wwPDB X-ray Structure Validation Summary Report (i)

May 14, 2020 – 10:45 pm BST

PDB ID	:	4H13
Title	:	Crystal Structure of the Cytochrome b6f Complex from Mastigocladus laminosus with TDS
Authors	:	Hasan, S.S.; Yamashita, E.; Baniulis, D.; Cramer, W.A.
Deposited on	:	2012-09-10
Resolution	:	3.07 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 (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

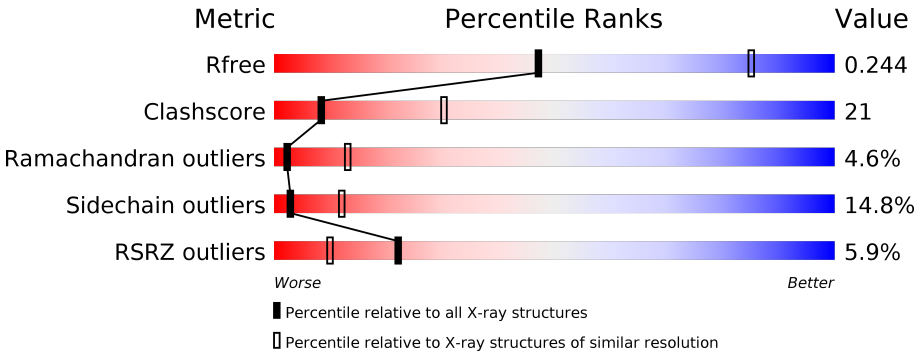
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.07 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	1447 (3.10-3.06)
Clashscore	141614	1546 (3.10-3.06)
Ramachandran outliers	138981	1487 (3.10-3.06)
Sidechain outliers	138945	1486 (3.10-3.06)
RSRZ outliers	127900	1416 (3.10-3.06)

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 <=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	215	% 68% 26% 6%
2	B	160	3% 60% 37% .
3	C	289	9% 50% 40% 10%
4	D	179	13% 48% 37% 7% . 6%
5	E	32	56% 34% 9%
6	F	35	3% 54% 17% 17% 11%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
7	G	37	3% 46% 32% 22%
8	H	29	66% 28% 7%

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
11	OPC	A	305	-	-	-	X
14	UMQ	A	308	X	-	-	X
14	UMQ	B	202	X	-	-	X
16	CLA	B	203	X	-	-	-
17	7PH	C	301	-	-	-	X
18	SQD	D	201	X	-	-	X
20	OCT	F	101	-	-	-	X
21	BCR	G	101	-	-	-	X

2 Entry composition

There are 22 unique types of molecules in this entry. The entry contains 16526 atoms, of which 8399 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 b6.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	215	Total	C	H	N	O	S	0	0	0
			3449	1140	1738	272	288	11			

- Molecule 2 is a protein called Cytochrome b6-f complex subunit 4.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	160	Total	C	H	N	O	S	0	0	0
			2558	841	1309	193	209	6			

- Molecule 3 is a protein called Apocytochrome f.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	C	288	Total	C	H	N	O	S	0	0	0
			4451	1415	2235	369	424	8			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	11	PRO	GLU	SEE REMARK 999	UNP P83793

- Molecule 4 is a protein called Cytochrome b6-f complex iron-sulfur subunit.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
4	D	168	Total	C	H	N	O	S	0	0	0
			2563	823	1275	221	237	7			

- Molecule 5 is a protein called Cytochrome b6-f complex subunit 6.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
5	E	32	Total	C	H	N	O	S	0	0	0
			532	179	284	34	34	1			

- Molecule 6 is a protein called Cytochrome b6-f complex subunit 7.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
6	F	31	483	160	249	34	39	1	0	0	0

- Molecule 7 is a protein called Cytochrome b6-f complex subunit 5.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
7	G	37	572	188	289	44	50	1	0	0	0

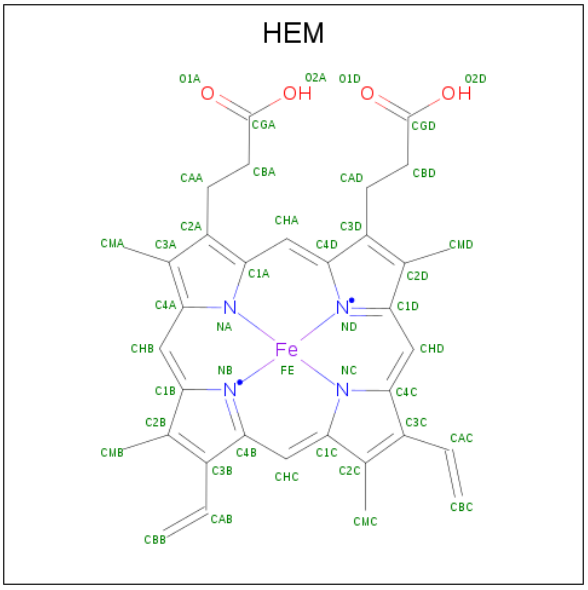
- Molecule 8 is a protein called Cytochrome b6-f complex subunit 8.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
8	H	29	469	156	239	36	36	2	0	0	0

- Molecule 9 is CADMIUM ION (three-letter code: CD) (formula: Cd).

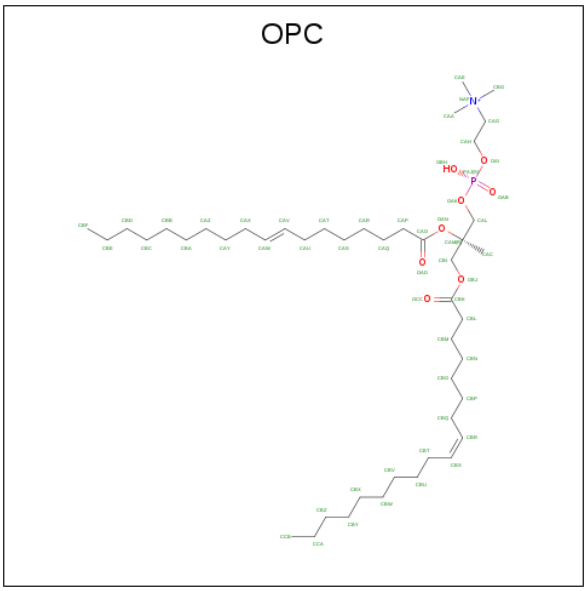
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Cd		
9	A	1	1	1	0	0

- Molecule 10 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (formula: $C_{34}H_{32}FeN_4O_4$).



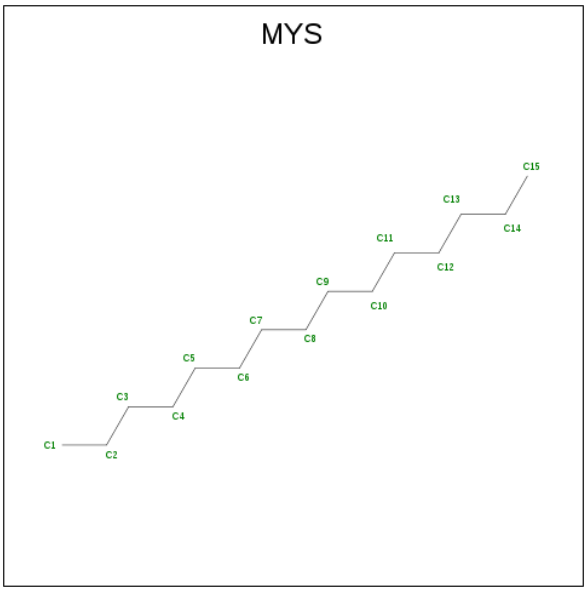
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
10	A	1	Total	C	Fe	H	N	O	0	0
			73	34	1	30	4	4		
10	A	1	Total	C	Fe	H	N	O	0	0
			73	34	1	30	4	4		
10	A	1	Total	C	Fe	H	N	O	0	0
			73	34	1	30	4	4		
10	C	1	Total	C	Fe	H	N	O	0	0
			73	34	1	30	4	4		

- Molecule 11 is (7R,17E)-4-HYDROXY-N,N,N,7-TETRAMETHYL-7-[(8E)-OCTADEC-8-ENOYLOXY]-10-OXO-3,5,9-TRIOXA-4-PHOSPHAHEPTACOS-17-EN-1-AMINIUM 4-OXIDE (three-letter code: OPC) (formula: $C_{45}H_{87}NO_8P$).



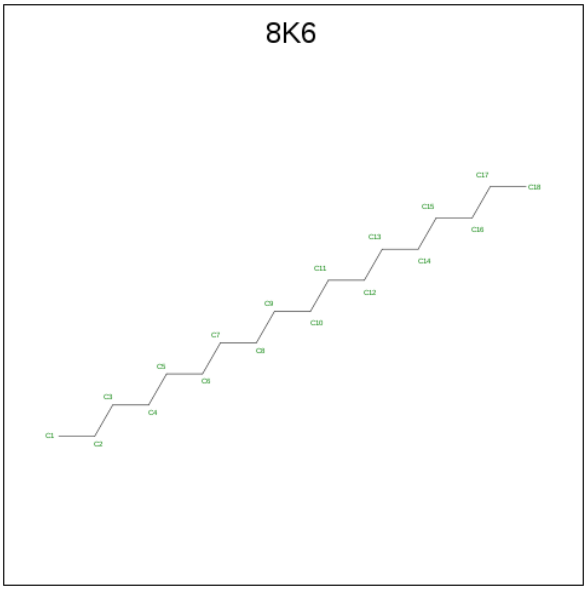
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
11	A	1	Total	C	H	N	O	P	0	0
			137	44	83	1	8	1		
11	B	1	Total	C	H	N	O	P	0	0
			137	44	83	1	8	1		

- Molecule 12 is PENTADECANE (three-letter code: MYS) (formula: $C_{15}H_{32}$).



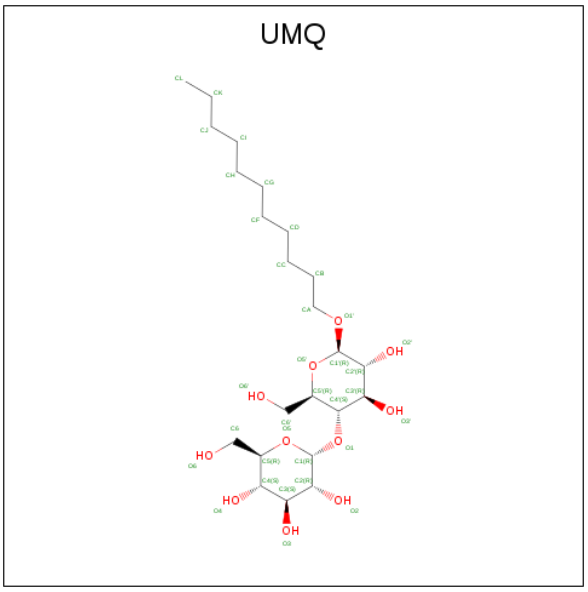
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	H		
12	A	1	47	15	32	0	0

- Molecule 13 is Octadecane (three-letter code: 8K6) (formula: $C_{18}H_{38}$).



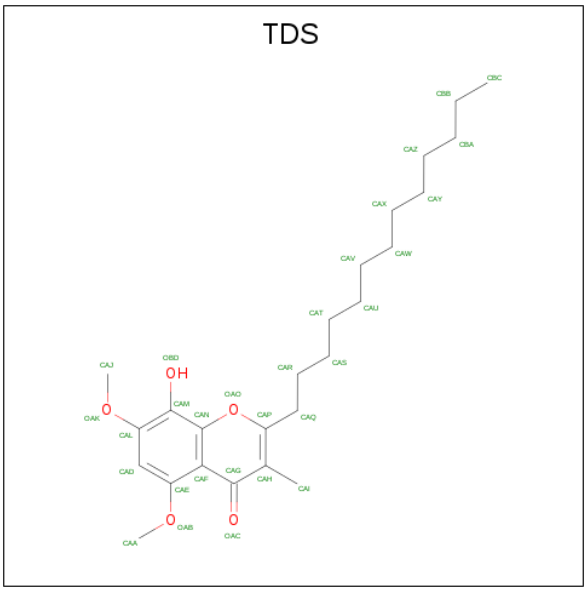
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	H		
13	A	1	56	18	38	0	0

- Molecule 14 is UNDECYL-MALTOSIDE (three-letter code: UMQ) (formula: $C_{23}H_{44}O_{11}$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
14	A	1	Total	C	H	O	0	0
			77	23	43	11		
14	B	1	Total	C	H	O	0	0
			77	23	43	11		

- Molecule 15 is 8-HYDROXY-5,7-DIMETHOXY-3-METHYL-2-TRIDECYL-4H-CHROMEN-4-ONE (three-letter code: TDS) (formula: $C_{25}H_{38}O_5$).



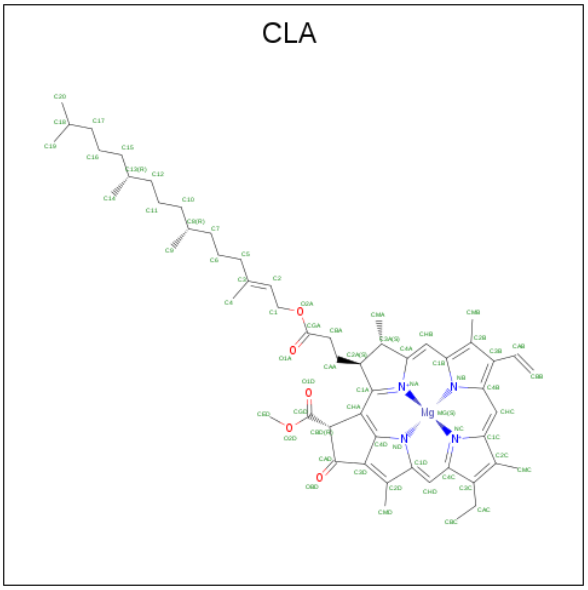
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
15	A	1	Total	C	H	O	0	0
			68	25	38	5		

Continued on next page...

Continued from previous page...

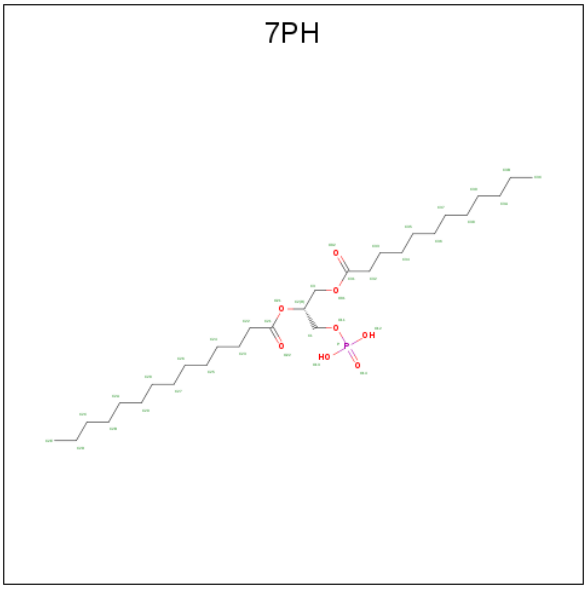
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	H	O		
15	B	1	68	25	38	5	0	0

- Molecule 16 is CHLOROPHYLL A (three-letter code: CLA) (formula: $C_{55}H_{72}MgN_4O_5$).



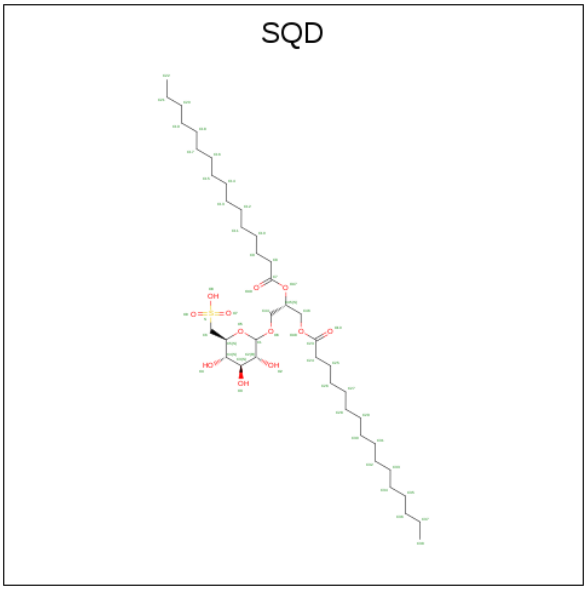
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
			Total	C	H	Mg	N	O		
16	B	1	127	55	62	1	4	5	0	0

- Molecule 17 is (1R)-2-(dodecanoyloxy)-1-[(phosphonooxy)methyl]ethyl tetradecanoate (three-letter code: 7PH) (formula: $C_{29}H_{57}O_8P$).



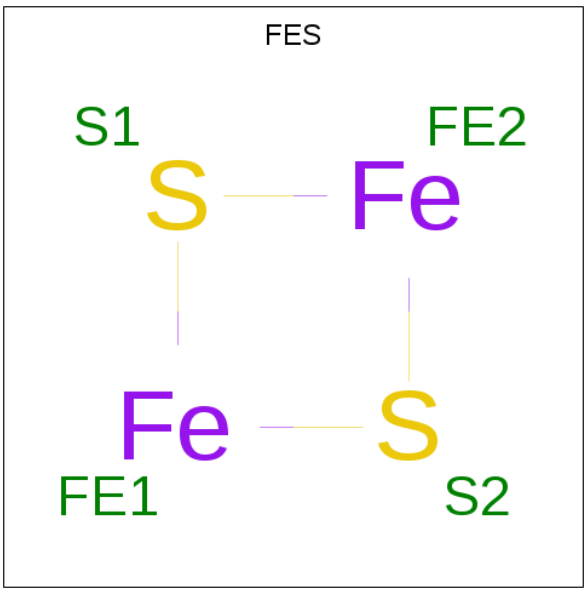
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
17	C	1	Total	C	H	O	0	0
			81	27	49	5		

- Molecule 18 is 1,2-DI-O-ACYL-3-O-[6-DEOXY-6-SULFO-ALPHA-D-GLUCOPYRANOSYL]-SN-GLYCEROL (three-letter code: SQD) (formula: $C_{41}H_{78}O_{12}S$).



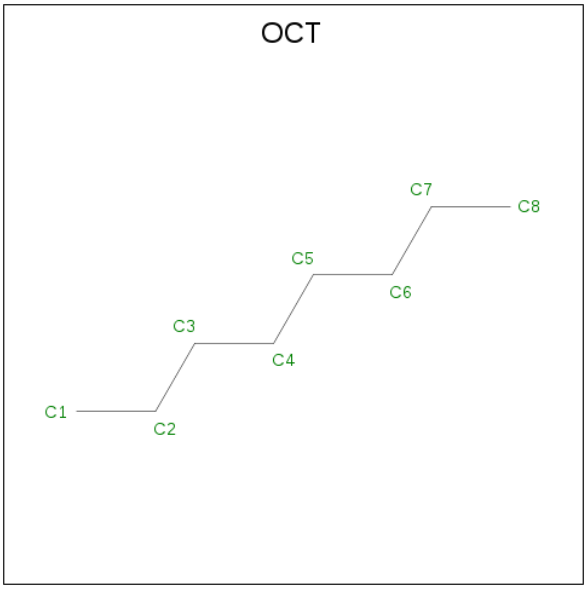
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
18	D	1	Total	C	H	O	S	1	0
			131	41	78	11	1		

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



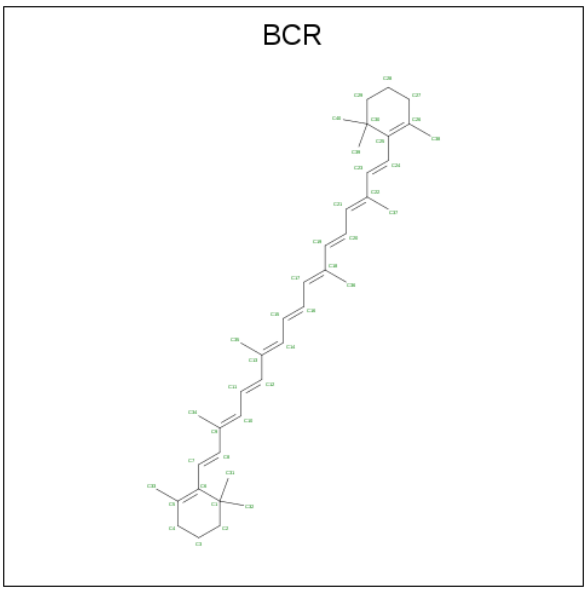
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
19	D	1	Total Fe S 4 2 2	0	0

- Molecule 20 is N-OCTANE (three-letter code: OCT) (formula: C_8H_{18}).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
20	F	1	Total C H 26 8 18	0	0

- Molecule 21 is BETA-CAROTENE (three-letter code: BCR) (formula: $C_{40}H_{56}$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
21	G	1	Total	C	H	0	0
			96	40	56		

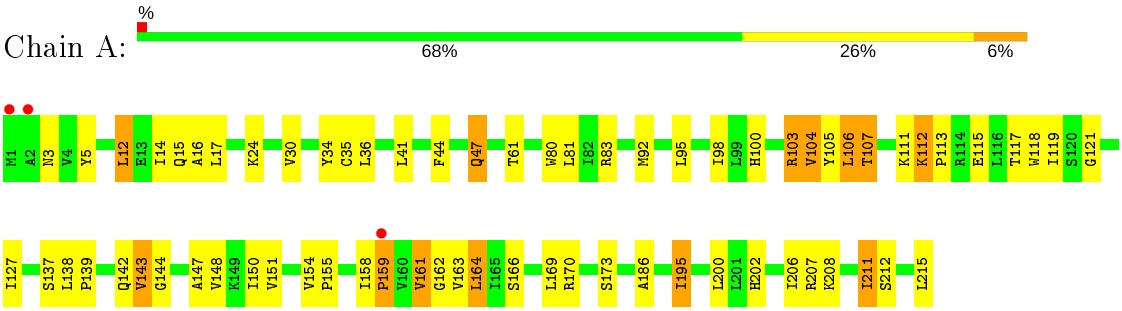
- Molecule 22 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
22	A	12	Total	O	0	0
			12	12		
22	B	6	Total	O	0	0
			6	6		
22	C	4	Total	O	0	0
			4	4		
22	F	1	Total	O	0	0
			1	1		
22	G	1	Total	O	0	0
			1	1		

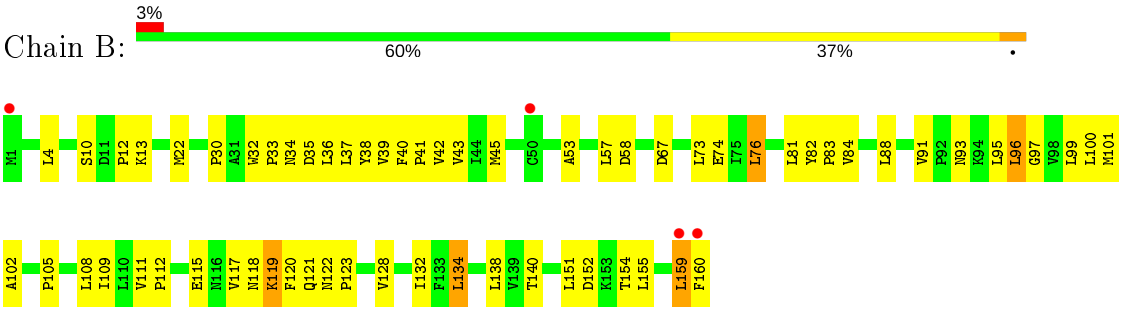
3 Residue-property plots (i)

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.

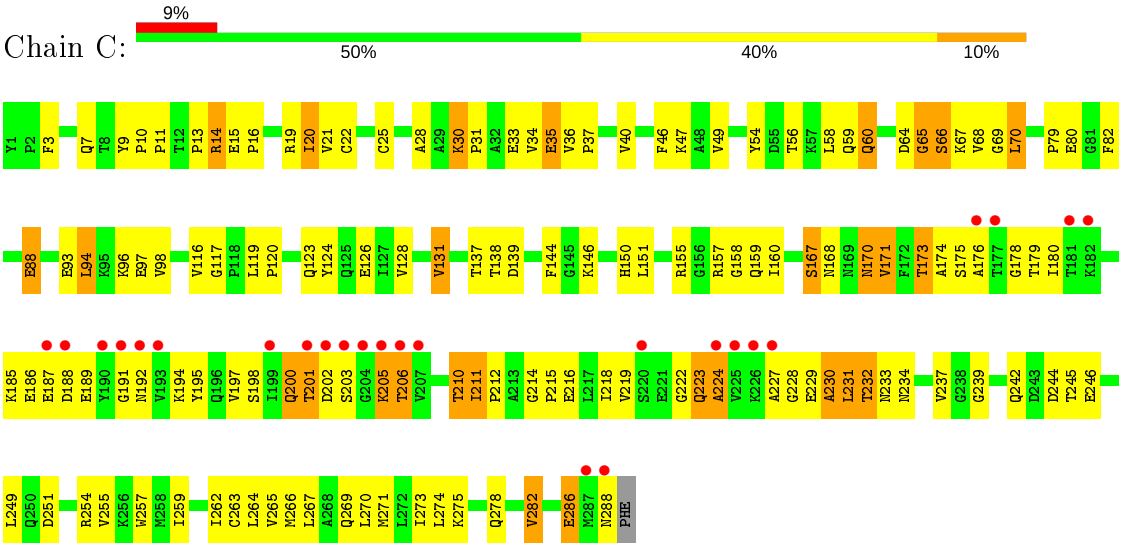
- Molecule 1: Cytochrome b6



- Molecule 2: Cytochrome b6-f complex subunit 4

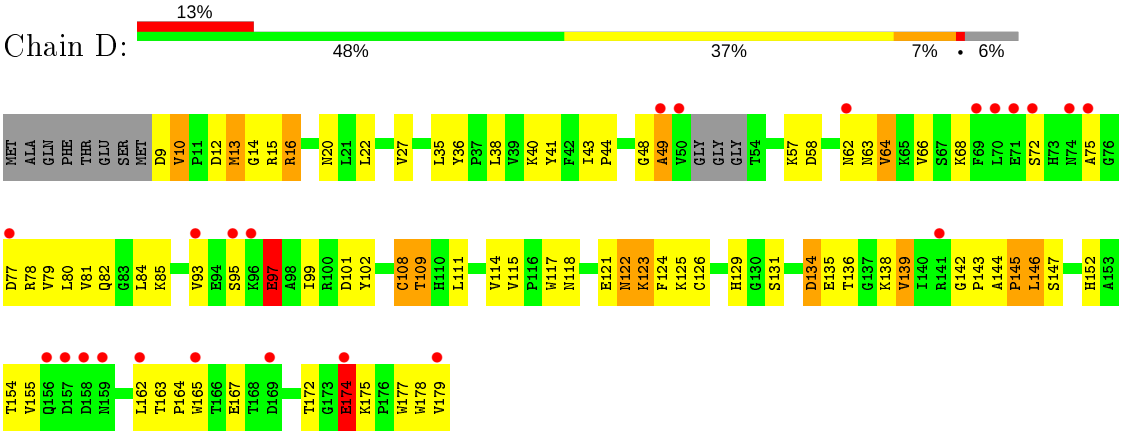


- Molecule 3: Apocytochrome f



• Molecule 4: Cytochrome b6-f complex iron-sulfur subunit

Chain D:

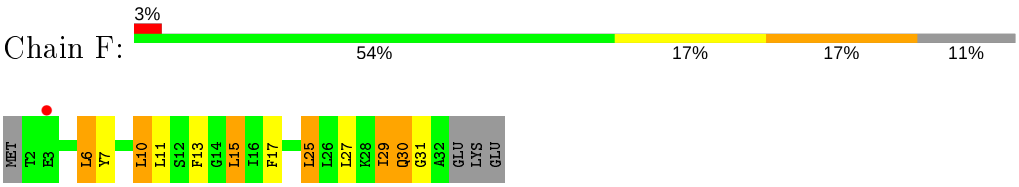


• Molecule 5: Cytochrome b6-f complex subunit 6

Chain E:

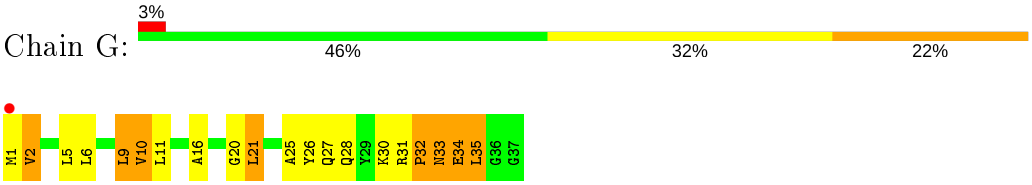
• Molecule 6: Cytochrome b6-f complex subunit 7

Chain F:



• Molecule 7: Cytochrome b6-f complex subunit 5

Chain G:



• Molecule 8: Cytochrome b6-f complex subunit 8

Chain H:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 61 2 2	Depositor
Cell constants a, b, c, α, β, γ	157.20Å 157.20Å 363.28Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	34.57 – 3.07 34.57 – 3.07	Depositor EDS
% Data completeness (in resolution range)	99.1 (34.57-3.07) 99.1 (34.57-3.07)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.69 (at 3.06Å)	Xtriage
Refinement program	REFMAC 5.7.0029, PHENIX 1.8.1_1168	Depositor
R, R_{free}	0.214 , 0.238 0.219 , 0.244	Depositor DCC
R_{free} test set	2544 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å^2)	91.8	Xtriage
Anisotropy	0.061	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 86.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	16526	wwPDB-VP
Average B, all atoms (Å^2)	115.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.06% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: UMQ, MYS, CLA, CD, 7PH, FES, OPC, TDS, HEM, 8K6, OCT, BCR, SQD

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	0.29	0/1763	0.50	0/2405
2	B	0.29	0/1288	0.53	0/1765
3	C	0.27	0/2264	0.56	0/3082
4	D	0.25	0/1320	0.51	0/1798
5	E	0.34	0/253	0.63	0/340
6	F	0.30	0/238	0.51	0/321
7	G	0.31	0/289	0.62	0/391
8	H	0.31	0/236	0.54	0/323
All	All	0.28	0/7651	0.53	0/10425

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
4	D	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
4	D	97	GLU	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	1711	1738	1737	67	0
2	B	1249	1309	1308	62	0
3	C	2216	2235	2233	109	0
4	D	1288	1275	1273	52	0
5	E	248	284	284	16	0
6	F	234	249	248	12	0
7	G	283	289	289	20	0
8	H	230	239	239	10	0
9	A	1	0	0	0	0
10	A	129	90	90	23	0
10	C	43	30	30	10	0
11	A	54	83	83	5	0
11	B	54	83	83	1	0
12	A	15	32	32	5	0
13	A	18	38	38	0	0
14	A	34	43	41	0	0
14	B	34	43	41	1	0
15	A	30	38	38	4	0
15	B	30	38	37	4	0
16	B	65	62	71	5	0
17	C	32	49	45	1	0
18	D	53	78	75	3	0
19	D	4	0	0	1	0
20	F	8	18	18	0	0
21	G	40	56	56	7	0
22	A	12	0	0	5	0
22	B	6	0	0	1	0
22	C	4	0	0	0	0
22	F	1	0	0	0	0
22	G	1	0	0	0	0
All	All	8127	8399	8389	345	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 21.

The worst 5 of 345 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:35:CYS:SG	10:A:304:HEM:CAB	2.57	0.92
3:C:25:CYS:SG	10:C:302:HEM:CAC	2.58	0.92
21:G:101:BCR:H321	21:G:101:BCR:HC8	1.53	0.89
10:A:303:HEM:O1A	22:A:403:HOH:O	1.98	0.81
1:A:117:THR:OG1	22:A:404:HOH:O	2.01	0.79

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	213/215 (99%)	190 (89%)	17 (8%)	6 (3%)	5 23
2	B	158/160 (99%)	137 (87%)	21 (13%)	0	100 100
3	C	286/289 (99%)	229 (80%)	36 (13%)	21 (7%)	1 5
4	D	164/179 (92%)	135 (82%)	21 (13%)	8 (5%)	2 12
5	E	30/32 (94%)	26 (87%)	3 (10%)	1 (3%)	4 19
6	F	29/35 (83%)	26 (90%)	3 (10%)	0	100 100
7	G	35/37 (95%)	22 (63%)	6 (17%)	7 (20%)	0 0
8	H	27/29 (93%)	26 (96%)	1 (4%)	0	100 100
All	All	942/976 (96%)	791 (84%)	108 (12%)	43 (5%)	2 13

5 of 43 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	112	LYS
1	A	162	GLY
3	C	66	SER
3	C	192	ASN
3	C	201	THR

5.3.2 Protein sidechains

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	184/184 (100%)	164 (89%)	20 (11%)	6 24
2	B	137/137 (100%)	123 (90%)	14 (10%)	7 26
3	C	242/243 (100%)	202 (84%)	40 (16%)	2 9
4	D	139/146 (95%)	112 (81%)	27 (19%)	1 5
5	E	25/25 (100%)	22 (88%)	3 (12%)	5 19
6	F	23/27 (85%)	17 (74%)	6 (26%)	0 1
7	G	28/28 (100%)	23 (82%)	5 (18%)	2 7
8	H	24/24 (100%)	20 (83%)	4 (17%)	2 9
All	All	802/814 (98%)	683 (85%)	119 (15%)	3 12

5 of 119 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	C	167	SER
3	C	232	THR
6	F	30	GLN
3	C	171	VAL
3	C	189	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
3	C	60	GLN
3	C	242	GLN

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 carbohydrates in this entry.

5.6 Ligand geometry (i)

Of 19 ligands modelled in this entry, 1 is monoatomic - leaving 18 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	$\#\|Z\| > 2$
11	OPC	A	305	-	53,53,54	0.98	4 (7%)	59,61,64	1.06	3 (5%)
12	MYS	A	306	-	14,14,14	0.30	0	13,13,13	0.78	0
19	FES	D	202	4	0,4,4	0.00	-	-		
14	UMQ	B	202	-	35,35,35	1.20	4 (11%)	46,46,46	2.58	15 (32%)
10	HEM	A	304	15,22	27,50,50	2.21	5 (18%)	17,82,82	1.38	3 (17%)
15	TDS	B	201	10	28,31,31	0.87	1 (3%)	35,40,40	2.25	10 (28%)
10	HEM	A	303	1	27,50,50	2.13	6 (22%)	17,82,82	1.53	4 (23%)
20	OCT	F	101	-	7,7,7	0.23	0	6,6,6	0.64	0
16	CLA	B	203	22	59,73,73	1.41	5 (8%)	67,113,113	1.55	12 (17%)
14	UMQ	A	308	-	35,35,35	1.24	5 (14%)	46,46,46	2.43	17 (36%)
13	8K6	A	307	-	17,17,17	0.20	0	16,16,16	0.56	0
10	HEM	A	302	1	27,50,50	2.10	5 (18%)	17,82,82	1.77	4 (23%)
10	HEM	C	302	3	27,50,50	2.09	5 (18%)	17,82,82	1.57	5 (29%)
15	TDS	A	309	-	28,31,31	0.87	1 (3%)	35,40,40	2.39	9 (25%)
18	SQD	D	201	-	52,53,54	3.25	19 (36%)	60,63,65	2.29	14 (23%)
17	7PH	C	301	-	31,31,37	0.95	2 (6%)	33,33,42	1.21	3 (9%)
21	BCR	G	101	-	41,41,41	2.28	24 (58%)	56,56,56	2.20	21 (37%)
11	OPC	B	204	-	53,53,54	1.05	3 (5%)	59,61,64	0.99	4 (6%)

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
11	OPC	A	305	-	-	32/57/57/60	-
12	MYS	A	306	-	-	3/12/12/12	-
19	FES	D	202	4	-	-	0/1/1/1
14	UMQ	B	202	-	2/2/10/10	12/20/60/60	0/2/2/2
10	HEM	A	304	15,22	-	2/6/54/54	-
13	8K6	A	307	-	-	5/15/15/15	-
10	HEM	A	303	1	-	2/6/54/54	-
20	OCT	F	101	-	-	0/5/5/5	-
16	CLA	B	203	22	3/3/22/25	20/37/135/135	-
14	UMQ	A	308	-	3/3/10/10	6/20/60/60	0/2/2/2
15	TDS	B	201	10	-	10/16/17/17	0/2/2/2
10	HEM	A	302	1	-	1/6/54/54	-
10	HEM	C	302	3	-	2/6/54/54	-
15	TDS	A	309	-	-	7/16/17/17	0/2/2/2
18	SQD	D	201	-	2/2/9/9	27/49/65/69	0/1/1/1
17	7PH	C	301	-	-	13/33/33/39	-
21	BCR	G	101	-	-	18/29/63/63	0/2/2/2
11	OPC	B	204	-	-	18/57/57/60	-

The worst 5 of 89 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
18	D	201	SQD	C2-C3	-10.69	1.34	1.52
18	D	201	SQD	O6-C44	9.42	1.61	1.43
18	D	201	SQD	C8-C7	7.81	1.73	1.50
16	B	203	CLA	C4B-NB	7.28	1.41	1.35
18	D	201	SQD	C4-C5	-6.74	1.38	1.53

The worst 5 of 124 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
18	D	201	SQD	C3-C4-C5	8.73	118.67	109.97
15	A	309	TDS	OAK-CAL-CAM	8.02	122.62	114.54
14	A	308	UMQ	O5-C5-C4	7.70	123.68	109.69
15	B	201	TDS	OAO-CAP-CAQ	7.14	120.38	111.91

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
14	B	202	UMQ	O1-C1-C2	7.03	126.31	108.10

5 of 10 chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
14	B	202	UMQ	C2'
14	B	202	UMQ	C2
16	B	203	CLA	NC
16	B	203	CLA	ND
16	B	203	CLA	NA

5 of 178 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
11	A	305	OPC	CAL-OAK-PAJ-OBH
11	A	305	OPC	CAL-OAK-PAJ-OAB
15	B	201	TDS	CAD-CAE-OAB-CAA
10	A	304	HEM	C2D-C3D-CAD-CBD
10	A	304	HEM	C4D-C3D-CAD-CBD

There are no ring outliers.

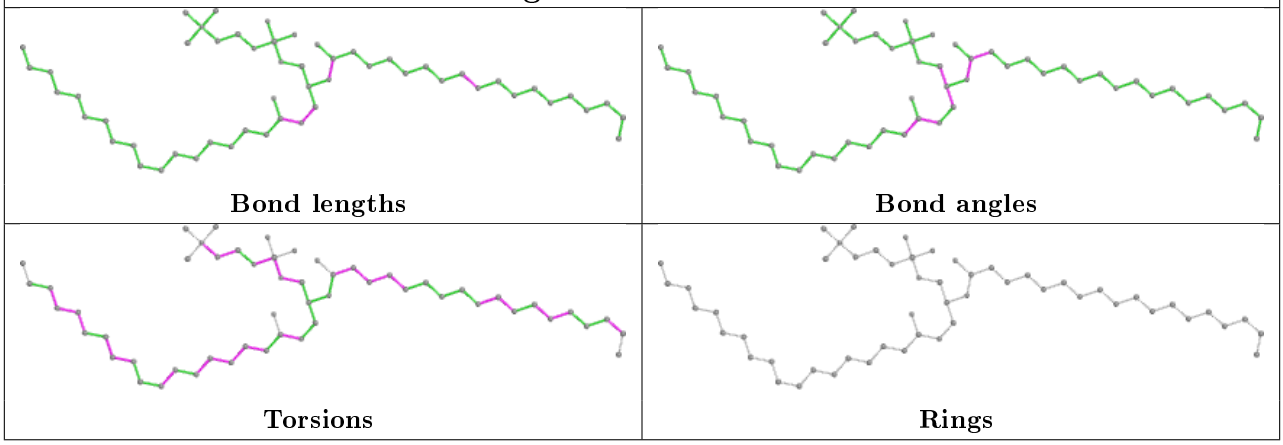
15 monomers are involved in 68 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
11	A	305	OPC	5	0
12	A	306	MYS	5	0
19	D	202	FES	1	0
14	B	202	UMQ	1	0
10	A	304	HEM	10	0
15	B	201	TDS	4	0
10	A	303	HEM	6	0
16	B	203	CLA	5	0
10	A	302	HEM	7	0
10	C	302	HEM	10	0
15	A	309	TDS	4	0
18	D	201	SQD	3	0
17	C	301	7PH	1	0
21	G	101	BCR	7	0
11	B	204	OPC	1	0

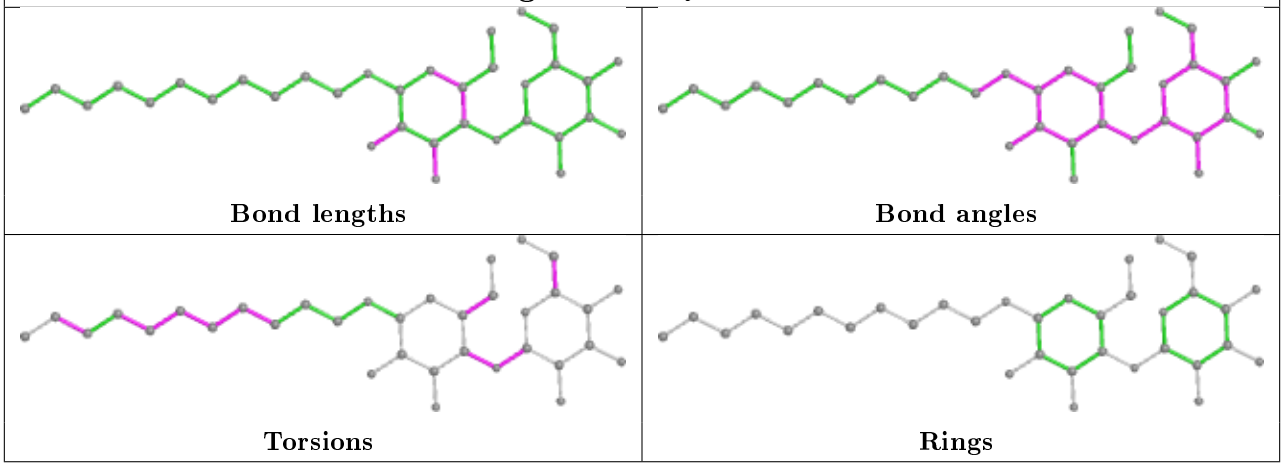
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In

addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

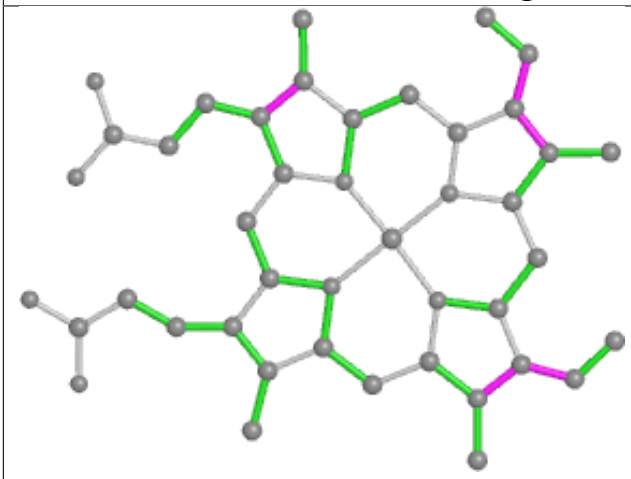
Ligand OPC A 305



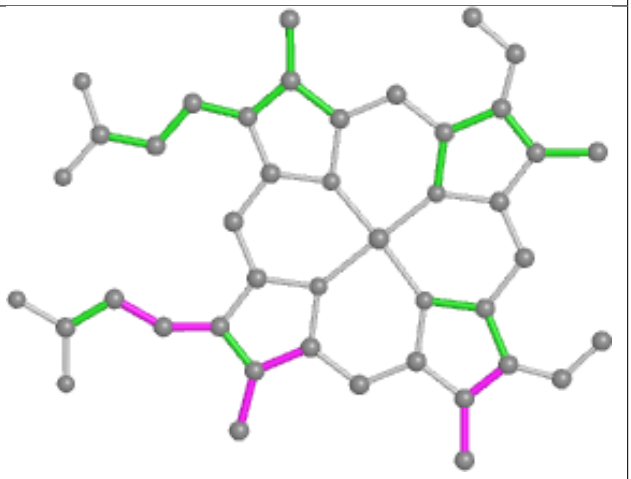
Ligand UMQ B 202



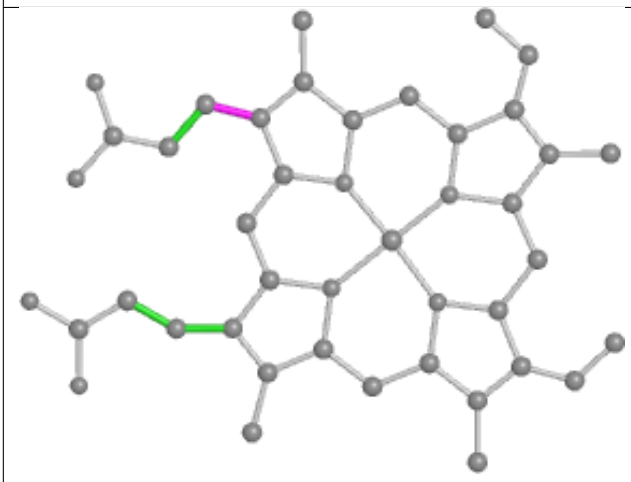
Ligand HEM A 304



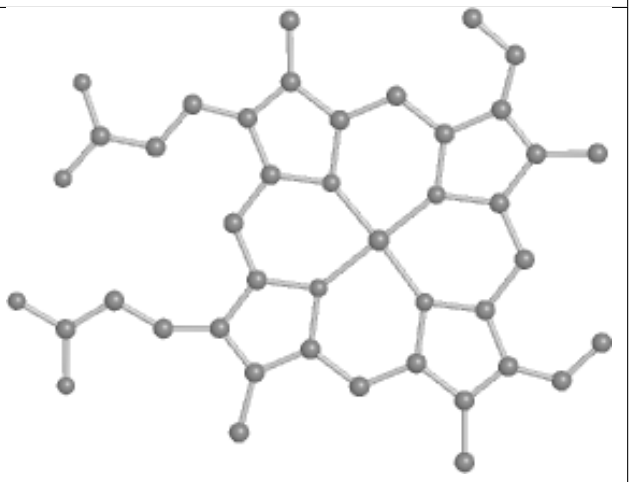
Bond lengths



Bond angles

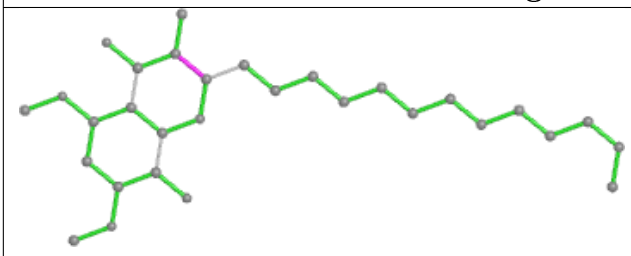


Torsions

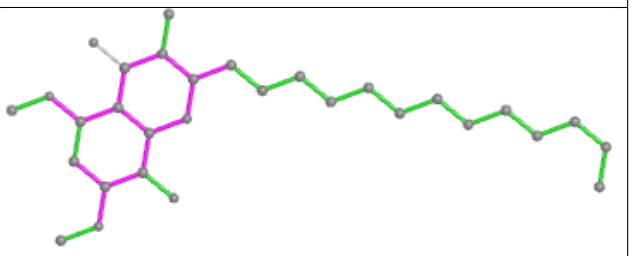


Rings

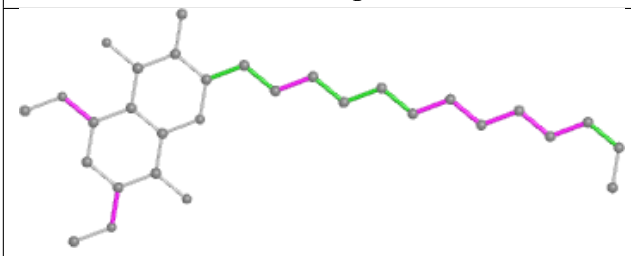
Ligand TDS B 201



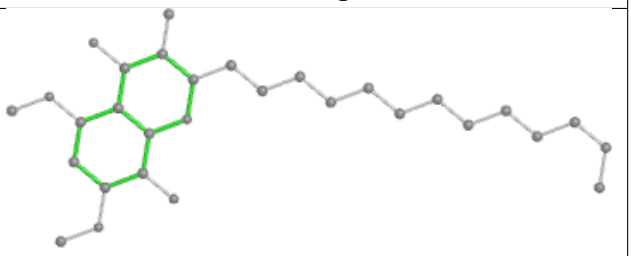
Bond lengths



Bond angles

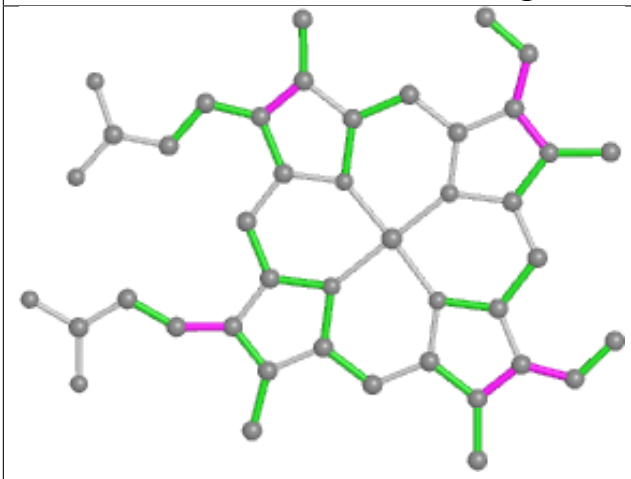


Torsions

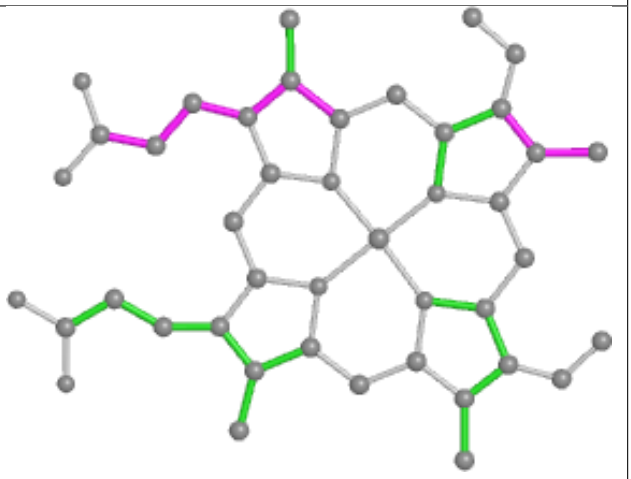


Rings

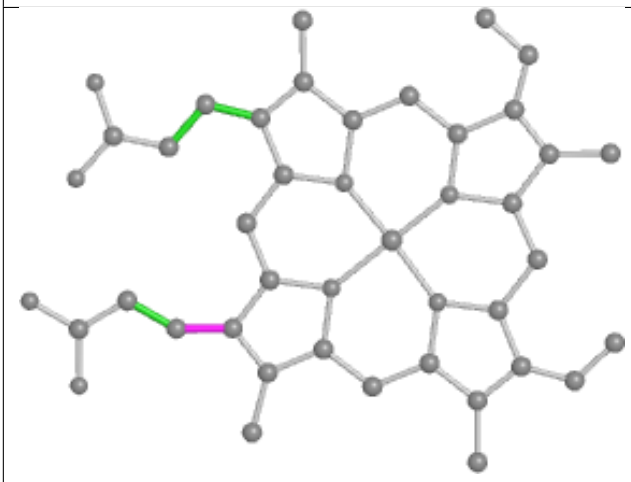
Ligand HEM A 303



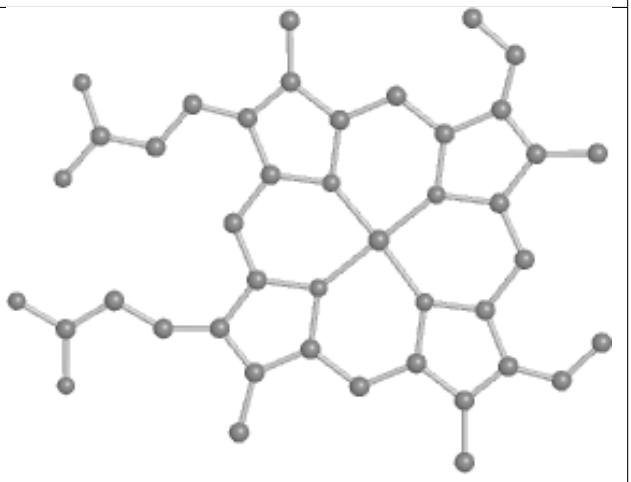
Bond lengths



Bond angles

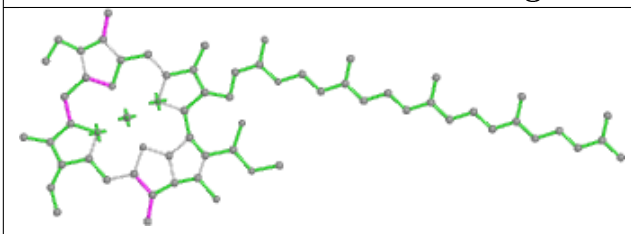


Torsions

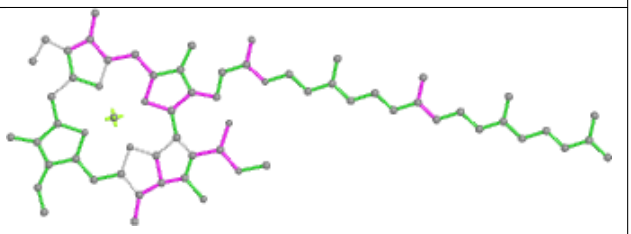


Rings

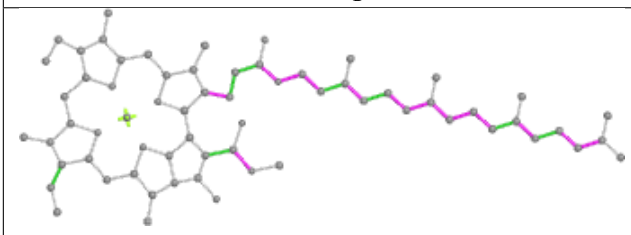
Ligand CLA B 203



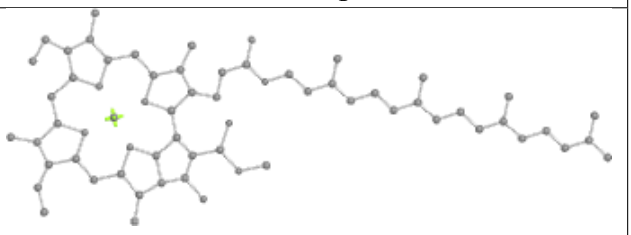
Bond lengths



Bond angles

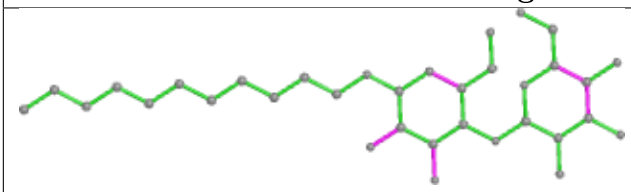


Torsions

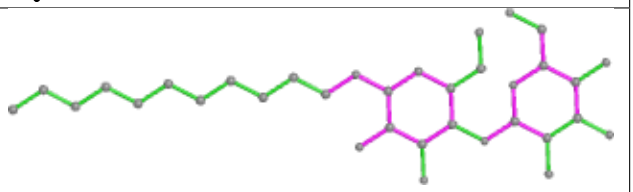


Rings

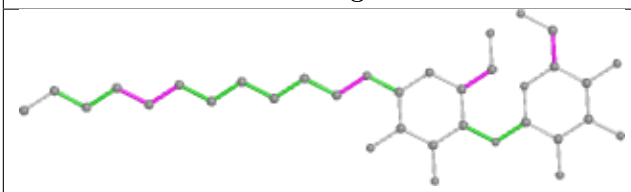
Ligand UMQ A 308



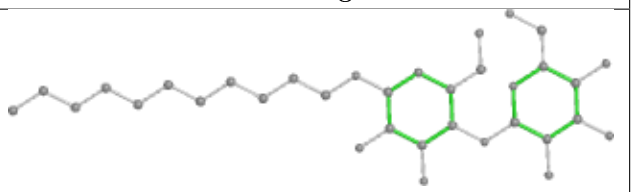
Bond lengths



Bond angles

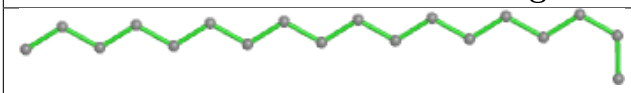


Torsions

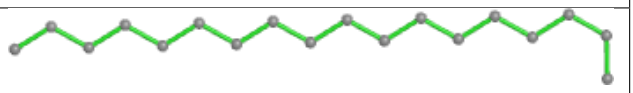


Rings

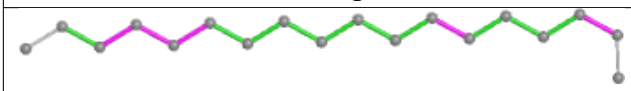
Ligand 8K6 A 307



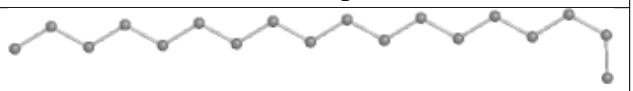
Bond lengths



Bond angles

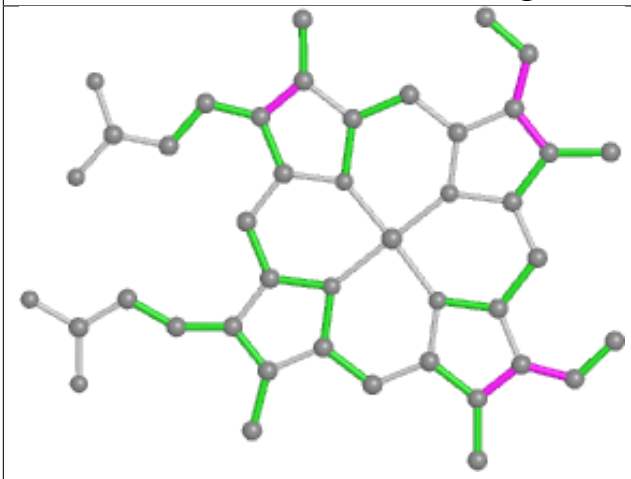


Torsions

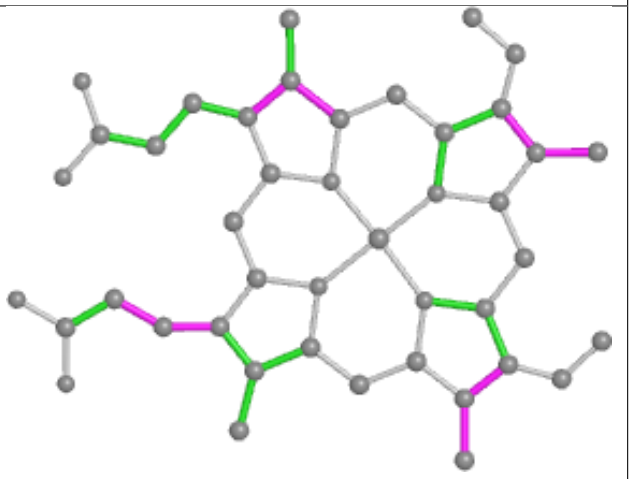


Rings

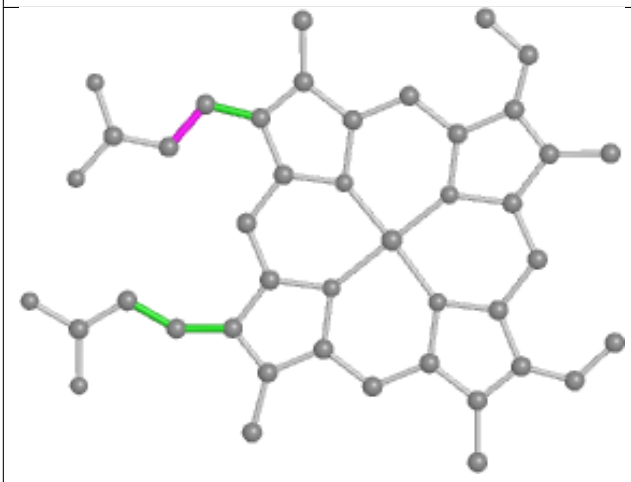
Ligand HEM A 302



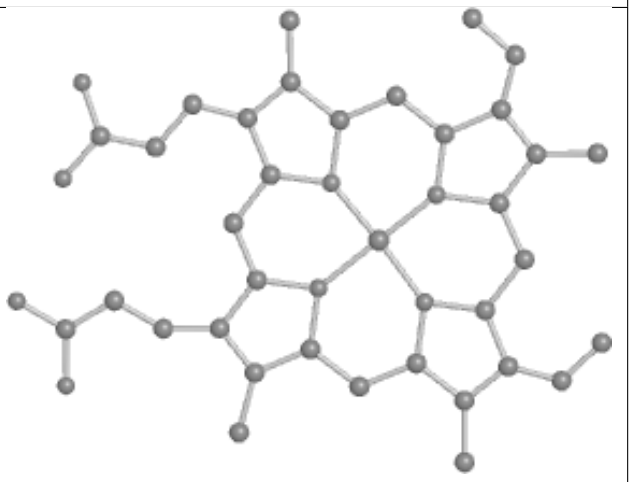
Bond lengths



Bond angles

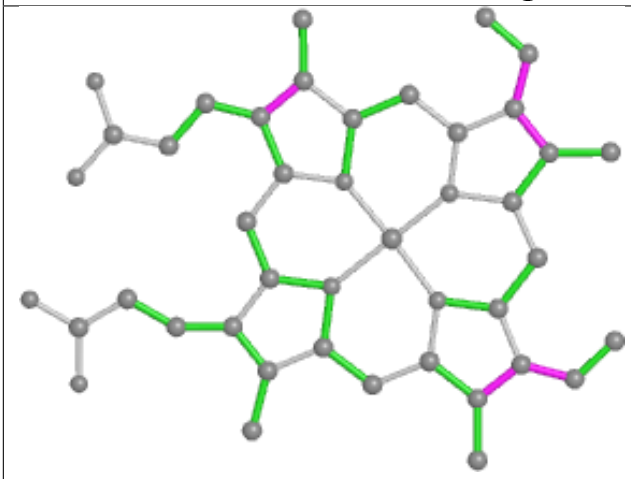


Torsions

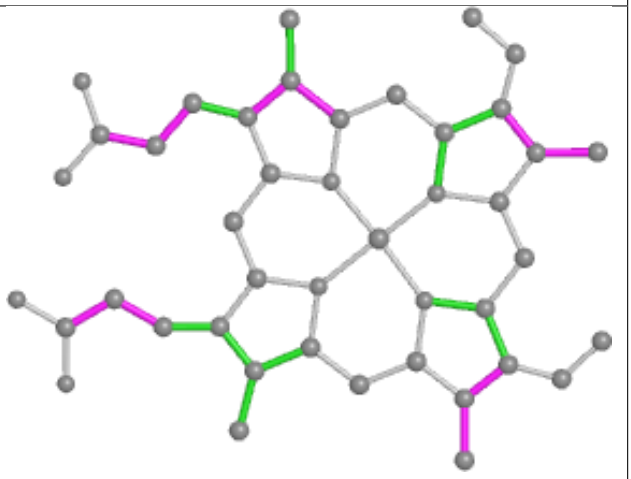


Rings

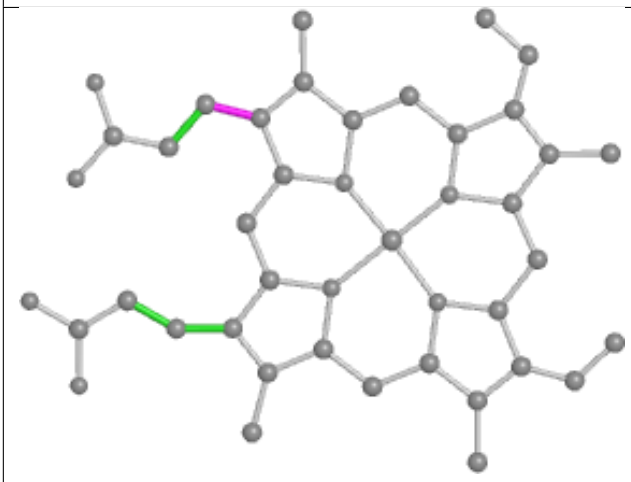
Ligand HEM C 302



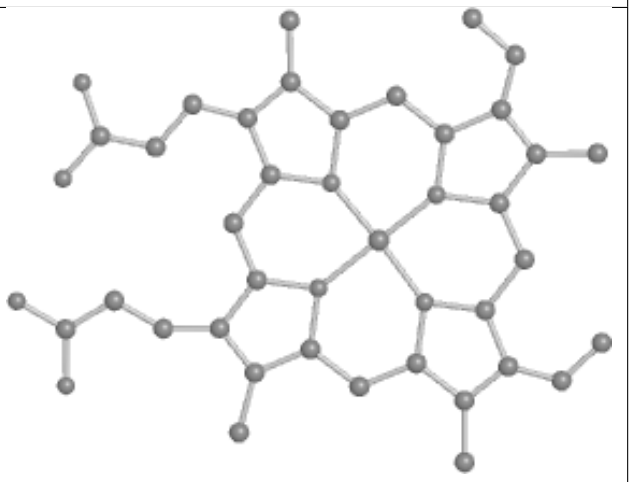
Bond lengths



Bond angles

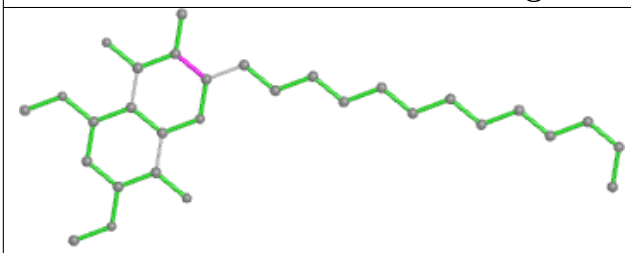


Torsions

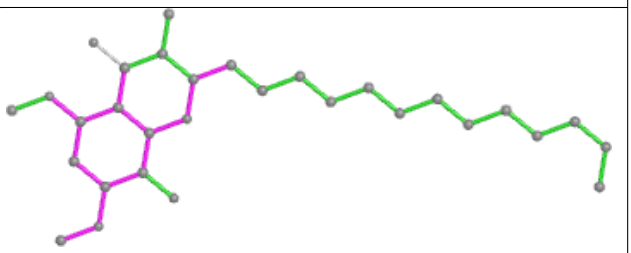


Rings

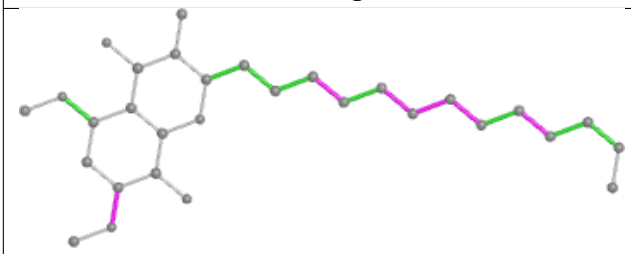
Ligand TDS A 309



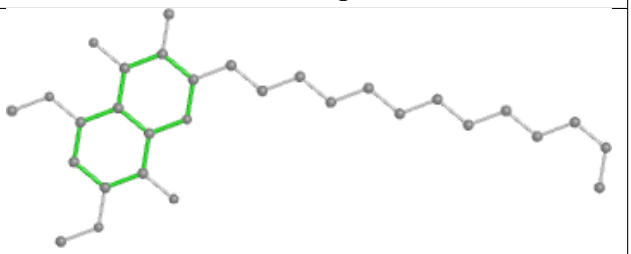
Bond lengths



Bond angles

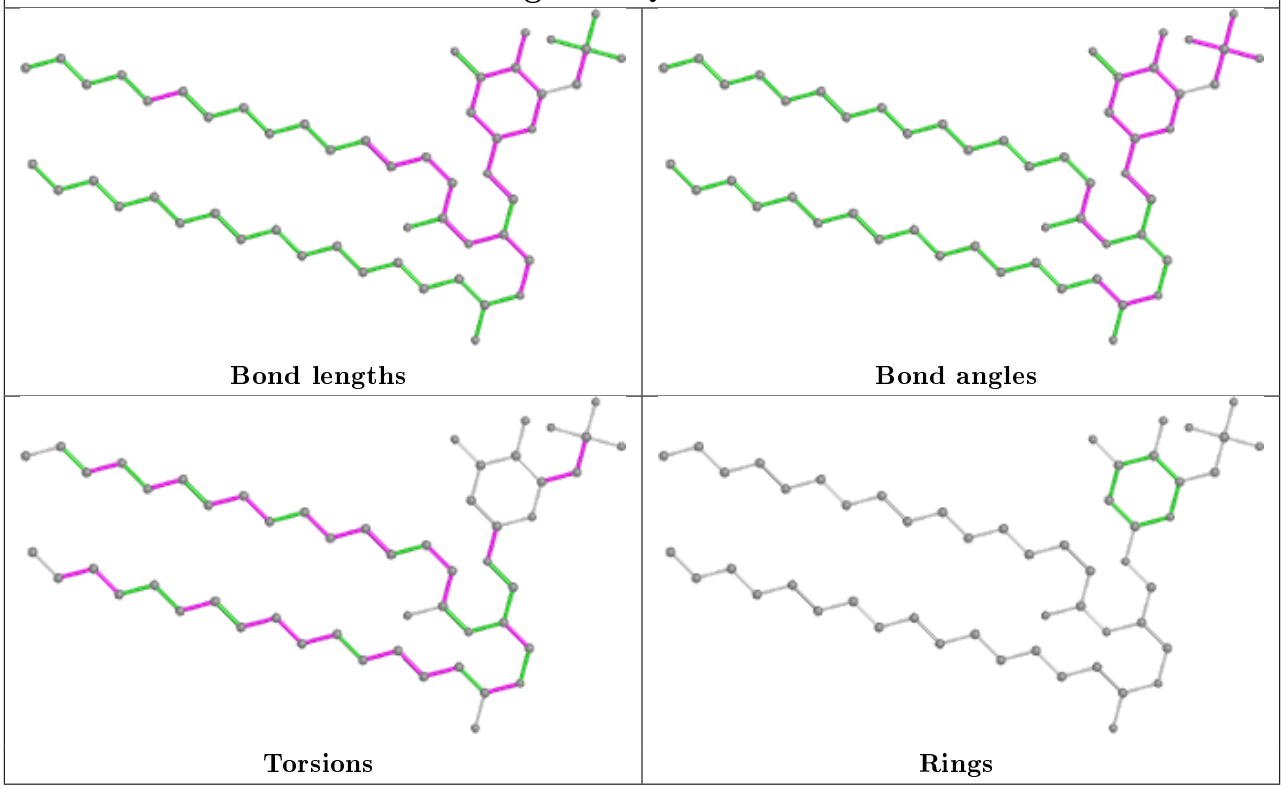


Torsions

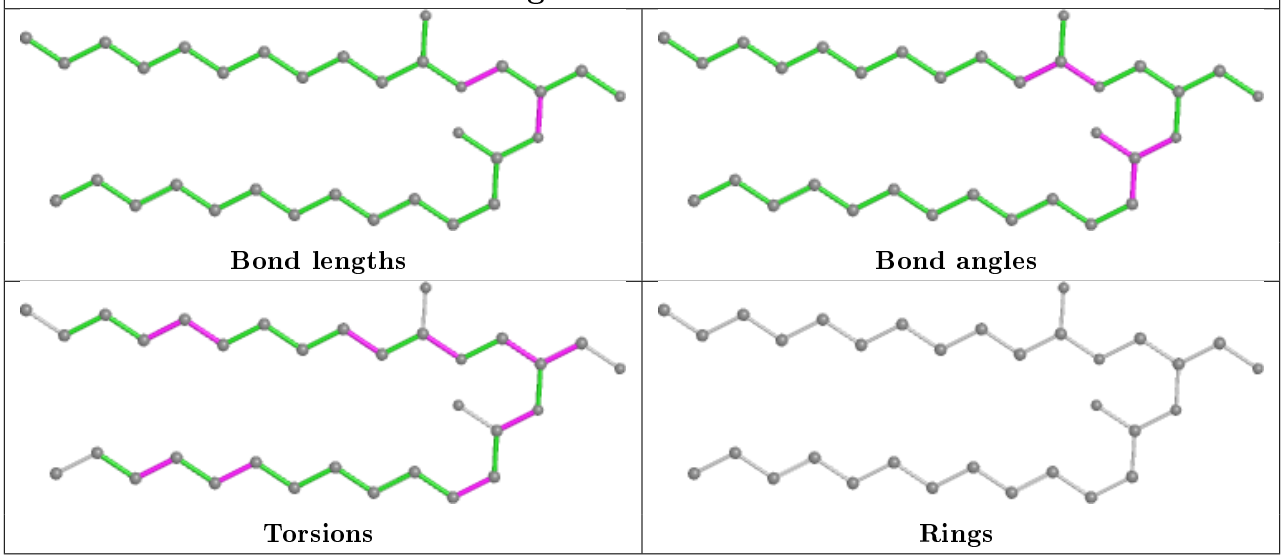


Rings

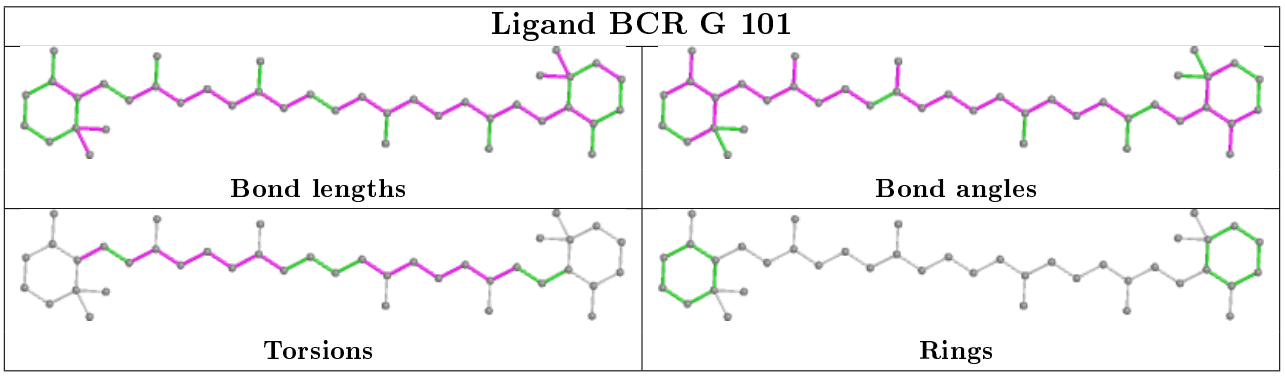
Ligand SQD D 201

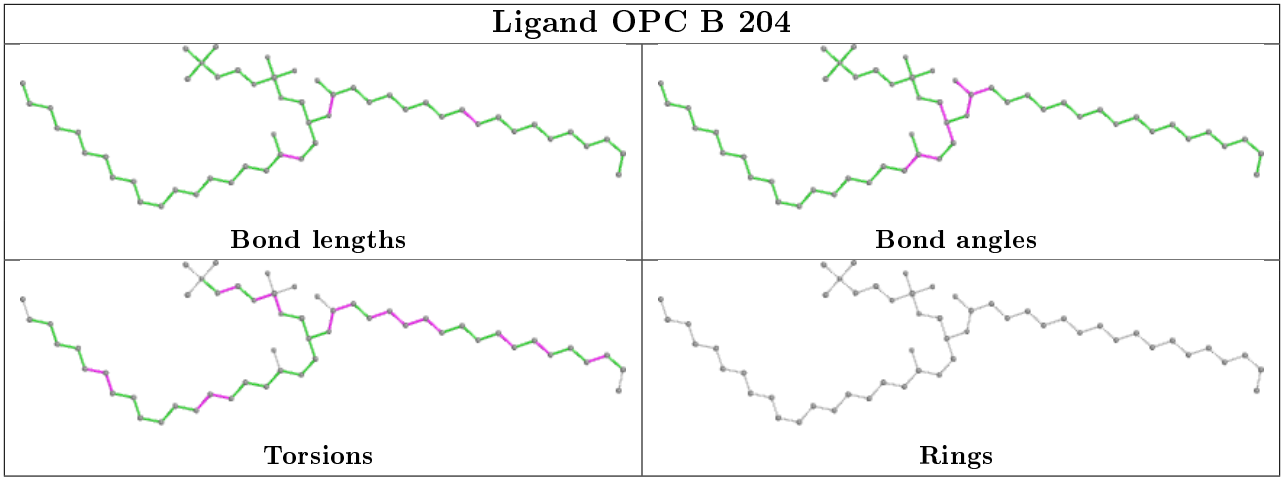


Ligand 7PH C 301



Ligand BCR G 101





5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

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^{th} 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^2)	Q<0.9
1	A	215/215 (100%)	-0.07	3 (1%) 75 55	49, 69, 105, 222	0
2	B	160/160 (100%)	-0.02	4 (2%) 57 33	61, 90, 131, 186	0
3	C	288/289 (99%)	0.18	25 (8%) 10 3	64, 100, 224, 256	1 (0%)
4	D	168/179 (93%)	0.64	23 (13%) 3 1	63, 138, 207, 236	0
5	E	32/32 (100%)	-0.02	0 100 100	86, 105, 135, 150	0
6	F	31/35 (88%)	-0.16	1 (3%) 47 25	78, 93, 150, 158	0
7	G	37/37 (100%)	0.02	1 (2%) 54 29	71, 87, 178, 205	0
8	H	29/29 (100%)	0.20	0 100 100	76, 85, 103, 141	0
All	All	960/976 (98%)	0.15	57 (5%) 22 10	49, 93, 201, 256	1 (0%)

The worst 5 of 57 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	D	95	SER	8.5
4	D	96	LYS	7.1
4	D	50	VAL	6.2
1	A	1	MET	5.6
4	D	157	ASP	5.4

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

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

6.3 Carbohydrates

There are no carbohydrates in this entry.

6.4 Ligands

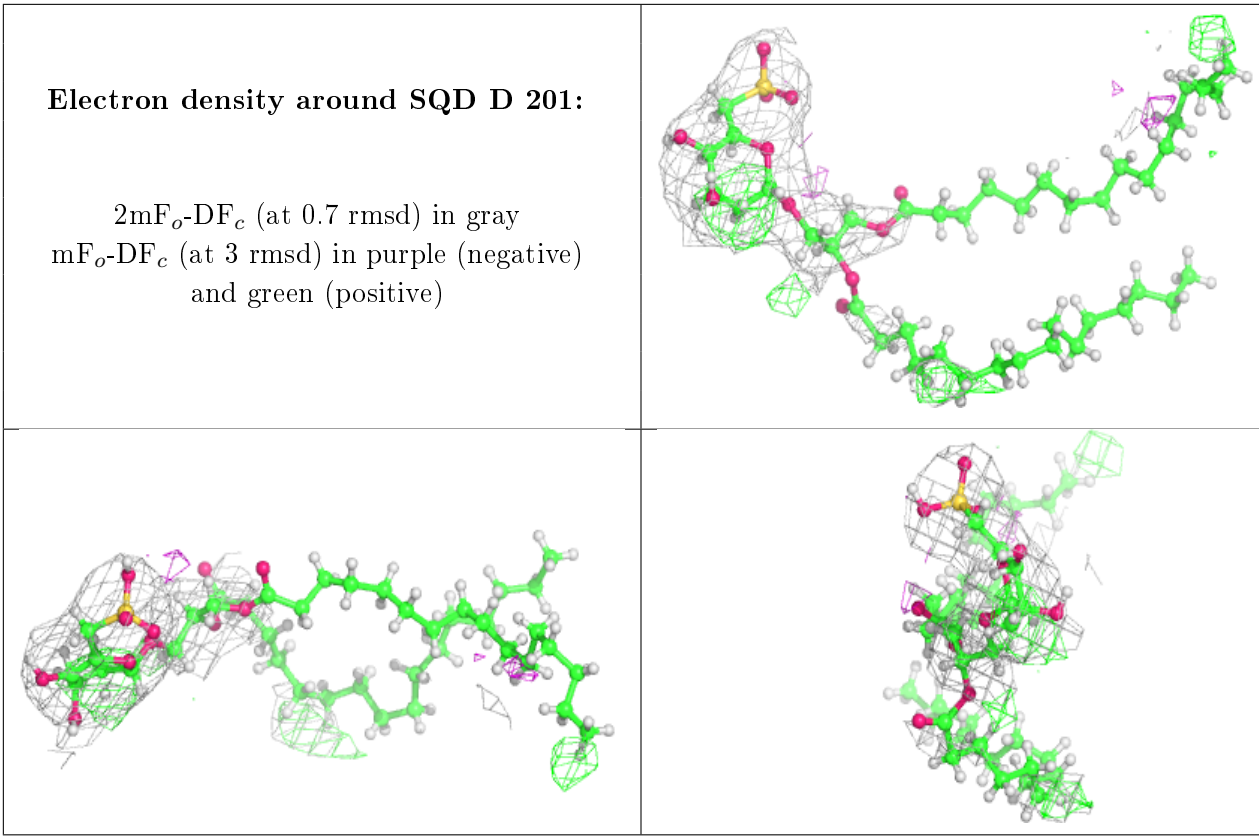
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
18	SQD	D	201	53/54	0.61	0.61	0,163,195,201	1
21	BCR	G	101	40/40	0.63	0.52	61,105,223,230	0
14	UMQ	B	202	34/34	0.65	0.49	104,171,232,247	0
13	8K6	A	307	18/18	0.67	0.37	83,115,138,142	0
17	7PH	C	301	32/38	0.70	0.46	73,108,165,198	0
11	OPC	A	305	54/55	0.70	0.50	77,127,224,237	0
20	OCT	F	101	8/8	0.71	0.53	98,124,131,132	0
14	UMQ	A	308	34/34	0.76	0.41	108,160,202,212	0
12	MYS	A	306	15/15	0.83	0.34	72,101,126,129	0
15	TDS	A	309	30/30	0.87	0.28	90,119,165,167	0
16	CLA	B	203	65/65	0.87	0.34	73,104,147,157	0
11	OPC	B	204	54/55	0.87	0.36	91,133,171,184	0
15	TDS	B	201	30/30	0.93	0.25	73,115,141,159	0
10	HEM	C	302	43/43	0.96	0.23	71,95,139,149	0
10	HEM	A	304	43/43	0.97	0.25	61,90,109,123	0
10	HEM	A	302	43/43	0.98	0.26	44,66,93,112	0
10	HEM	A	303	43/43	0.98	0.26	49,68,84,87	0
19	FES	D	202	4/4	0.99	0.12	80,95,101,103	0
9	CD	A	301	1/1	1.00	0.11	97,97,97,97	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.

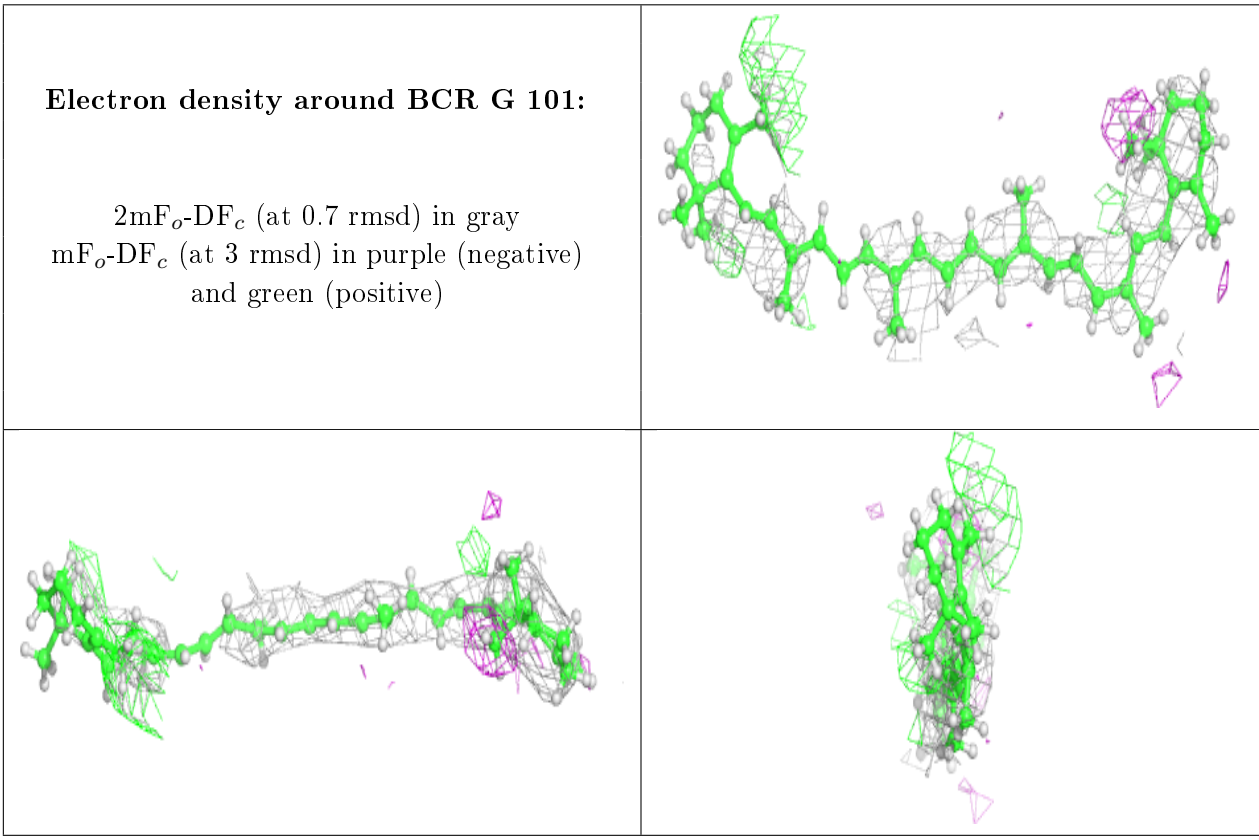
Electron density around SQD D 201:

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



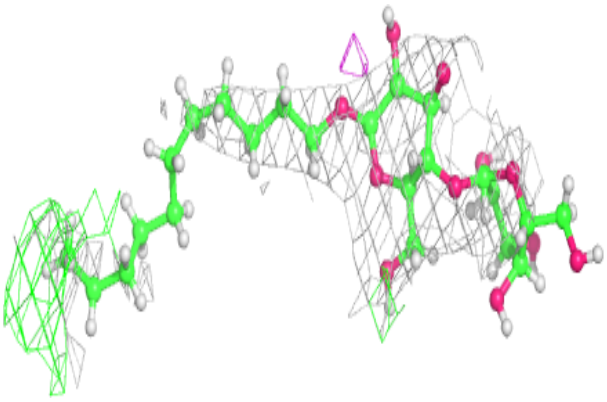
Electron density around BCR G 101:

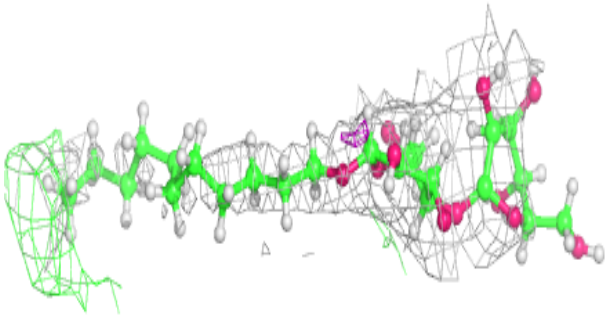
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

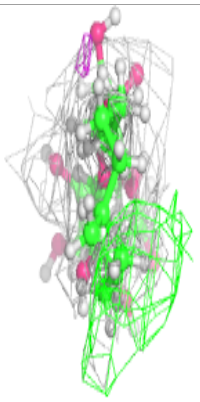


Electron density around UMQ B 202:

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

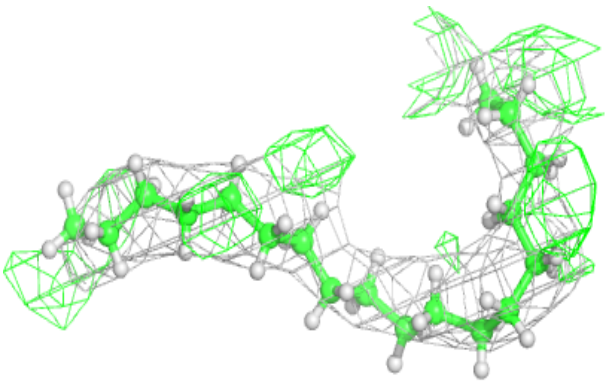


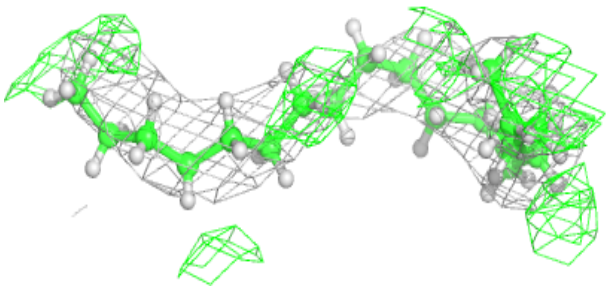


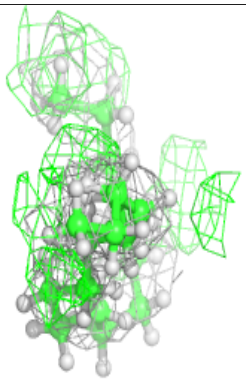


Electron density around 8K6 A 307:

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

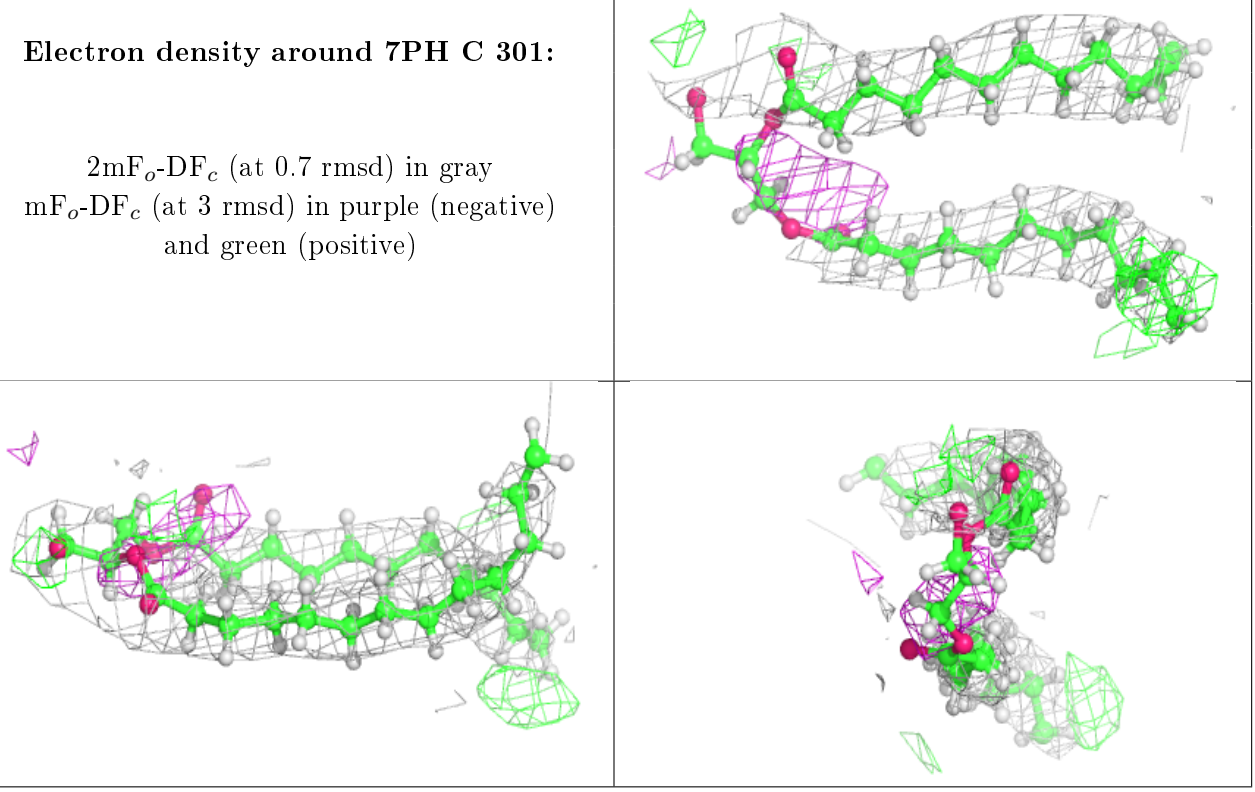






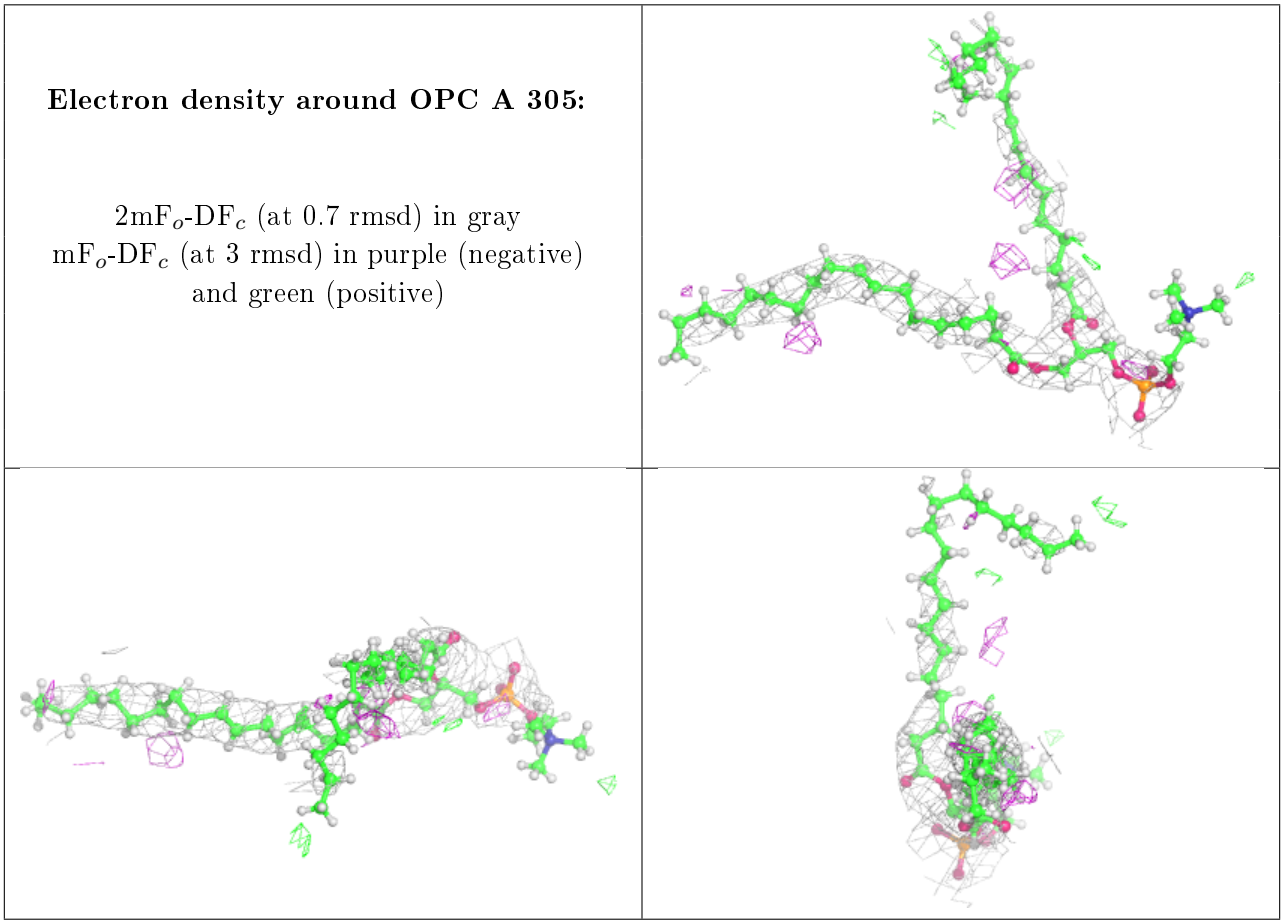
Electron density around 7PH C 301:

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



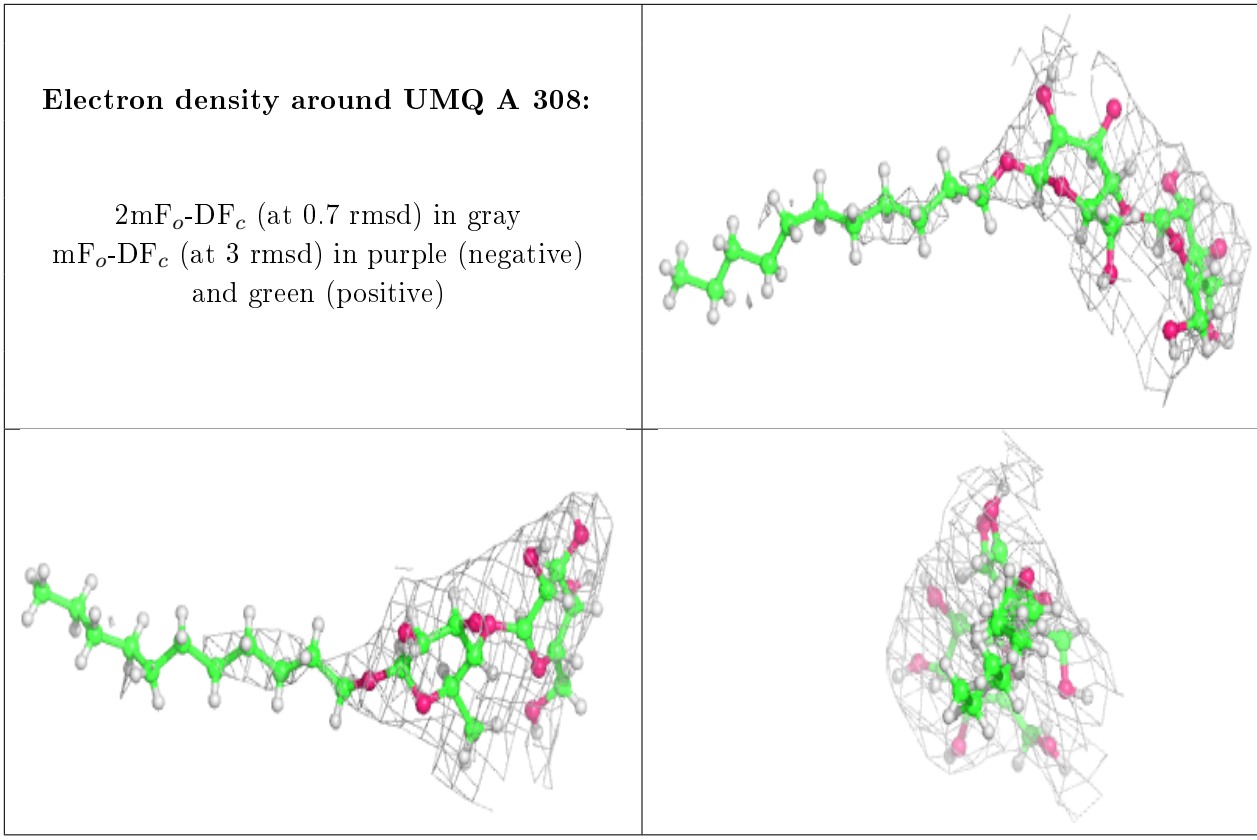
Electron density around OPC A 305:

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



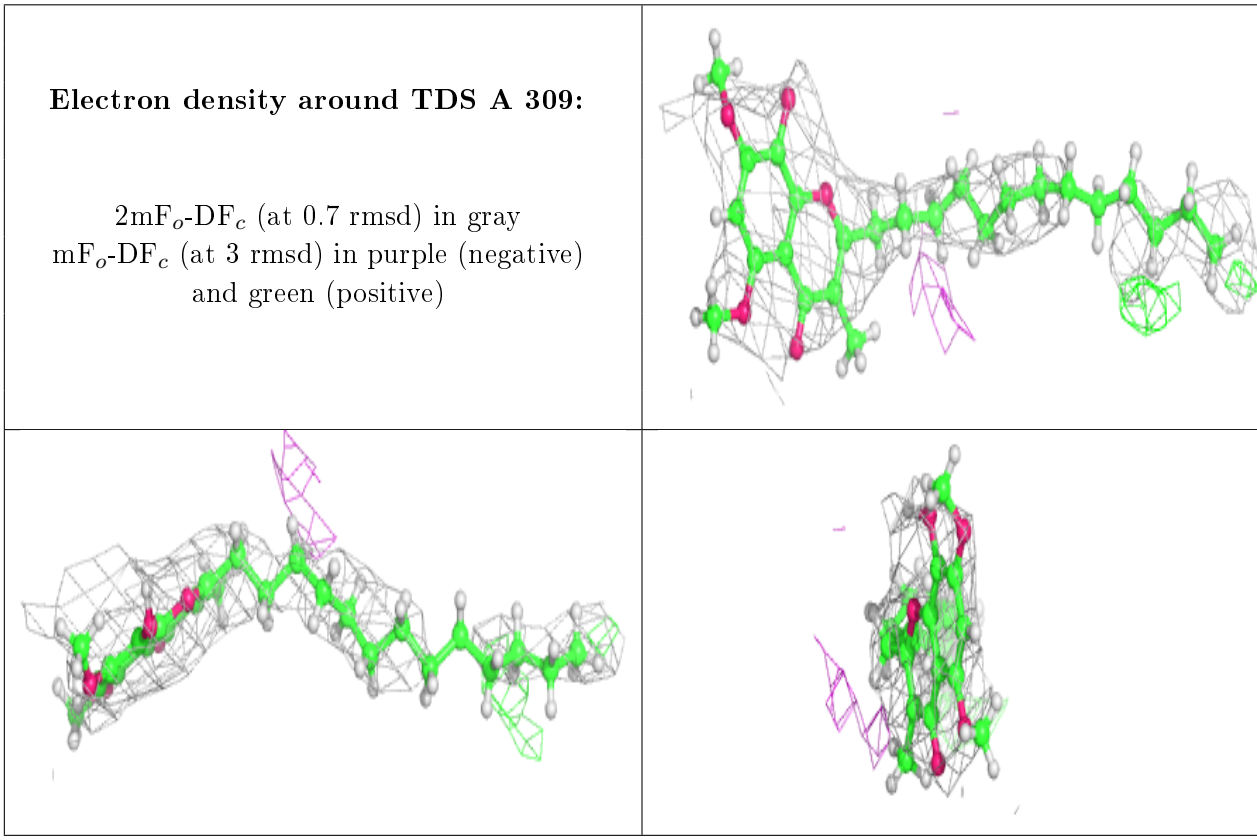
Electron density around UMQ A 308:

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



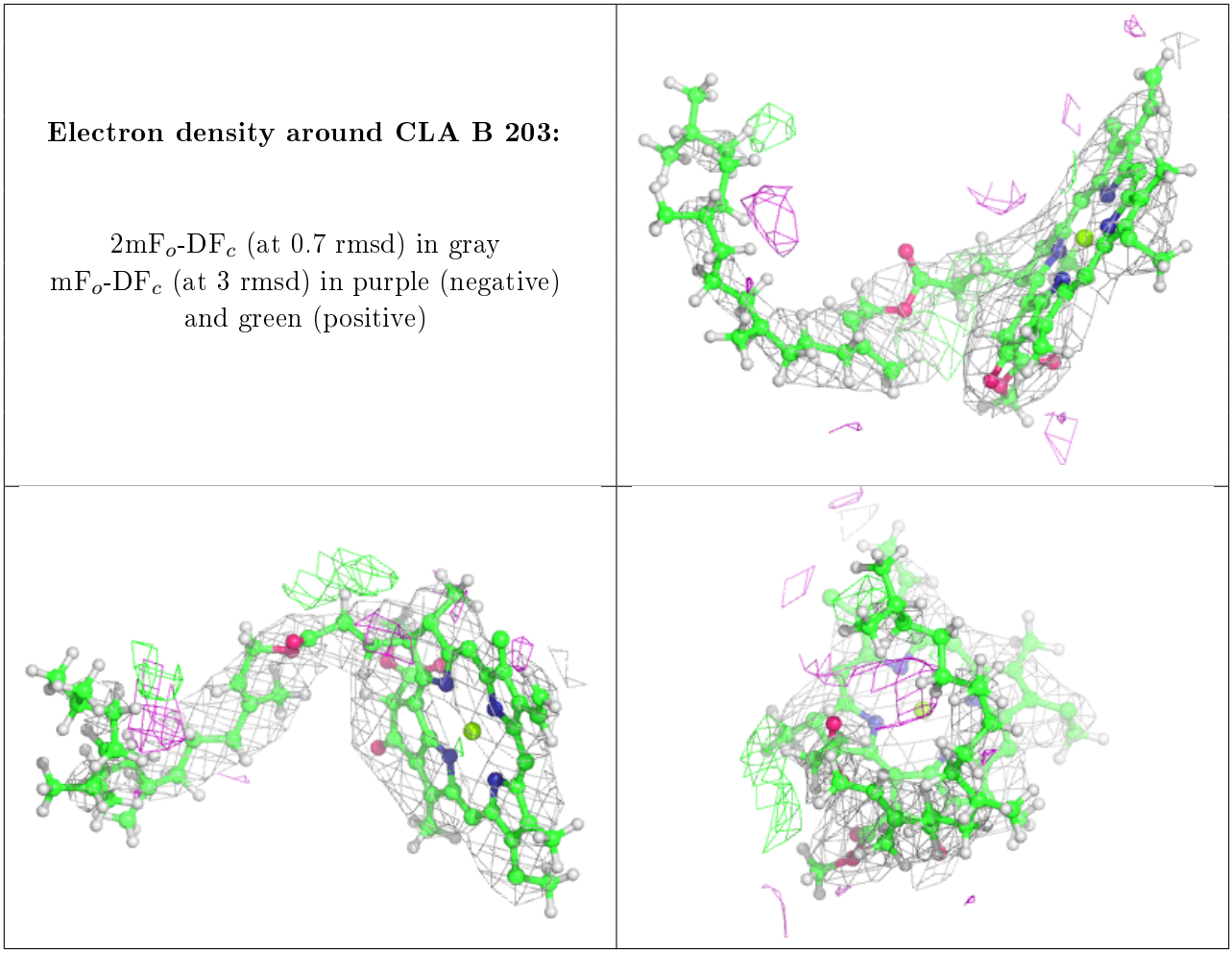
Electron density around TDS A 309:

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



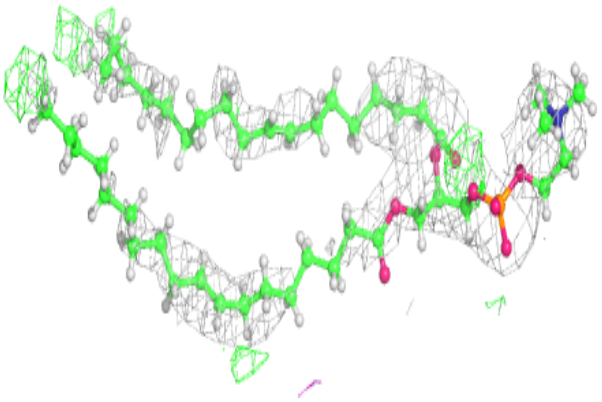
Electron density around CLA B 203:

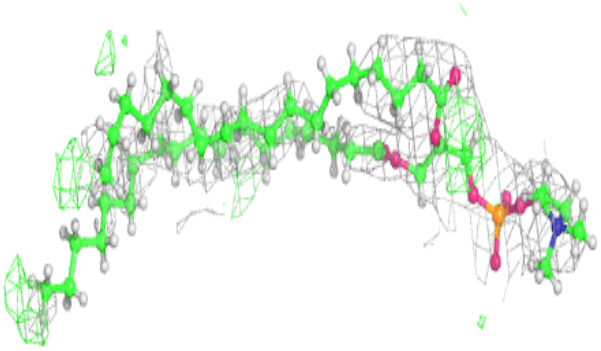
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

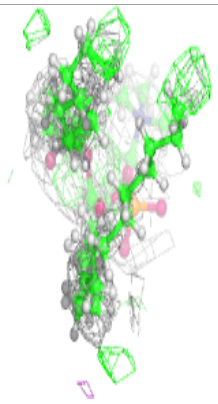


Electron density around OPC B 204:

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

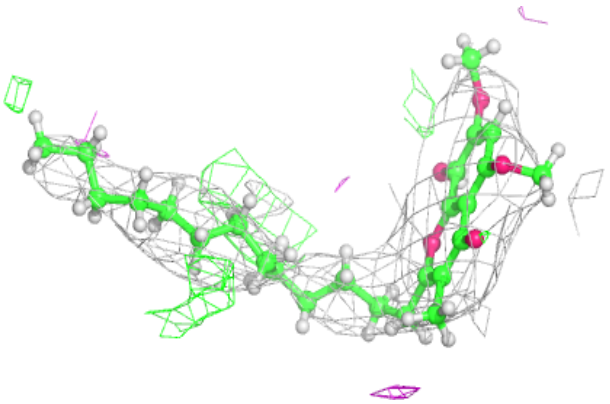


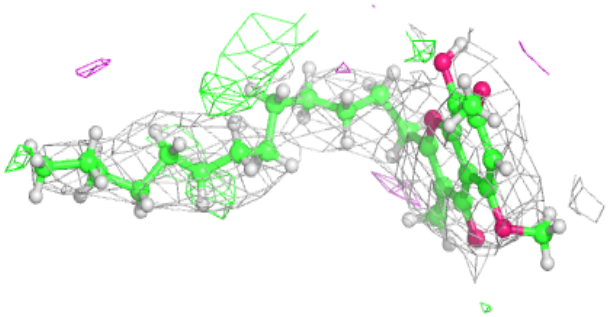


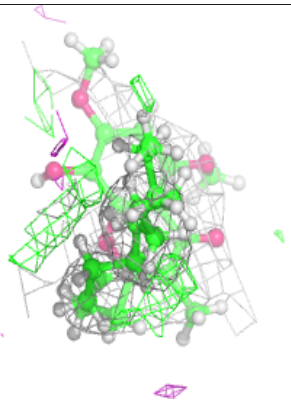


Electron density around TDS B 201:

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

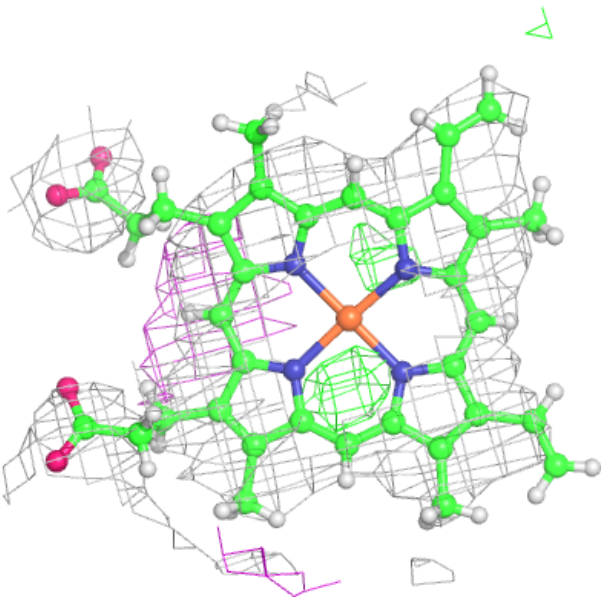


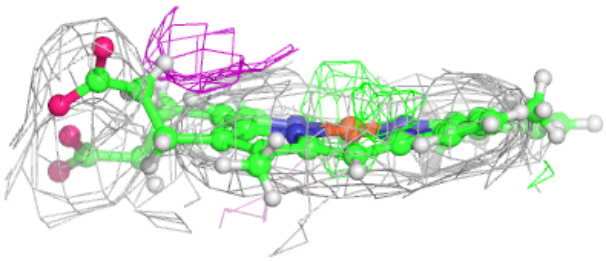


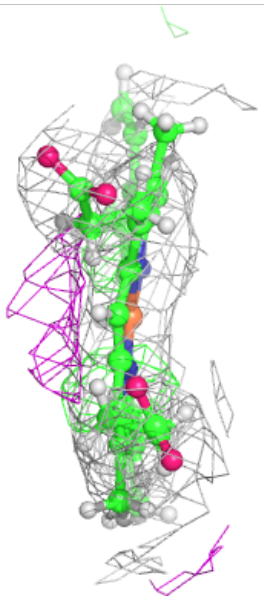


Electron density around HEM C 302:

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

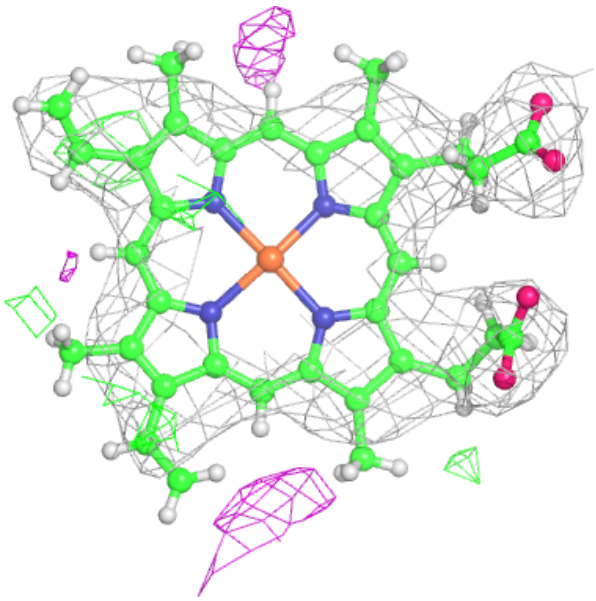


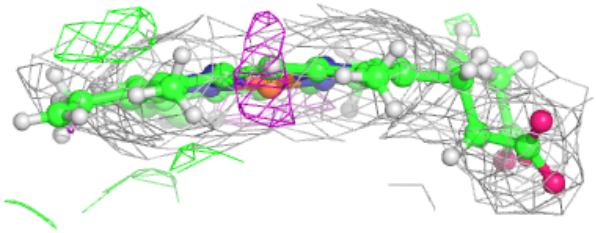


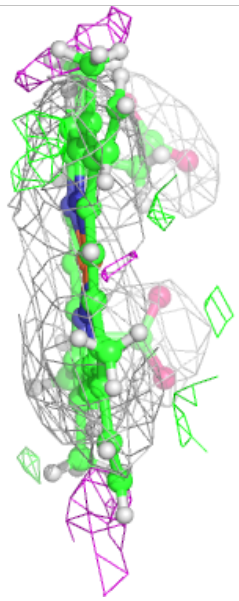


Electron density around HEM A 304:

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

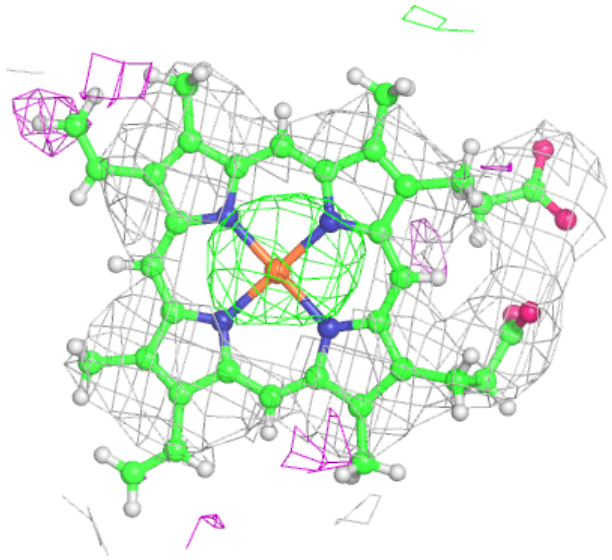


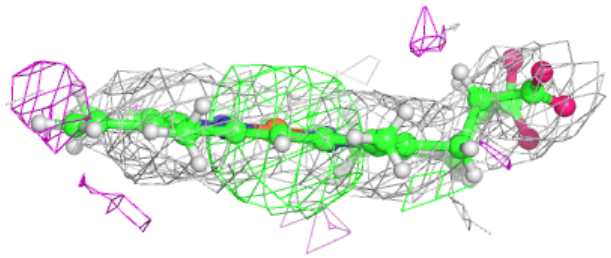


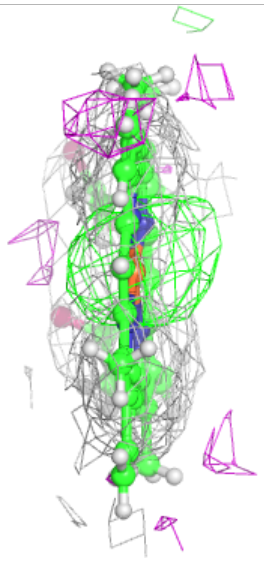


Electron density around HEM A 302:

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

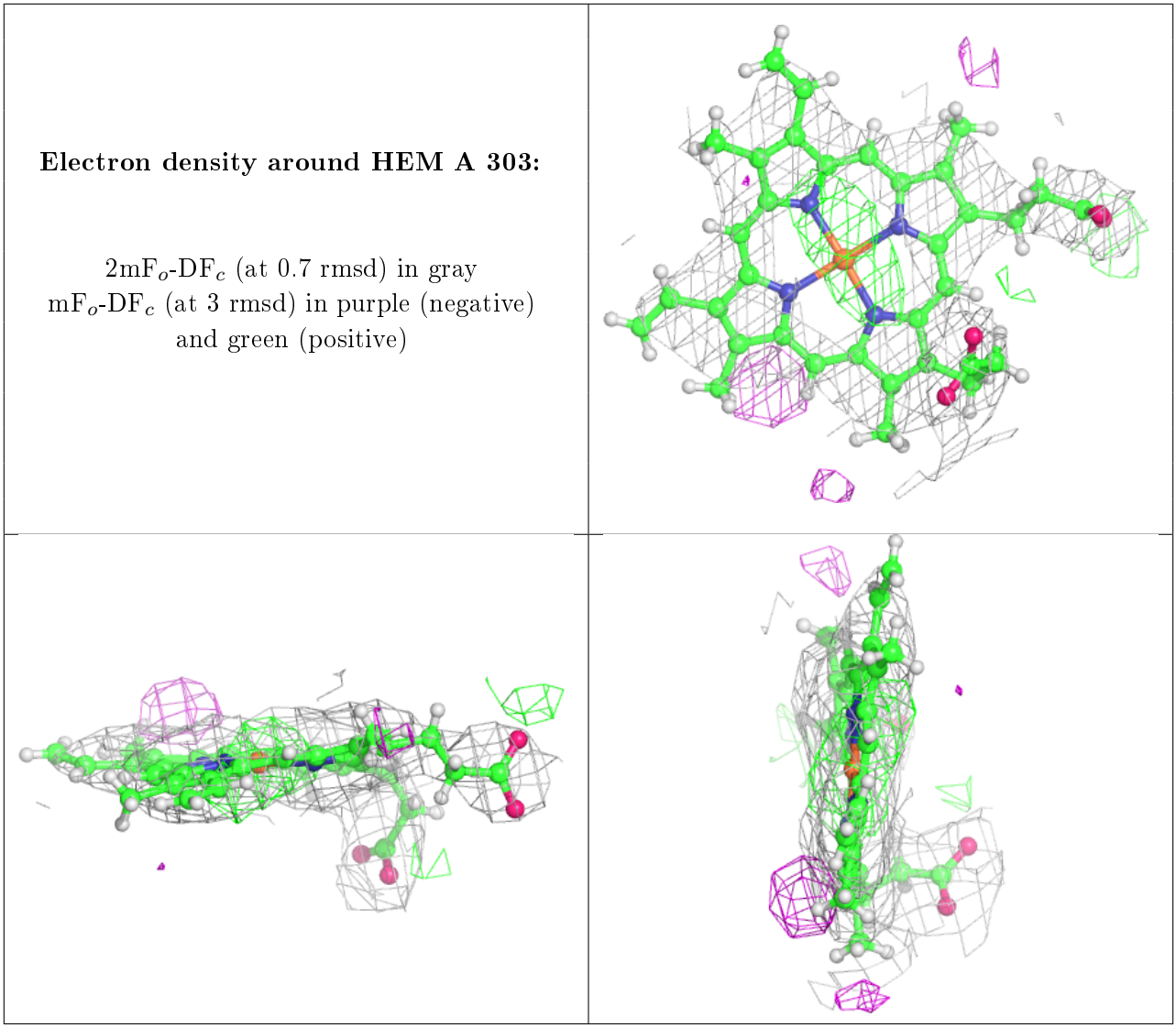






Electron density around HEM A 303:

$2mF_o$-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.