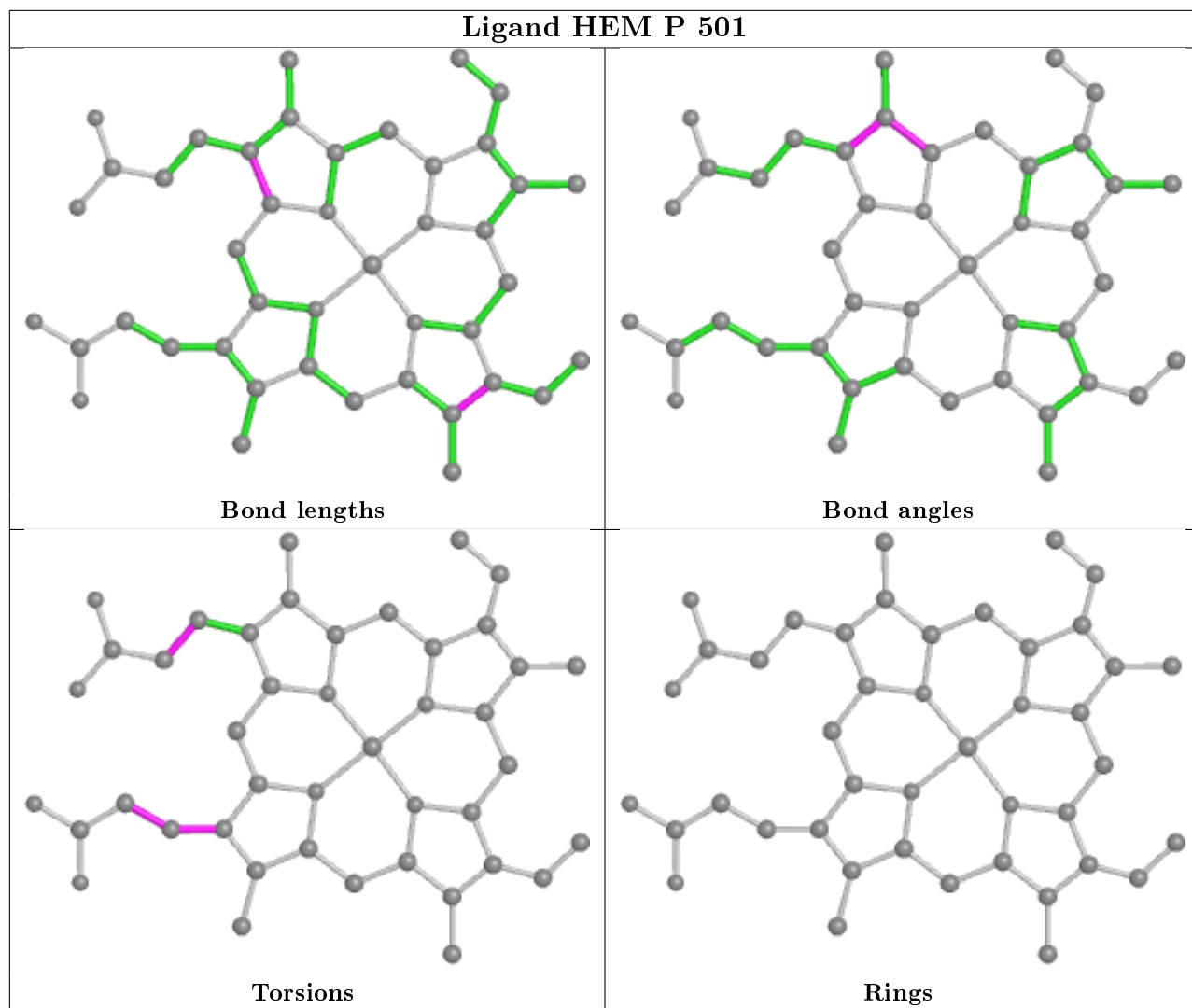
*Continued from previous page...*

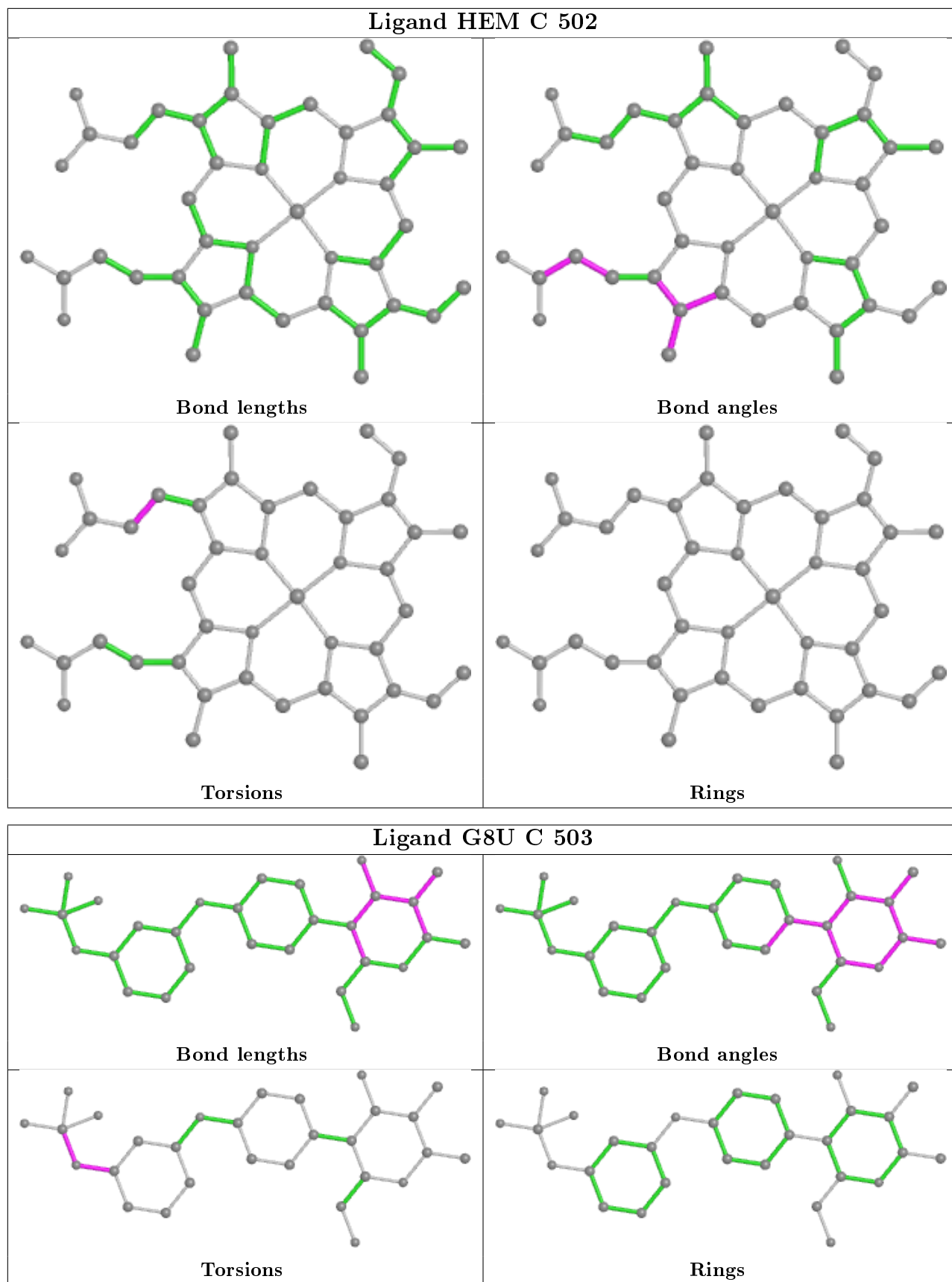
Mol	Chain	Res	Type	Clashes	Symm-Clashes
14	C	503	G8U	13	0
16	Q	506	PEE	3	0
13	C	501	HEM	5	0
14	P	503	G8U	9	0
13	P	502	HEM	2	0
18	T	501	CDL	5	0
16	D	506	PEE	1	0

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.

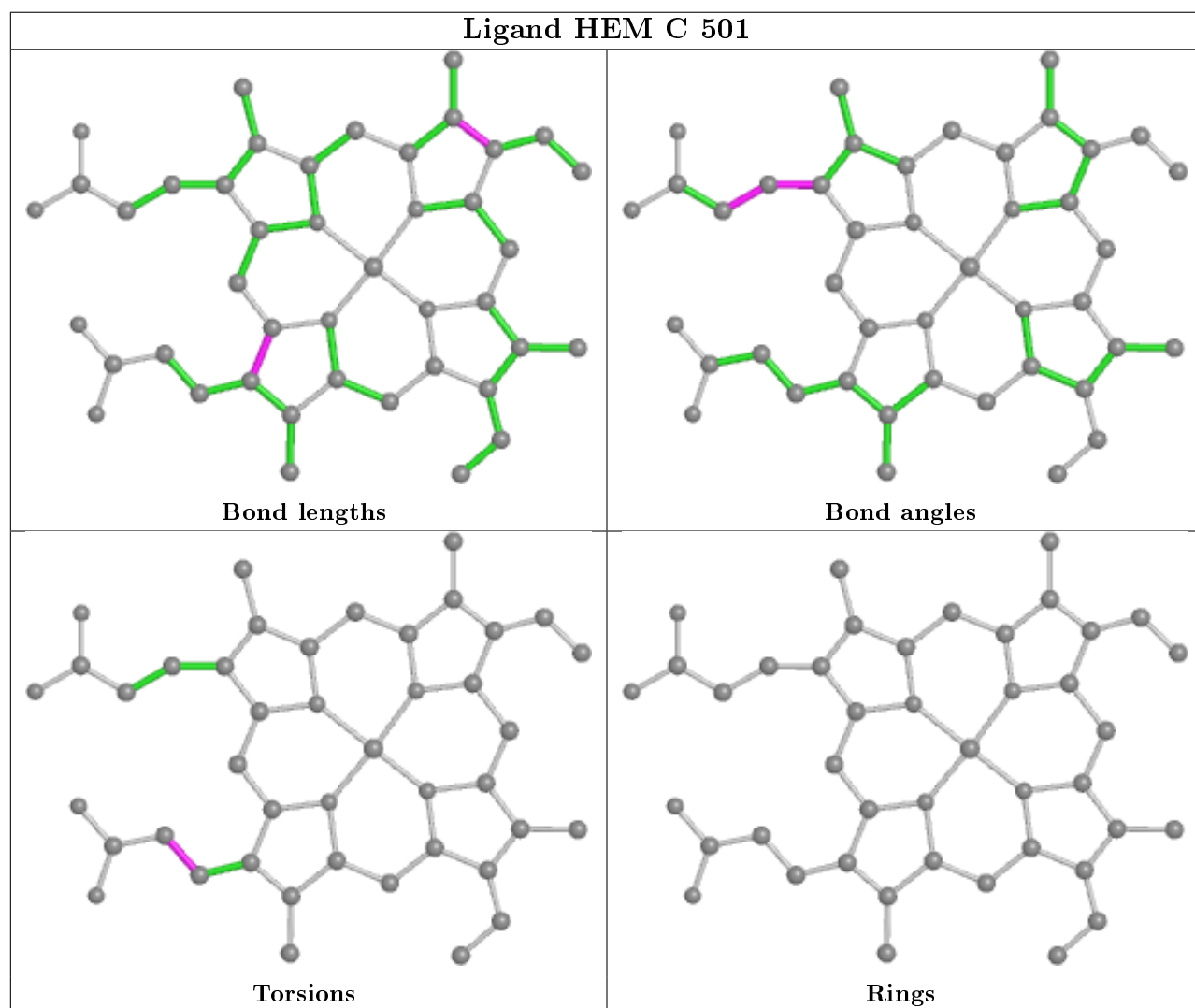
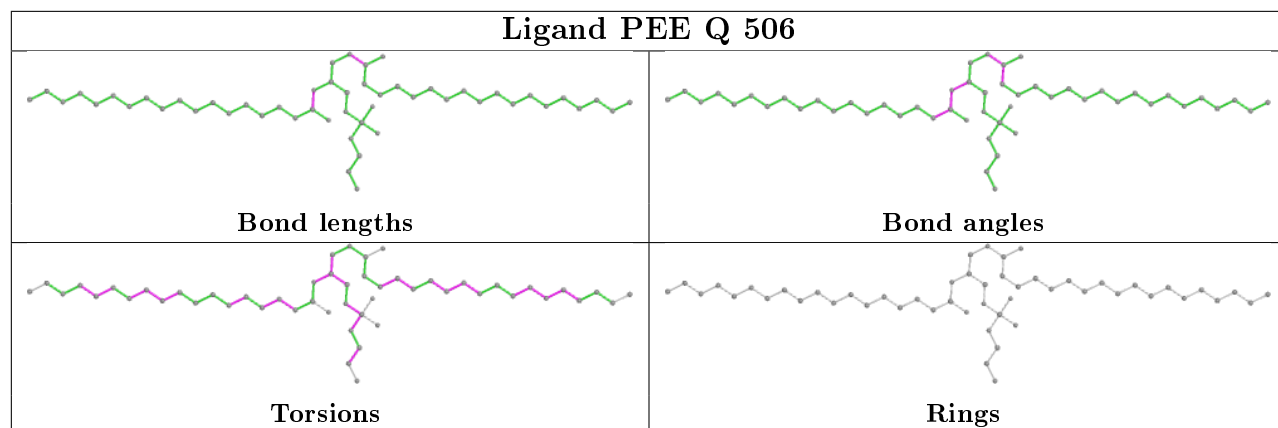


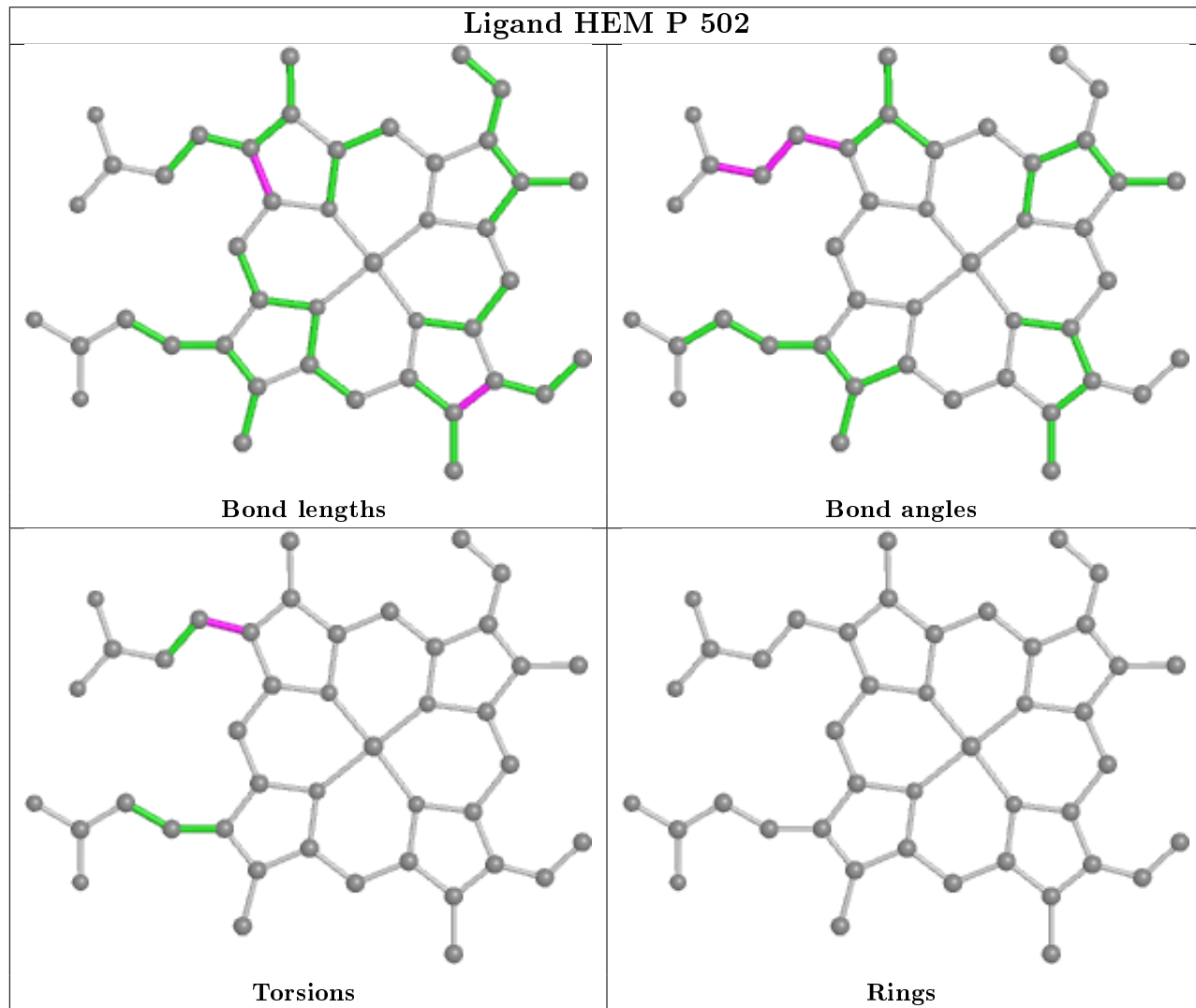
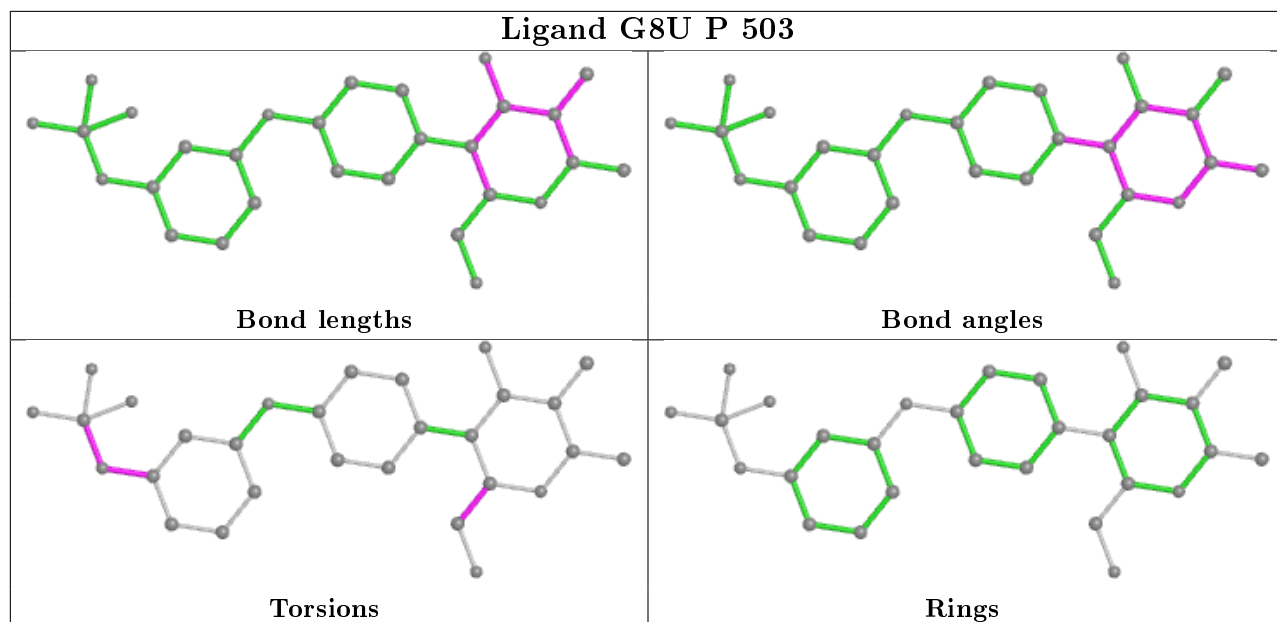


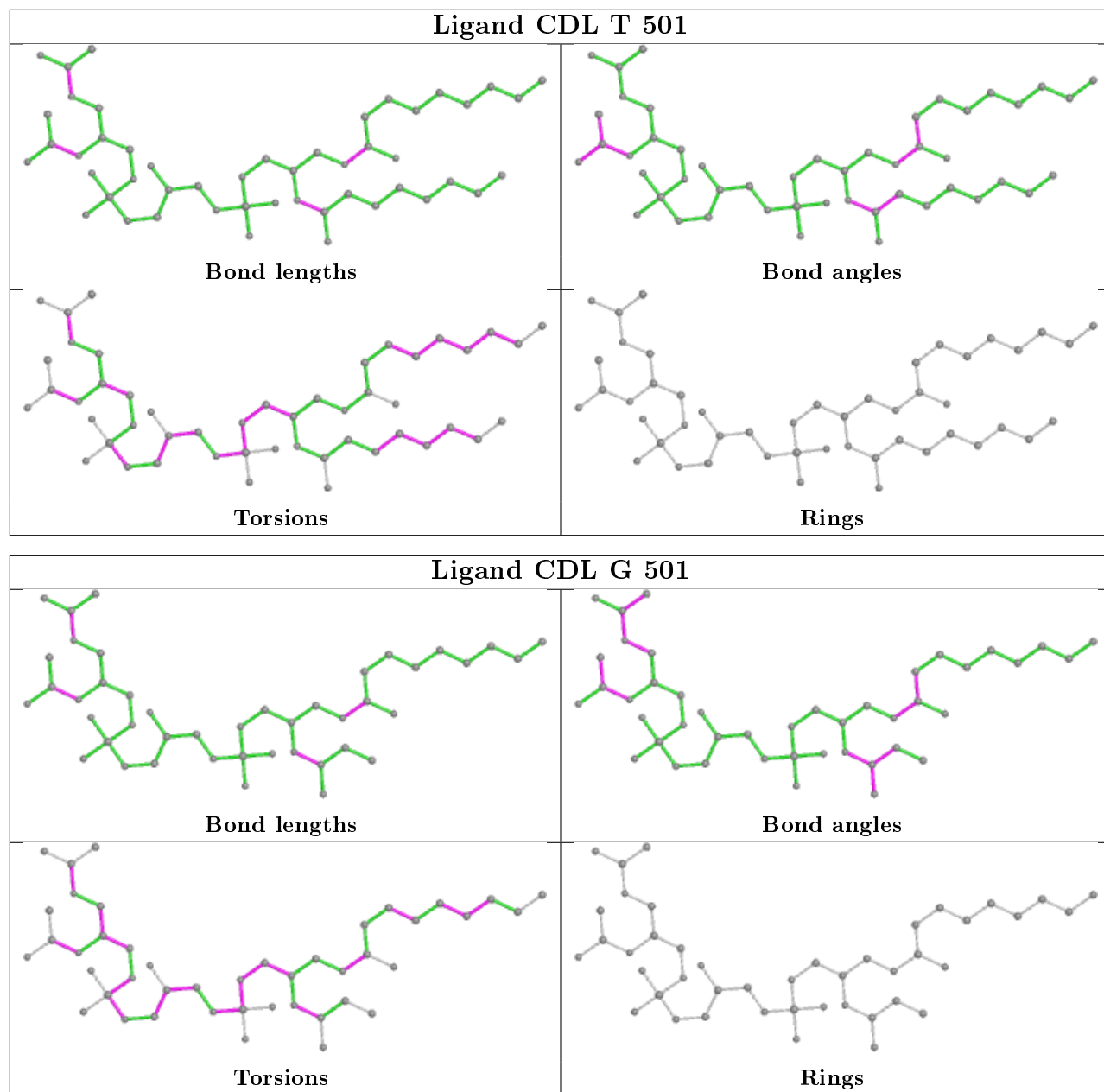


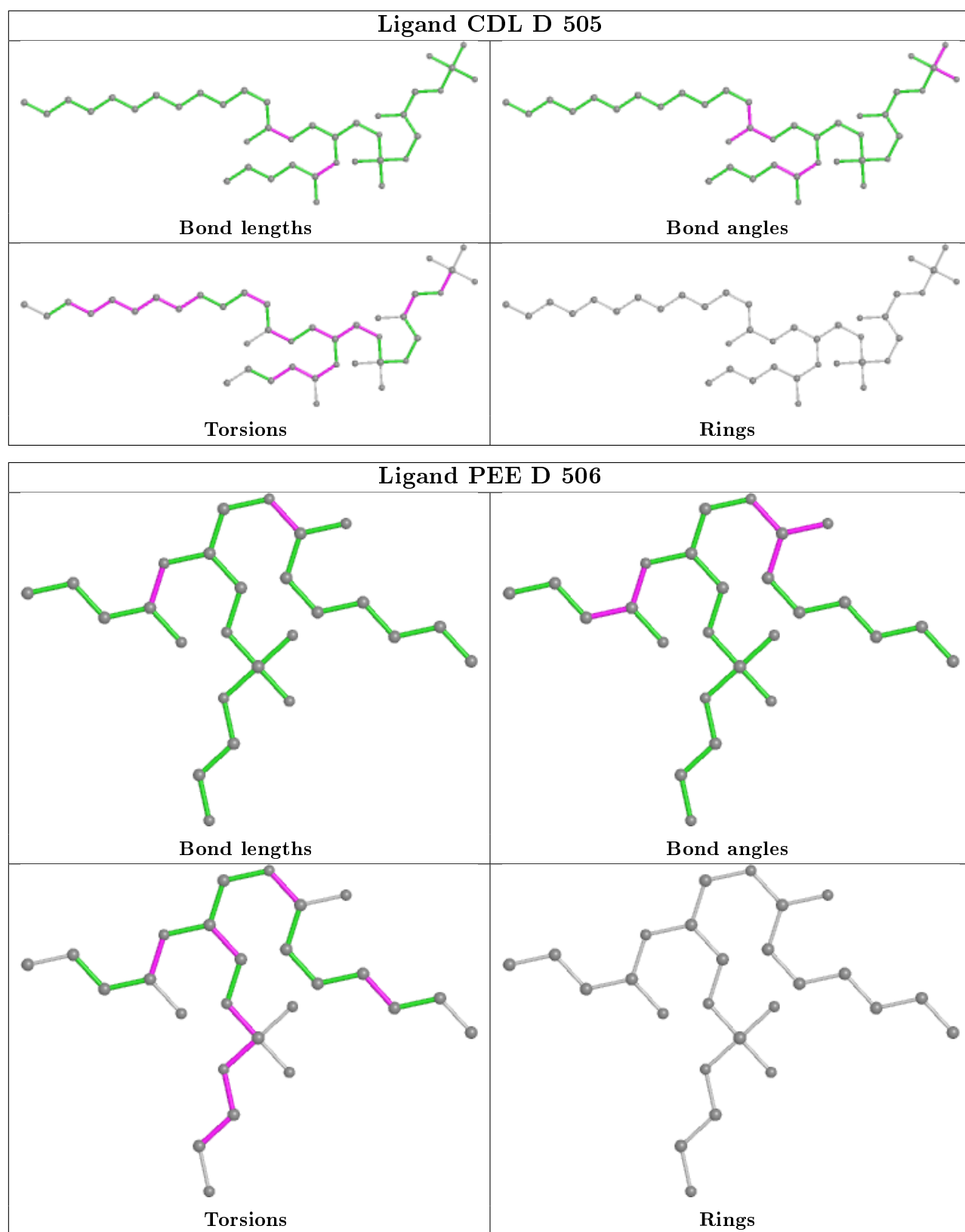












## 5.7 Other polymers [\(i\)](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues

The following chains have linkage breaks:

Mol	Chain	Number of breaks
10	N	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	N	444:LEU	C	445:LYS	N	0.87

## 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	444/480 (92%)	-0.06	7 (1%) 72 62	24, 182, 228, 270	0
2	B	422/453 (93%)	0.14	20 (4%) 31 26	126, 187, 223, 271	0
3	C	374/379 (98%)	-0.45	0 100 100	23, 131, 164, 249	0
3	P	370/379 (97%)	-0.39	0 100 100	99, 138, 177, 195	0
4	D	240/325 (73%)	-0.28	0 100 100	119, 149, 186, 210	0
4	Q	241/325 (74%)	-0.17	0 100 100	115, 162, 198, 232	0
5	E	73/274 (26%)	-0.34	0 100 100	123, 155, 194, 211	0
5	I	27/274 (9%)	0.29	1 (3%) 41 33	168, 214, 255, 306	0
5	R	196/274 (71%)	0.07	10 (5%) 28 24	30, 190, 238, 288	0
6	F	98/111 (88%)	-0.34	1 (1%) 82 74	121, 153, 190, 200	0
6	S	99/111 (89%)	-0.50	0 100 100	117, 148, 182, 208	0
7	G	80/82 (97%)	-0.23	2 (2%) 57 47	30, 149, 204, 273	0
7	T	74/82 (90%)	-0.26	0 100 100	114, 156, 198, 212	0
8	H	65/91 (71%)	-0.17	0 100 100	130, 160, 190, 202	0
8	U	66/91 (72%)	-0.35	1 (1%) 73 63	161, 190, 238, 284	0
9	J	58/64 (90%)	0.08	1 (1%) 70 60	131, 158, 191, 222	0
9	W	59/64 (92%)	-0.02	1 (1%) 70 60	115, 150, 186, 198	0
10	N	444/480 (92%)	-0.15	0 100 100	24, 162, 208, 261	0
11	O	419/453 (92%)	-0.09	6 (1%) 75 65	30, 176, 212, 247	0
12	V	17/274 (6%)	0.95	2 (11%) 4 5	174, 195, 225, 228	0
All	All	3866/5066 (76%)	-0.16	52 (1%) 77 68	23, 162, 216, 306	0

The worst 5 of 52 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
12	V	62	ARG	6.6
5	R	118	ARG	4.8
2	B	19	PRO	4.5
2	B	376	GLU	4.2
2	B	232	LEU	4.2

## 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 carbohydrates 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.

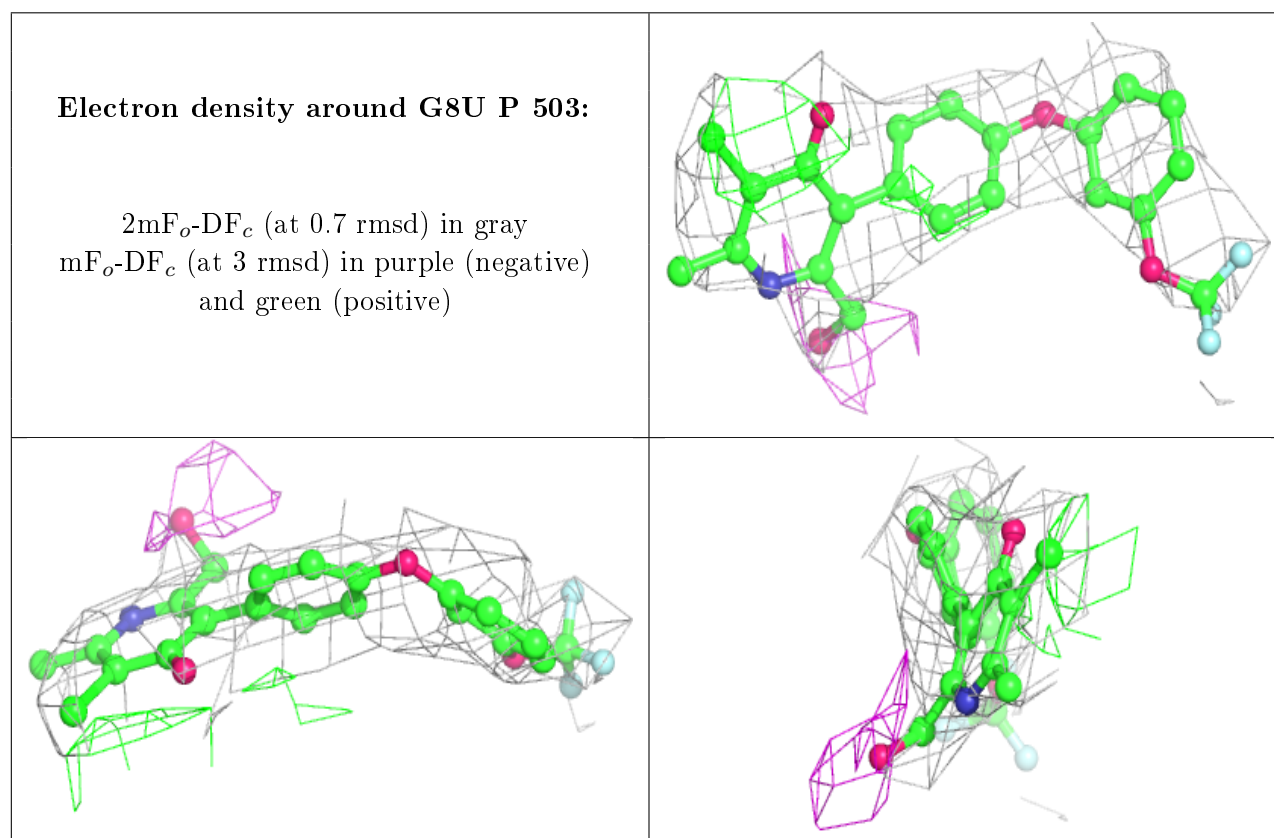
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
15	PO4	D	504	5/5	0.26	0.50	246,252,254,257	0
15	PO4	S	501	5/5	0.40	0.52	268,268,275,278	0
15	PO4	F	501	5/5	0.45	0.58	251,252,258,259	0
15	PO4	N	501	5/5	0.52	0.58	155,167,169,176	0
15	PO4	D	502	5/5	0.64	0.39	234,240,241,242	0
20	GOL	R	502	6/6	0.64	1.37	215,221,224,225	0
15	PO4	D	503	5/5	0.66	1.20	212,225,230,231	0
14	G8U	P	503	29/29	0.69	0.50	125,164,212,219	0
14	G8U	C	503	29/29	0.74	0.67	128,171,211,230	0
15	PO4	E	501	5/5	0.79	0.29	142,147,152,152	0
16	PEE	Q	506	51/51	0.80	0.55	125,148,166,171	0
15	PO4	C	504	5/5	0.80	0.35	156,159,161,168	0
18	CDL	D	505	39/100	0.81	0.37	112,152,165,174	0
18	CDL	Q	505	39/100	0.86	0.38	116,140,173,174	0
15	PO4	P	504	5/5	0.86	0.49	159,166,173,174	0
16	PEE	D	506	26/51	0.88	0.43	124,151,176,178	0
16	PEE	C	505	49/51	0.89	0.43	113,144,173,183	0
18	CDL	G	501	44/100	0.90	0.29	111,137,175,179	0
18	CDL	T	501	49/100	0.91	0.29	111,152,190,204	0

*Continued on next page...*

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
16	PEE	P	505	49/51	0.92	0.40	129,146,172,180	0
19	FES	R	501	4/4	0.94	0.10	200,213,218,224	0
17	HEC	D	501	43/43	0.96	0.30	125,134,153,161	0
17	HEC	Q	501	43/43	0.96	0.37	119,145,161,164	0
13	HEM	P	501	43/43	0.98	0.30	113,130,145,151	0
13	HEM	C	501	43/43	0.98	0.33	110,118,132,142	0
13	HEM	P	502	43/43	0.98	0.34	100,107,130,132	0
13	HEM	C	502	43/43	0.98	0.36	96,108,120,147	0

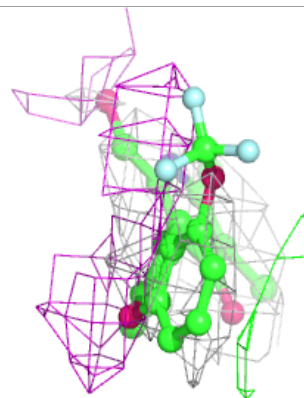
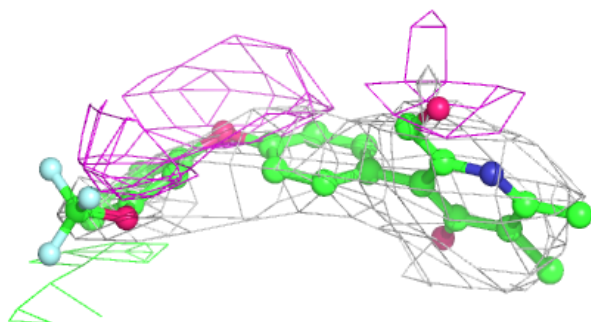
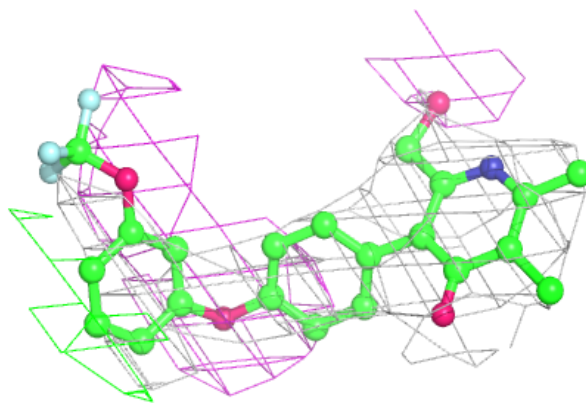
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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



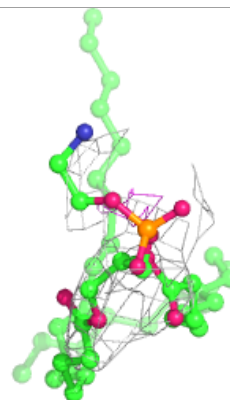
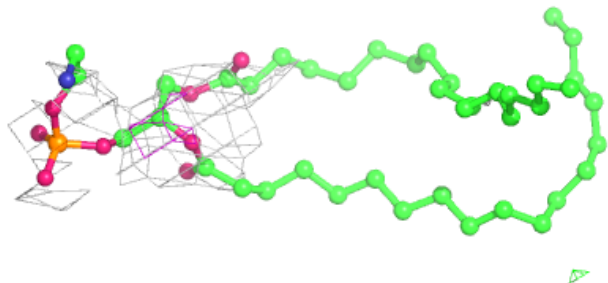
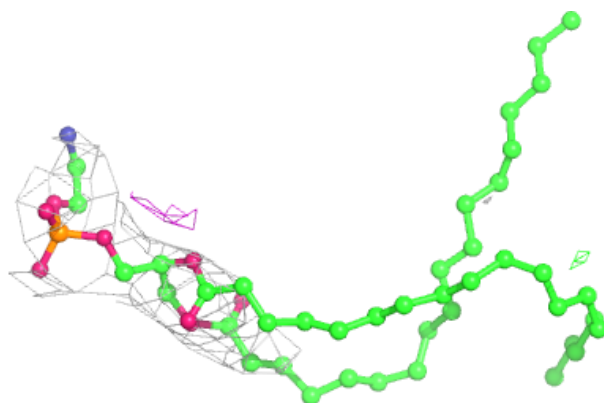


**Electron density around G8U C 503:**

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

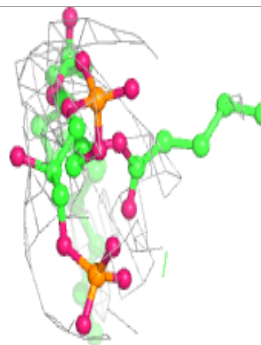
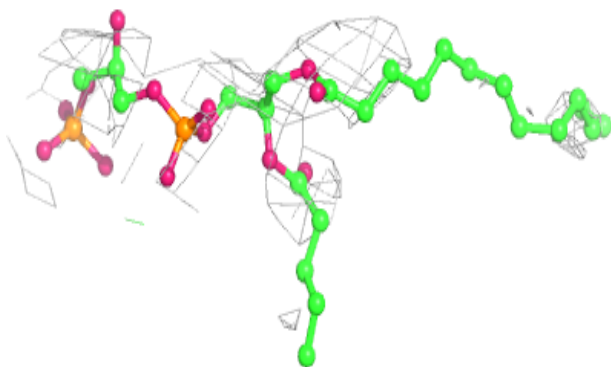
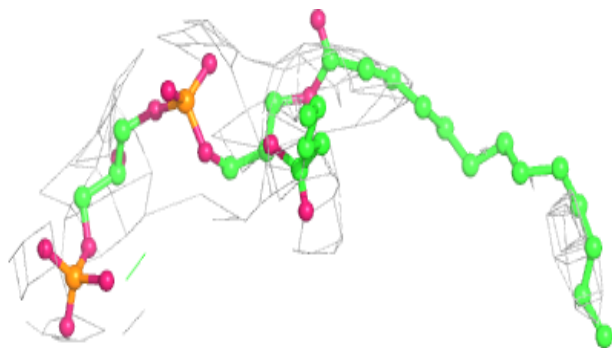
**Electron density around PEE Q 506:**

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

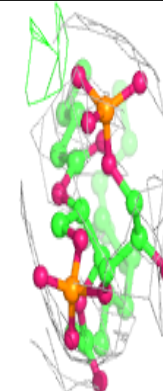
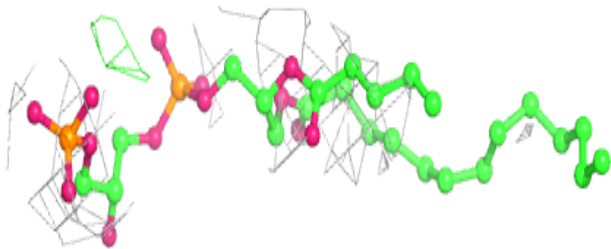
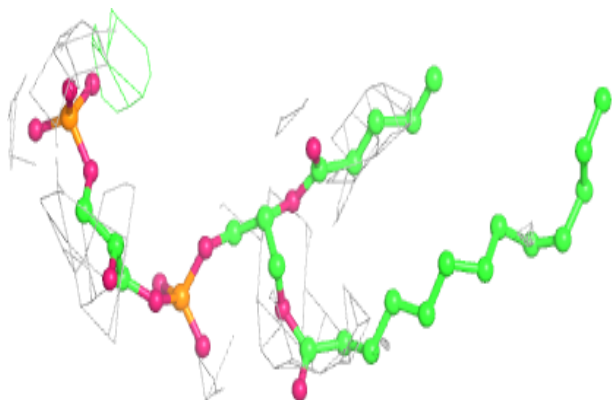


**Electron density around CDL D 505:**

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

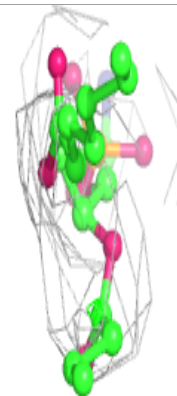
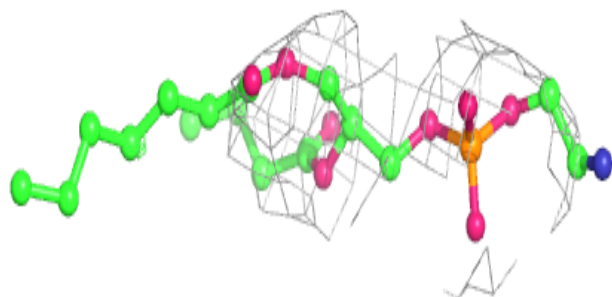
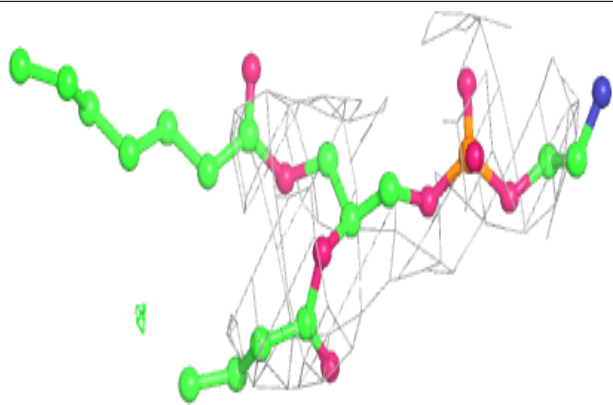
**Electron density around CDL Q 505:**

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

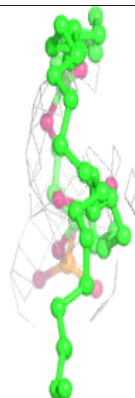
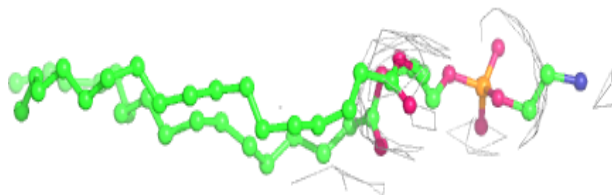
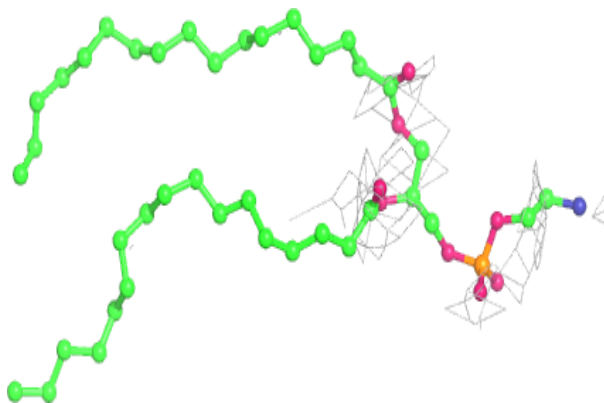


**Electron density around PEE D 506:**

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

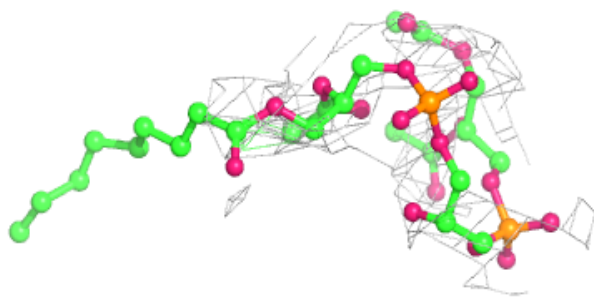
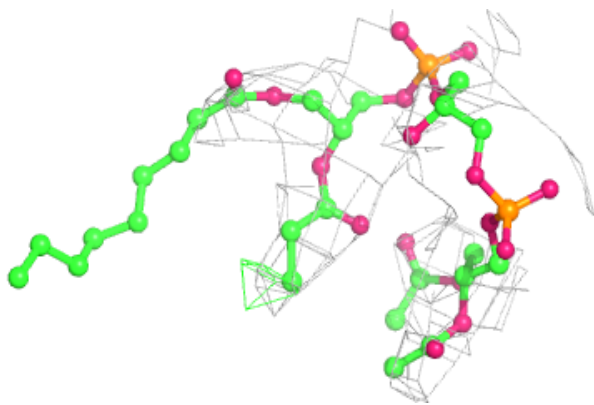
**Electron density around PEE C 505:**

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

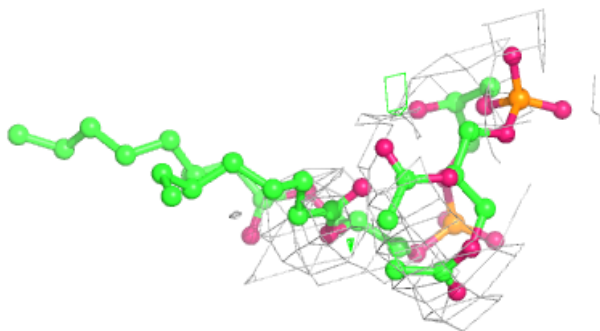
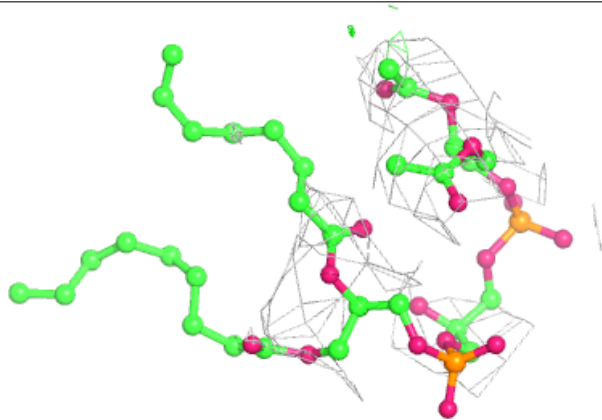


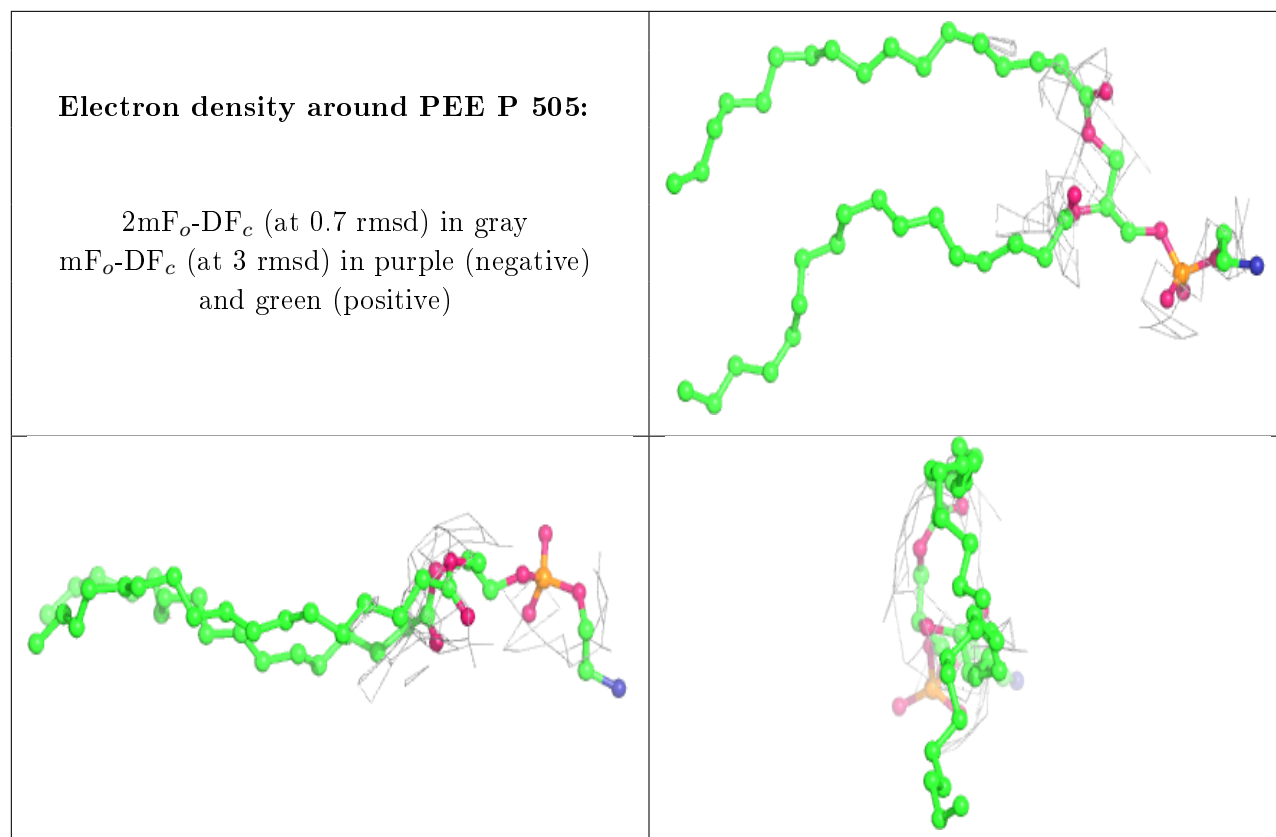
**Electron density around CDL G 501:**

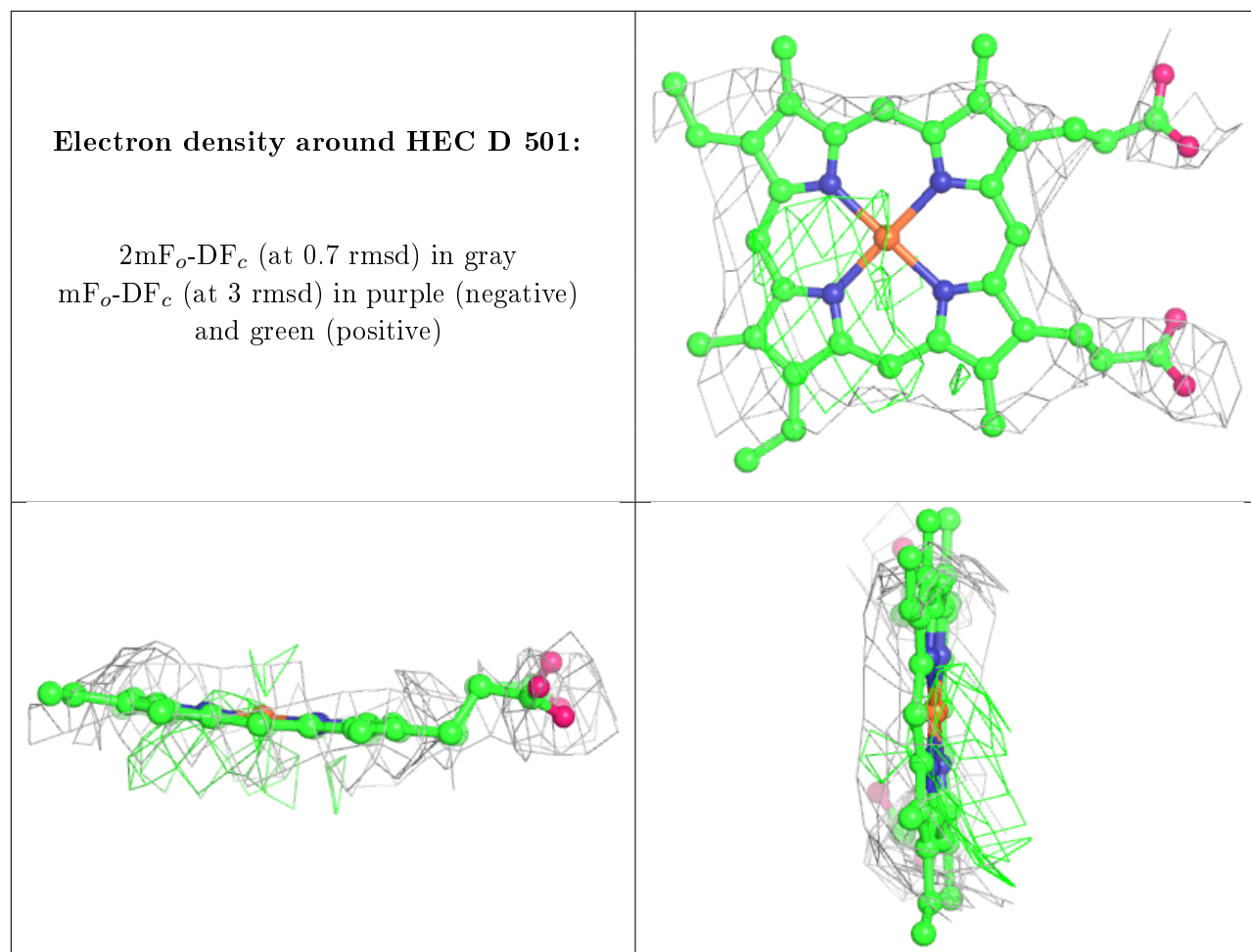
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)

**Electron density around CDL T 501:**

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

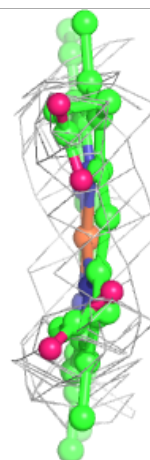
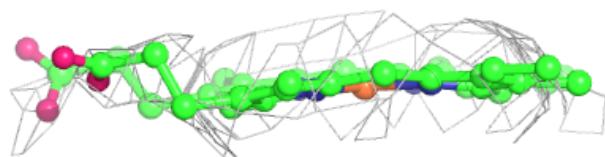
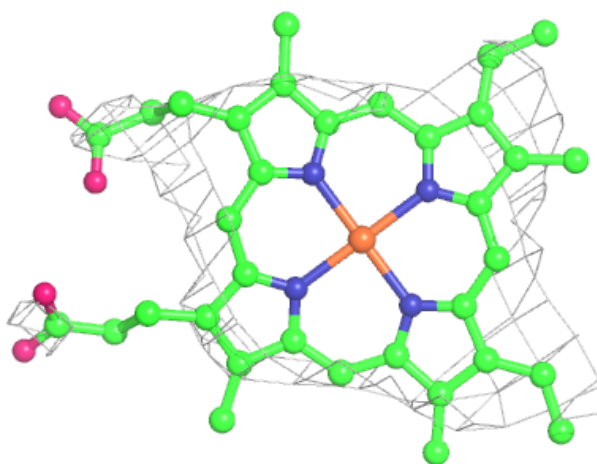


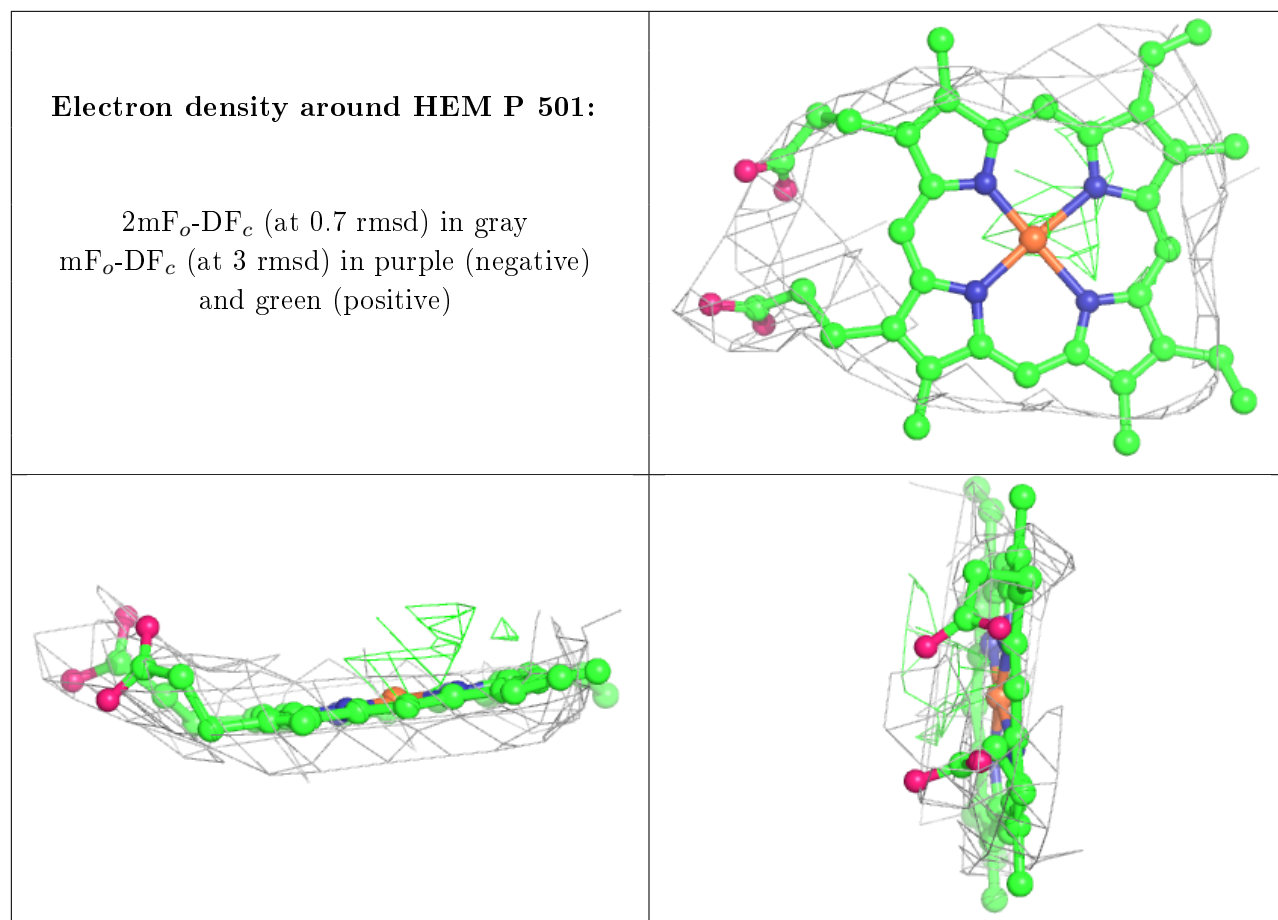




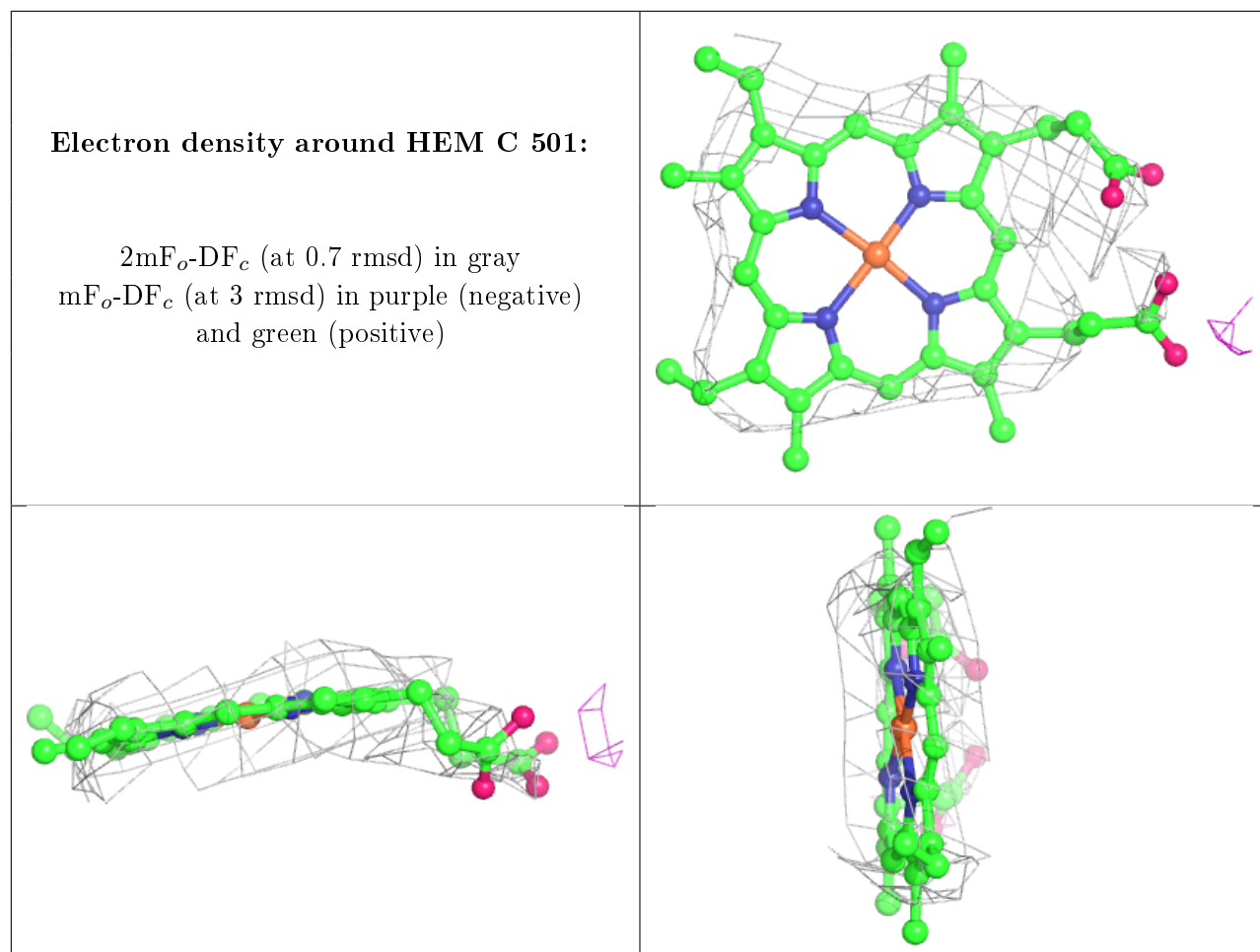
**Electron density around HEC Q 501:**

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



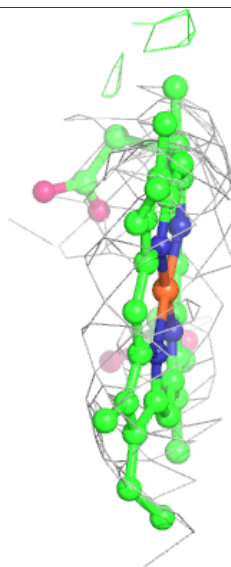
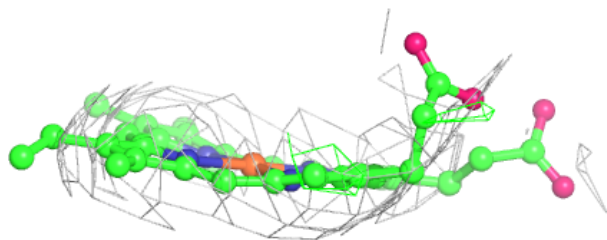
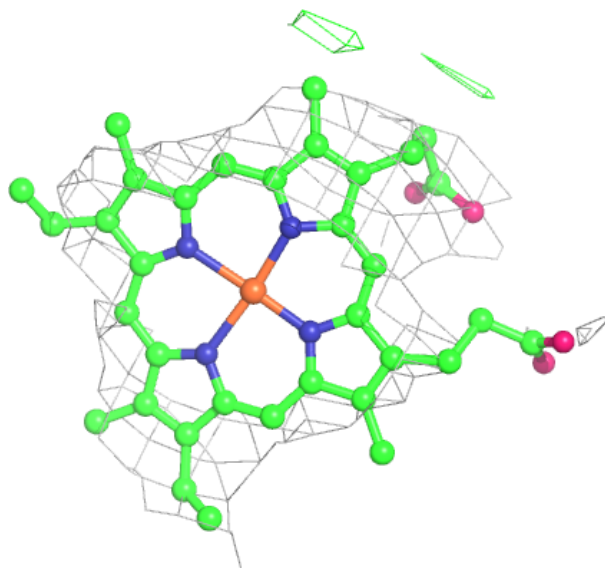


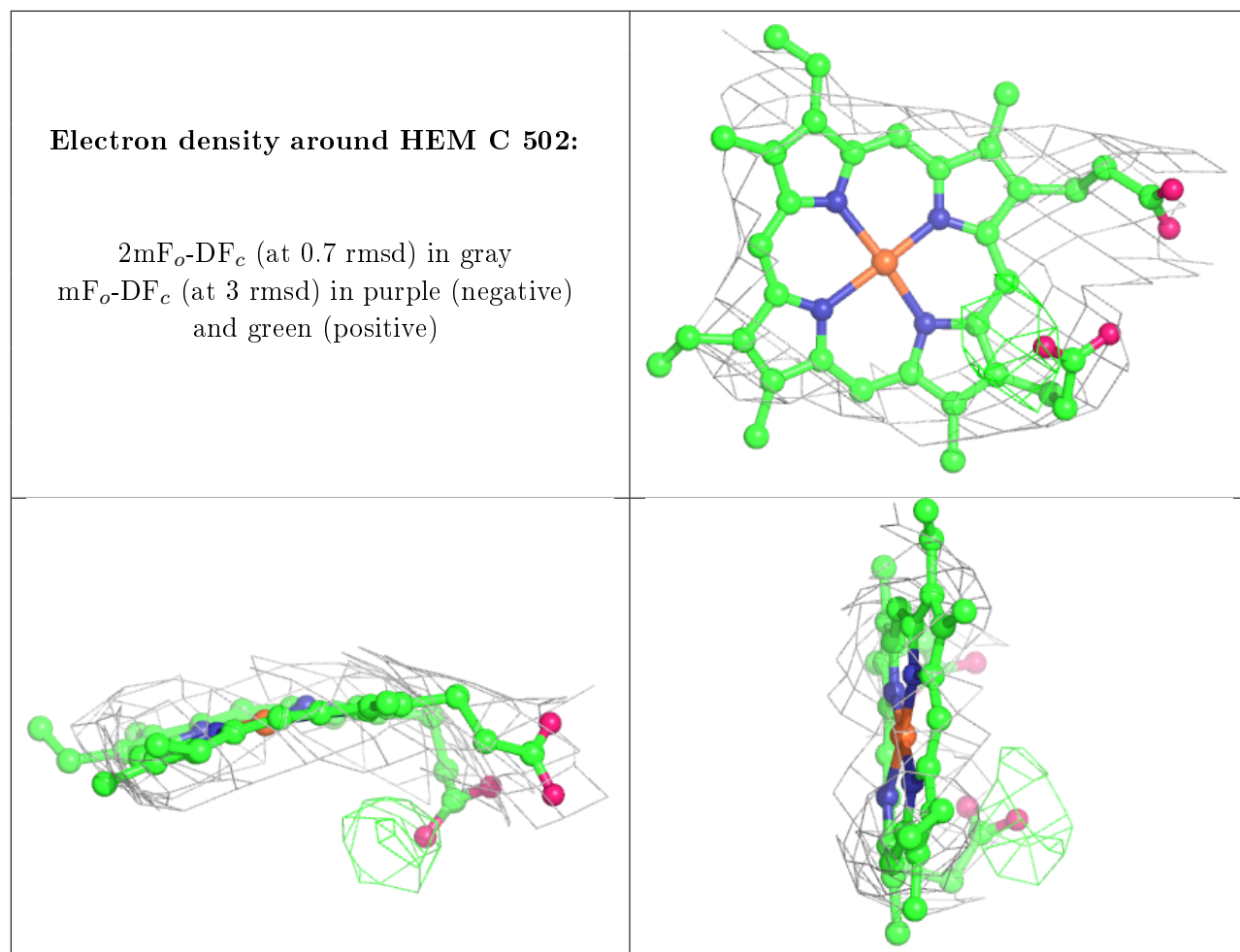




**Electron density around HEM P 502:**

$2mF_o-DF_c$  (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.