wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 20, 2023 – 12:47 PM EDT

PDB ID	:	5E58
Title	:	Crystal Structure Of Cytochrome P450 2B35 from Desert Woodrat Neotoma Lepida in complex with 4-(4-chlorophenyl)imidazole
Authors	:	Shah, M.B.; Stout, C.D.; Halpert, J.R.
Deposited on	:	2015-10-08
Resolution	:	2.40 Å(reported)

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

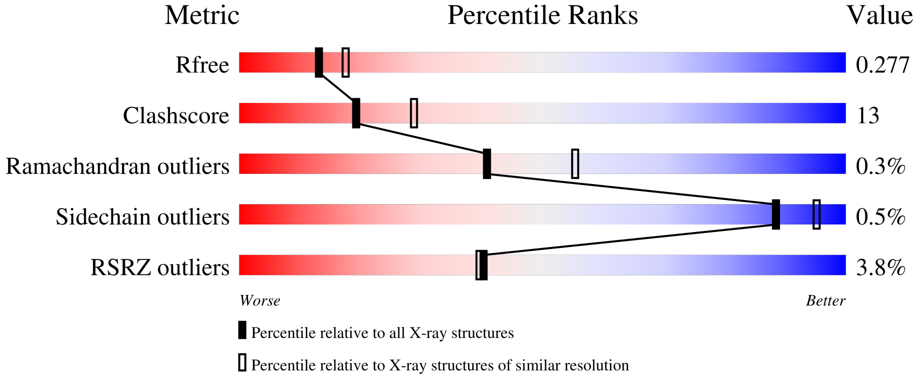
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

1 Overall quality at a glance ⓘ

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

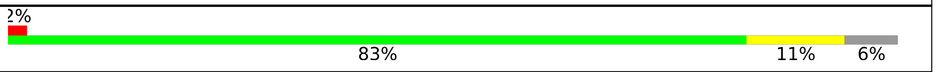
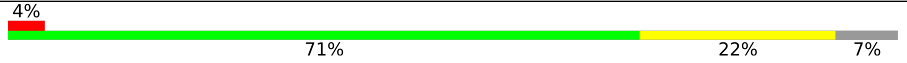
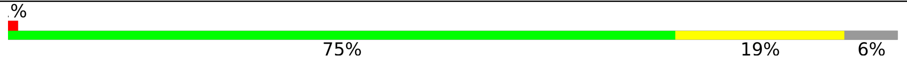
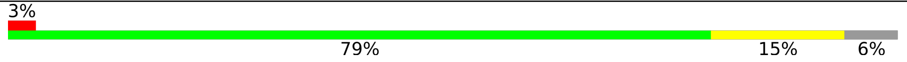
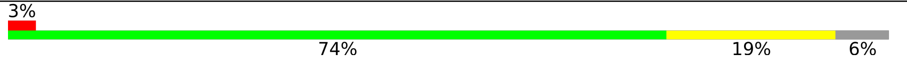
The reported resolution of this entry is 2.40 Å.

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.



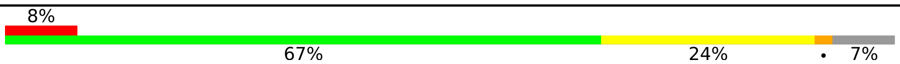
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=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	493	
1	B	493	
1	C	493	
1	D	493	
1	E	493	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	493	
2	G	2	

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
2	GLC	G	1	-	X	-	-
2	FRU	G	2	-	-	X	-
4	CPZ	A	502	-	-	X	-
4	CPZ	B	502	-	-	X	-
4	CPZ	B	503	-	-	X	-
4	CPZ	C	502	-	-	X	-
4	CPZ	D	502	-	-	X	-
4	CPZ	E	502	-	-	X	-
4	CPZ	F	502	-	-	X	-
4	CPZ	F	503	-	-	X	-

2 Entry composition ⓘ

There are 7 unique types of molecules in this entry. The entry contains 23024 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.

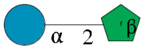
- Molecule 1 is a protein called Cytochrome P450 family 2 subfamily B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	464	Total 3727	C 2396	N 638	O 677	S 16	0	0	0
1	B	458	Total 3602	C 2321	N 607	O 659	S 15	0	0	0
1	C	465	Total 3685	C 2371	N 631	O 667	S 16	0	0	0
1	D	463	Total 3665	C 2360	N 623	O 666	S 16	0	0	0
1	E	462	Total 3661	C 2352	N 626	O 667	S 16	0	0	0
1	F	458	Total 3558	C 2284	N 608	O 651	S 15	0	1	0

There are 12 discrepancies between the modelled and reference sequences:

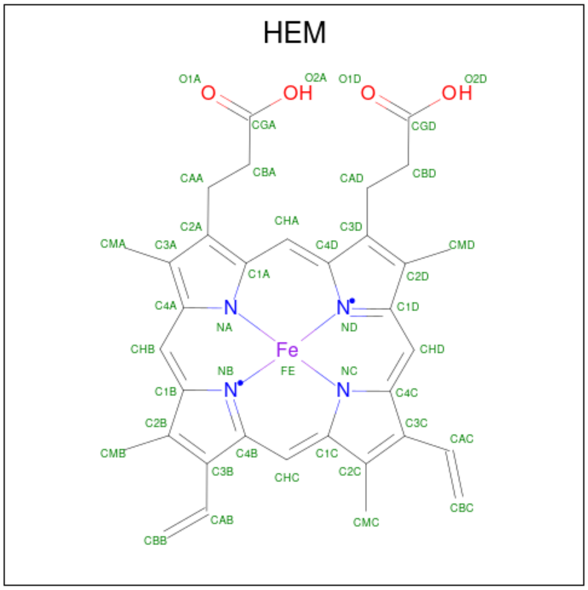
Chain	Residue	Modelled	Actual	Comment	Reference
A	492	HIS	-	expression tag	UNP J9JD66
A	493	HIS	-	expression tag	UNP J9JD66
B	492	HIS	-	expression tag	UNP J9JD66
B	493	HIS	-	expression tag	UNP J9JD66
C	492	HIS	-	expression tag	UNP J9JD66
C	493	HIS	-	expression tag	UNP J9JD66
D	492	HIS	-	expression tag	UNP J9JD66
D	493	HIS	-	expression tag	UNP J9JD66
E	492	HIS	-	expression tag	UNP J9JD66
E	493	HIS	-	expression tag	UNP J9JD66
F	492	HIS	-	expression tag	UNP J9JD66
F	493	HIS	-	expression tag	UNP J9JD66

- Molecule 2 is an oligosaccharide called beta-D-fructofuranose-(2-1)-alpha-D-glucopyranose.



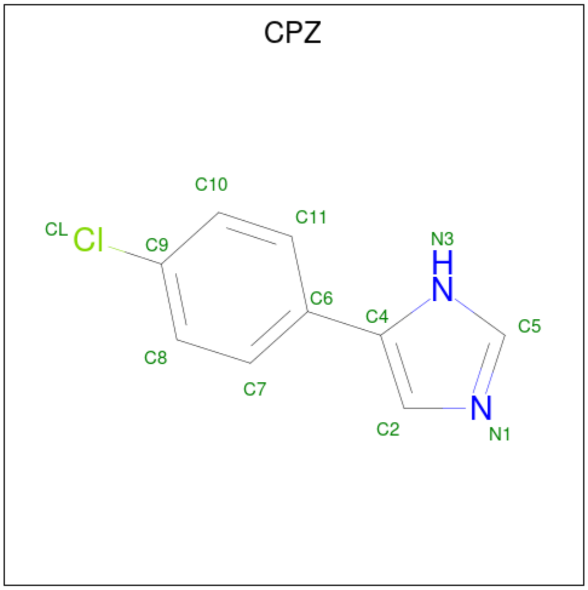
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
			Total	C	O			
2	G	2	23	12	11	0	0	0

- Molecule 3 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (formula: $C_{34}H_{32}FeN_4O_4$).



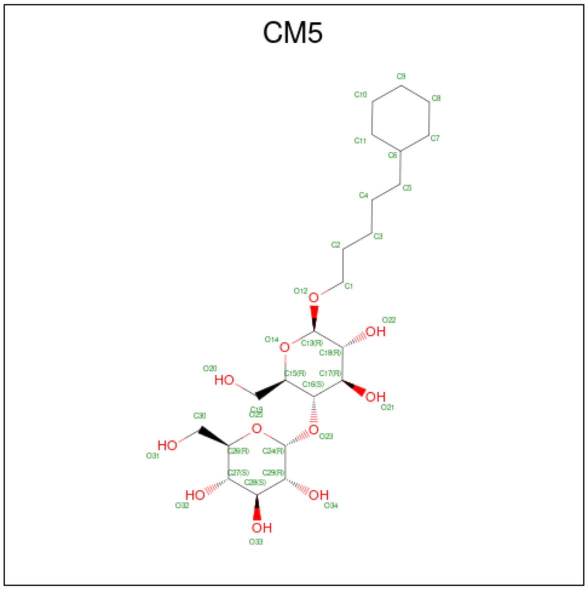
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	Fe	N	O		
3	A	1	43	34	1	4	4	0	0
3	B	1	43	34	1	4	4	0	0
3	C	1	43	34	1	4	4	0	0
3	D	1	43	34	1	4	4	0	0
3	E	1	43	34	1	4	4	0	0
3	F	1	43	34	1	4	4	0	0

- Molecule 4 is 4-(4-CHLOROPHENYL)IMIDAZOLE (three-letter code: CPZ) (formula: $C_9H_7ClN_2$).



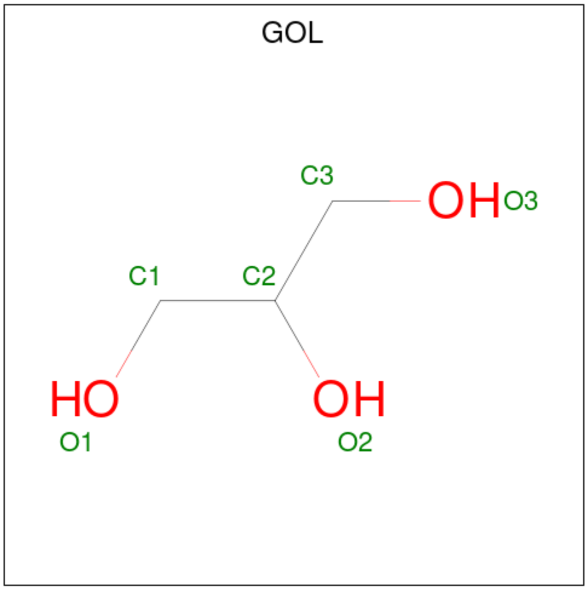
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C Cl N 12 9 1 2	0	0
4	A	1	Total C Cl N 12 9 1 2	0	0
4	B	1	Total C Cl N 12 9 1 2	0	0
4	B	1	Total C Cl N 12 9 1 2	0	0
4	C	1	Total C Cl N 12 9 1 2	0	0
4	C	1	Total C Cl N 12 9 1 2	0	0
4	D	1	Total C Cl N 12 9 1 2	0	0
4	D	1	Total C Cl N 12 9 1 2	0	0
4	E	1	Total C Cl N 12 9 1 2	0	0
4	E	1	Total C Cl N 12 9 1 2	0	0
4	F	1	Total C Cl N 12 9 1 2	0	0
4	F	1	Total C Cl N 12 9 1 2	0	0

- Molecule 5 is 5-CYCLOHEXYL-1-PENTYL-BETA-D-MALTOSIDE (three-letter code: CM5) (formula: $C_{23}H_{42}O_{11}$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C 6 6	0	0
5	E	1	Total C O 12 11 1	0	0

- Molecule 6 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total C O 6 3 3	0	0
6	B	1	Total C O 6 3 3	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	C	1	Total C O 6 3 3	0	0
6	E	1	Total C O 6 3 3	0	0

- Molecule 7 is water.

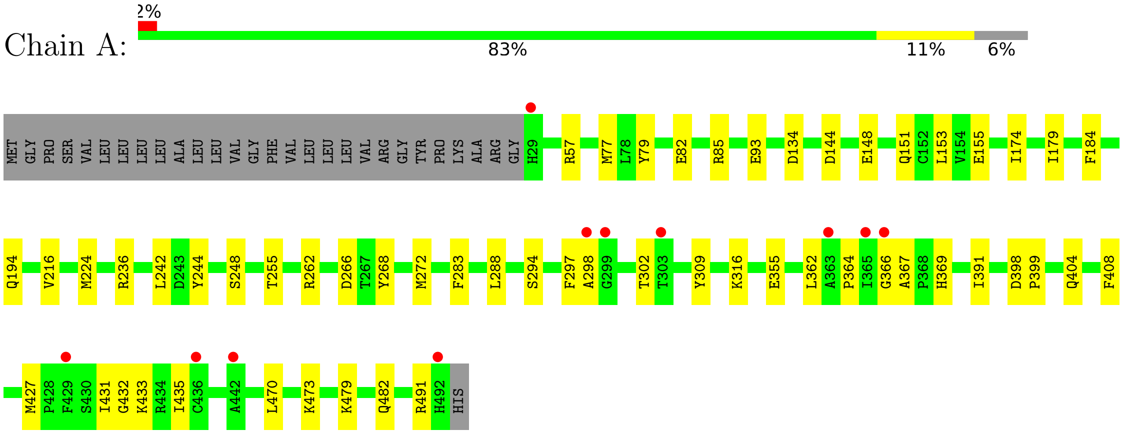
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	211	Total O 211 211	0	0
7	B	59	Total O 59 59	0	0
7	C	124	Total O 124 124	0	0
7	D	106	Total O 106 106	0	0
7	E	99	Total O 99 99	0	0
7	F	60	Total O 60 60	0	0

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

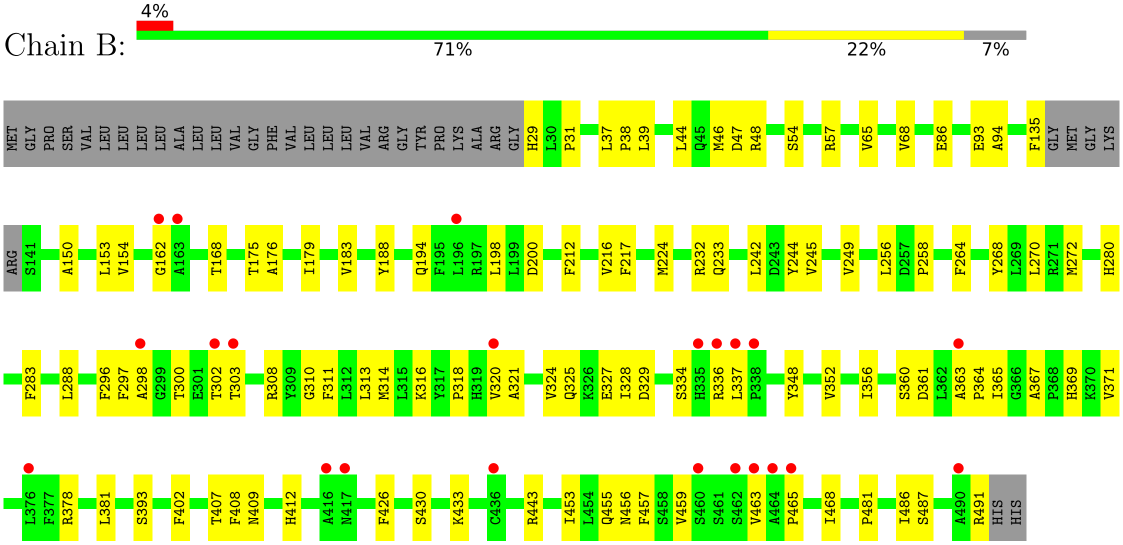
- Molecule 1: Cytochrome P450 family 2 subfamily B

Chain A:



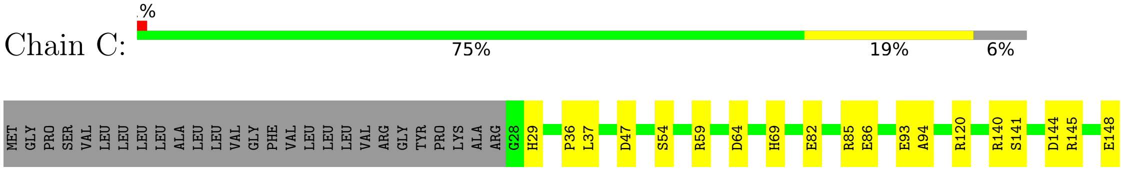
- Molecule 1: Cytochrome P450 family 2 subfamily B

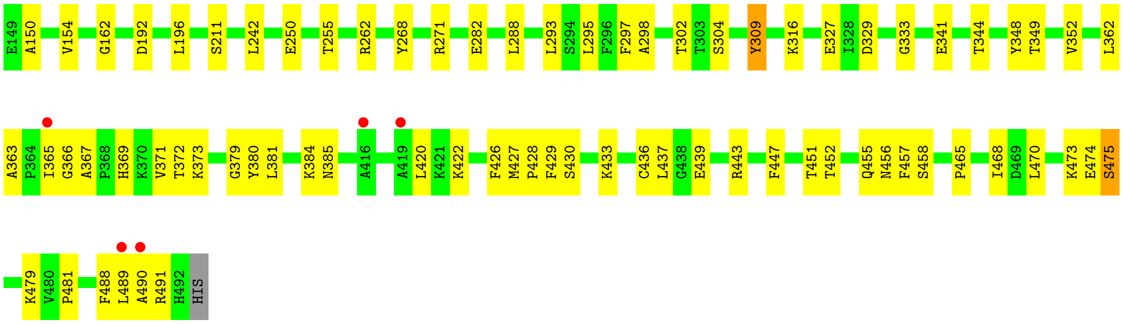
Chain B:



- Molecule 1: Cytochrome P450 family 2 subfamily B

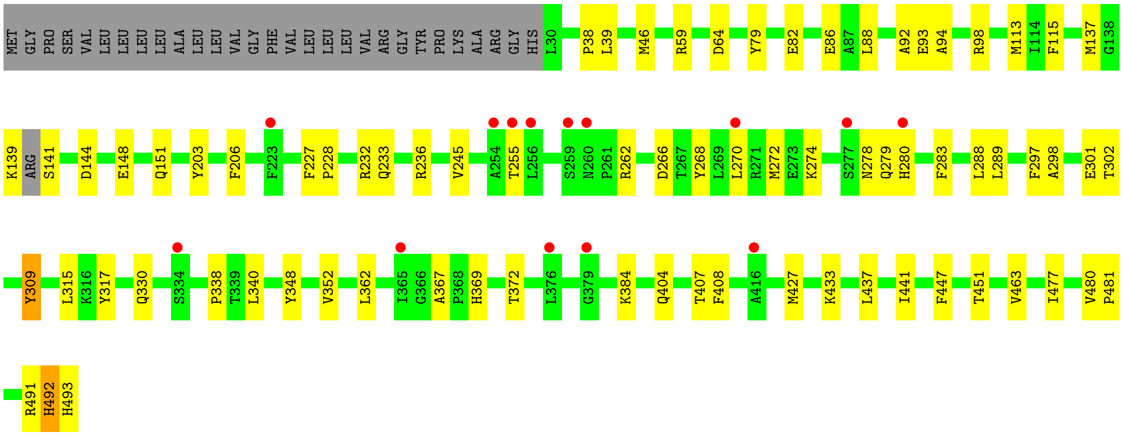
Chain C:





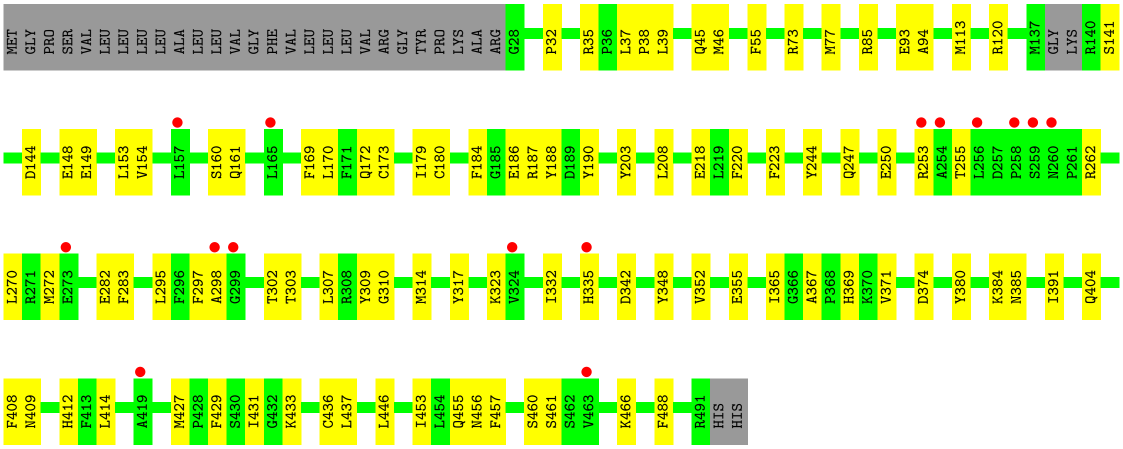
- Molecule 1: Cytochrome P450 family 2 subfamily B

Chain D:



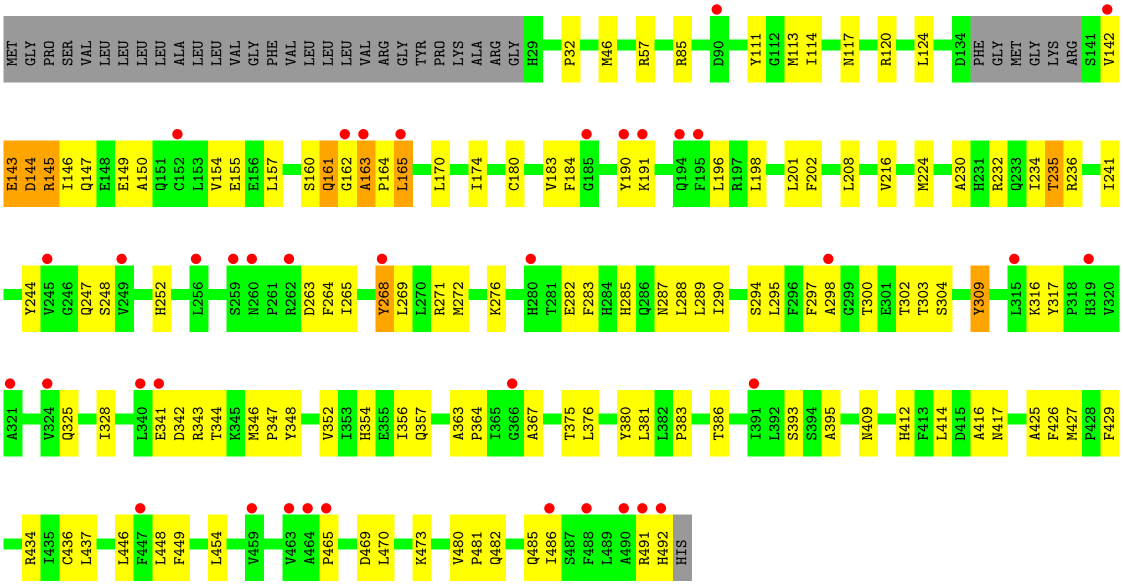
- Molecule 1: Cytochrome P450 family 2 subfamily B

Chain E:



- Molecule 1: Cytochrome P450 family 2 subfamily B

Chain F:



- Molecule 2: beta-D-fructofuranose-(2-1)-alpha-D-glucopyranose

Chain G:

4 Data and refinement statistics ⓘ

Property	**Value**	**Source**
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	98.08Å 106.07Å 106.20Å 64.61° 82.53° 69.93°	Depositor
Resolution (Å)	40.00 – 2.40 39.72 – 2.40	Depositor EDS
% Data completeness (in resolution range)	94.9 (40.00-2.40) 94.9 (39.72-2.40)	Depositor EDS
R$_{merge}$	0.24	Depositor
R$_{sym}$	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.57 (at 2.39Å)	Xtriage
Refinement program	REFMAC 5.5.0072	Depositor
R, R$_{free}$	0.216 , 0.283 0.229 , 0.277	Depositor DCC
R$_{free}$ test set	6567 reflections (4.88%)	wwPDB-VP
Wilson B-factor (Å^2)	43.0	Xtriage
Anisotropy	0.038	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 63.4	EDS
L-test for twinning[2]	$< \lvert L\rvert > = 0.48$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F$_o$,F$_c$ correlation	0.93	EDS
Total number of atoms	23024	wwPDB-VP
Average B, all atoms (Å^2)	47.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.90% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, CM5, CPZ, HEM, FRU, GLC

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	0.57	0/3819	0.67	0/5163
1	B	0.61	0/3692	0.67	0/5008
1	C	0.57	0/3778	0.67	0/5114
1	D	0.56	0/3757	0.64	0/5090
1	E	0.51	0/3752	0.64	0/5080
1	F	0.60	0/3637	0.72	0/4936
All	All	0.57	0/22435	0.67	0/30391

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	3727	0	3708	49	0
1	B	3602	0	3500	96	1
1	C	3685	0	3625	94	0
1	D	3665	0	3594	77	0
1	E	3661	0	3581	91	1
1	F	3558	0	3406	134	0
2	G	23	0	18	13	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	43	0	30	9	0
3	B	43	0	30	9	0
3	C	43	0	30	7	0
3	D	43	0	30	2	0
3	E	43	0	30	8	0
3	F	43	0	30	7	0
4	A	24	0	14	4	0
4	B	24	0	14	9	0
4	C	24	0	14	6	0
4	D	24	0	14	5	0
4	E	24	0	14	8	0
4	F	24	0	14	10	0
5	A	6	0	10	0	0
5	E	12	0	21	2	0
6	A	6	0	8	0	0
6	B	6	0	8	1	0
6	C	6	0	8	0	0
6	E	6	0	8	0	0
7	A	211	0	0	12	0
7	B	59	0	0	8	0
7	C	124	0	0	13	0
7	D	106	0	0	13	0
7	E	99	0	0	14	0
7	F	60	0	0	16	0
All	All	23024	0	21759	571	1

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 13.

The worst 5 of 571 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:E:455:GLN:NE2	7:E:601:HOH:O	1.79	1.15
2:G:1:GLC:O2	2:G:2:FRU:O3	1.70	1.06
1:A:297:PHE:O	1:A:298:ALA:HB3	1.60	1.02
1:B:297:PHE:O	1:B:298:ALA:HB3	1.60	1.01
1:F:157:LEU:O	1:F:160:SER:HB3	1.62	1.00

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:280:HIS:O	1:E:335:HIS:NE2[1_554]	2.03	0.17

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	462/493 (94%)	447 (97%)	15 (3%)	0	100	100
1	B	454/493 (92%)	432 (95%)	22 (5%)	0	100	100
1	C	463/493 (94%)	450 (97%)	12 (3%)	1 (0%)	47	62
1	D	459/493 (93%)	446 (97%)	12 (3%)	1 (0%)	47	62
1	E	458/493 (93%)	444 (97%)	14 (3%)	0	100	100
1	F	455/493 (92%)	424 (93%)	26 (6%)	5 (1%)	14	20
All	All	2751/2958 (93%)	2643 (96%)	101 (4%)	7 (0%)	41	55

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	475	SER
1	F	163	ALA
1	F	144	ASP
1	F	235	THR
1	D	492	HIS

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	408/433 (94%)	407 (100%)	1 (0%)	93	97
1	B	384/433 (89%)	383 (100%)	1 (0%)	92	97
1	C	396/433 (92%)	394 (100%)	2 (0%)	88	95
1	D	395/433 (91%)	393 (100%)	2 (0%)	88	95
1	E	393/433 (91%)	392 (100%)	1 (0%)	92	97
1	F	370/433 (86%)	365 (99%)	5 (1%)	67	82
All	All	2346/2598 (90%)	2334 (100%)	12 (0%)	88	95

5 of 12 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	145	ARG
1	F	165	LEU
1	F	343	ARG
1	F	268	TYR
1	C	458	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	F	325	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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

5.5 Carbohydrates ⓘ

2 monosaccharides 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	$\#\|Z\| > 2$
2	GLC	G	1	2	11,11,12	3.30	8 (72%)	15,15,17	3.30	10 (66%)
2	FRU	G	2	2	11,12,12	0.58	0	10,18,18	0.54	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
2	GLC	G	1	2	-	2/2/19/22	0/1/1/1
2	FRU	G	2	2	-	3/5/24/24	0/1/1/1

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	G	1	GLC	O5-C1	-5.00	1.35	1.43
2	G	1	GLC	C2-C3	-4.10	1.46	1.52
2	G	1	GLC	C4-C3	-4.05	1.42	1.52
2	G	1	GLC	O2-C2	-3.92	1.35	1.43
2	G	1	GLC	O4-C4	-3.75	1.34	1.43

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	G	1	GLC	O5-C1-C2	-7.64	98.98	110.77
2	G	1	GLC	O4-C4-C3	-4.45	100.06	110.35
2	G	1	GLC	O5-C5-C6	3.85	113.23	107.20
2	G	1	GLC	C1-C2-C3	3.83	114.38	109.67
2	G	1	GLC	O2-C2-C1	3.61	116.53	109.15

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	G	2	FRU	O1-C1-C2-C3
2	G	2	FRU	O1-C1-C2-O2
2	G	1	GLC	C4-C5-C6-O6
2	G	1	GLC	O5-C5-C6-O6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
2	G	2	FRU	O1-C1-C2-O5

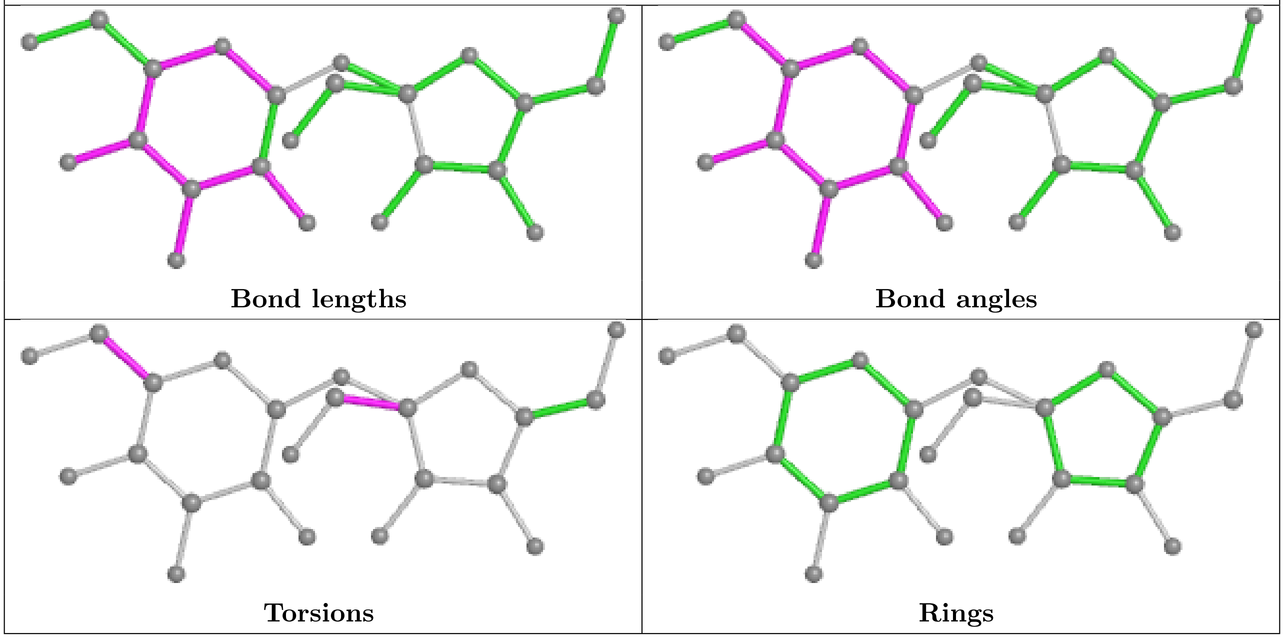
There are no ring outliers.

2 monomers are involved in 13 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	G	1	GLC	4	0
2	G	2	FRU	13	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.

Oligosaccharide Chain G



5.6 Ligand geometry ⓘ

24 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
4	CPZ	C	502	3	10,13,13	1.01	0	9,17,17	0.81	0
4	CPZ	E	503	-	10,13,13	1.00	0	9,17,17	0.80	0
4	CPZ	A	503	-	10,13,13	1.00	0	9,17,17	0.79	0
4	CPZ	C	503	-	10,13,13	1.00	0	9,17,17	0.79	0
6	GOL	A	506	-	5,5,5	0.57	0	5,5,5	0.96	0
3	HEM	C	501	4,1	41,50,50	2.10	9 (21%)	45,82,82	1.41	7 (15%)
4	CPZ	A	502	3	10,13,13	1.00	0	9,17,17	0.80	0
6	GOL	C	504	-	5,5,5	0.41	0	5,5,5	0.60	0
4	CPZ	E	502	3	10,13,13	1.00	0	9,17,17	0.80	0
4	CPZ	B	502	3	10,13,13	1.24	1 (10%)	9,17,17	1.04	0
3	HEM	A	501	4,1	41,50,50	2.24	13 (31%)	45,82,82	1.53	9 (20%)
6	GOL	E	505	-	5,5,5	0.47	0	5,5,5	0.57	0
4	CPZ	B	503	-	10,13,13	0.98	0	9,17,17	1.53	1 (11%)
3	HEM	E	501	4,1	41,50,50	2.11	6 (14%)	45,82,82	1.68	7 (15%)
3	HEM	F	501	4	41,50,50	2.00	8 (19%)	45,82,82	1.92	8 (17%)
3	HEM	B	501	4,1	41,50,50	1.90	6 (14%)	45,82,82	1.57	9 (20%)
4	CPZ	D	502	-	10,13,13	1.00	0	9,17,17	0.80	0
4	CPZ	F	502	3	10,13,13	1.00	0	9,17,17	0.80	0
4	CPZ	D	503	3	10,13,13	0.99	0	9,17,17	0.77	0
4	CPZ	F	503	-	10,13,13	1.01	0	9,17,17	0.80	0
5	CM5	E	504	-	12,12,36	1.15	1 (8%)	13,13,49	1.16	0
3	HEM	D	501	4,1	41,50,50	1.87	6 (14%)	45,82,82	1.78	8 (17%)
6	GOL	B	504	-	5,5,5	0.49	0	5,5,5	0.78	0
5	CM5	A	505	-	6,6,36	0.38	0	6,6,49	0.62	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
4	CPZ	C	502	3	-	0/0/4/4	0/2/2/2
4	CPZ	E	503	-	-	0/0/4/4	0/2/2/2
4	CPZ	A	503	-	-	0/0/4/4	0/2/2/2
4	CPZ	C	503	-	-	0/0/4/4	0/2/2/2
6	GOL	A	506	-	-	2/4/4/4	-
3	HEM	C	501	4,1	-	2/12/54/54	-
4	CPZ	A	502	3	-	0/0/4/4	0/2/2/2
6	GOL	C	504	-	-	0/4/4/4	-
4	CPZ	E	502	3	-	0/0/4/4	0/2/2/2

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	CPZ	B	502	3	-	0/0/4/4	0/2/2/2
3	HEM	A	501	4,1	-	4/12/54/54	-
6	GOL	E	505	-	-	4/4/4/4	-
4	CPZ	B	503	-	-	0/0/4/4	0/2/2/2
3	HEM	E	501	4,1	-	4/12/54/54	-
3	HEM	F	501	4	-	8/12/54/54	-
3	HEM	B	501	4,1	-	4/12/54/54	-
4	CPZ	D	502	-	-	0/0/4/4	0/2/2/2
4	CPZ	F	502	3	-	0/0/4/4	0/2/2/2
4	CPZ	D	503	3	-	0/0/4/4	0/2/2/2
4	CPZ	F	503	-	-	0/0/4/4	0/2/2/2
5	CM5	E	504	-	-	2/6/14/65	0/1/1/3
3	HEM	D	501	4,1	-	4/12/54/54	-
6	GOL	B	504	-	-	4/4/4/4	-
5	CM5	A	505	-	-	-	0/1/1/3

The worst 5 of 50 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	E	501	HEM	C3D-C2D	8.27	1.54	1.36
3	C	501	HEM	C3D-C2D	7.98	1.53	1.36
3	B	501	HEM	C3D-C2D	7.77	1.53	1.36
3	F	501	HEM	C3D-C2D	7.50	1.52	1.36
3	D	501	HEM	C3D-C2D	7.01	1.51	1.36

The worst 5 of 49 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	E	501	HEM	C4D-ND-C1D	6.26	111.54	105.07
3	D	501	HEM	C4D-ND-C1D	5.72	110.98	105.07
3	F	501	HEM	C4D-ND-C1D	5.34	110.59	105.07
3	F	501	HEM	CAD-CBD-CGD	-4.62	103.67	113.60
3	C	501	HEM	C4D-ND-C1D	4.51	109.73	105.07

There are no chirality outliers.

5 of 38 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	F	501	HEM	C1A-C2A-CAA-CBA
3	F	501	HEM	C3A-C2A-CAA-CBA

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
6	A	506	GOL	C1-C2-C3-O3
6	A	506	GOL	O2-C2-C3-O3
6	B	504	GOL	O1-C1-C2-C3

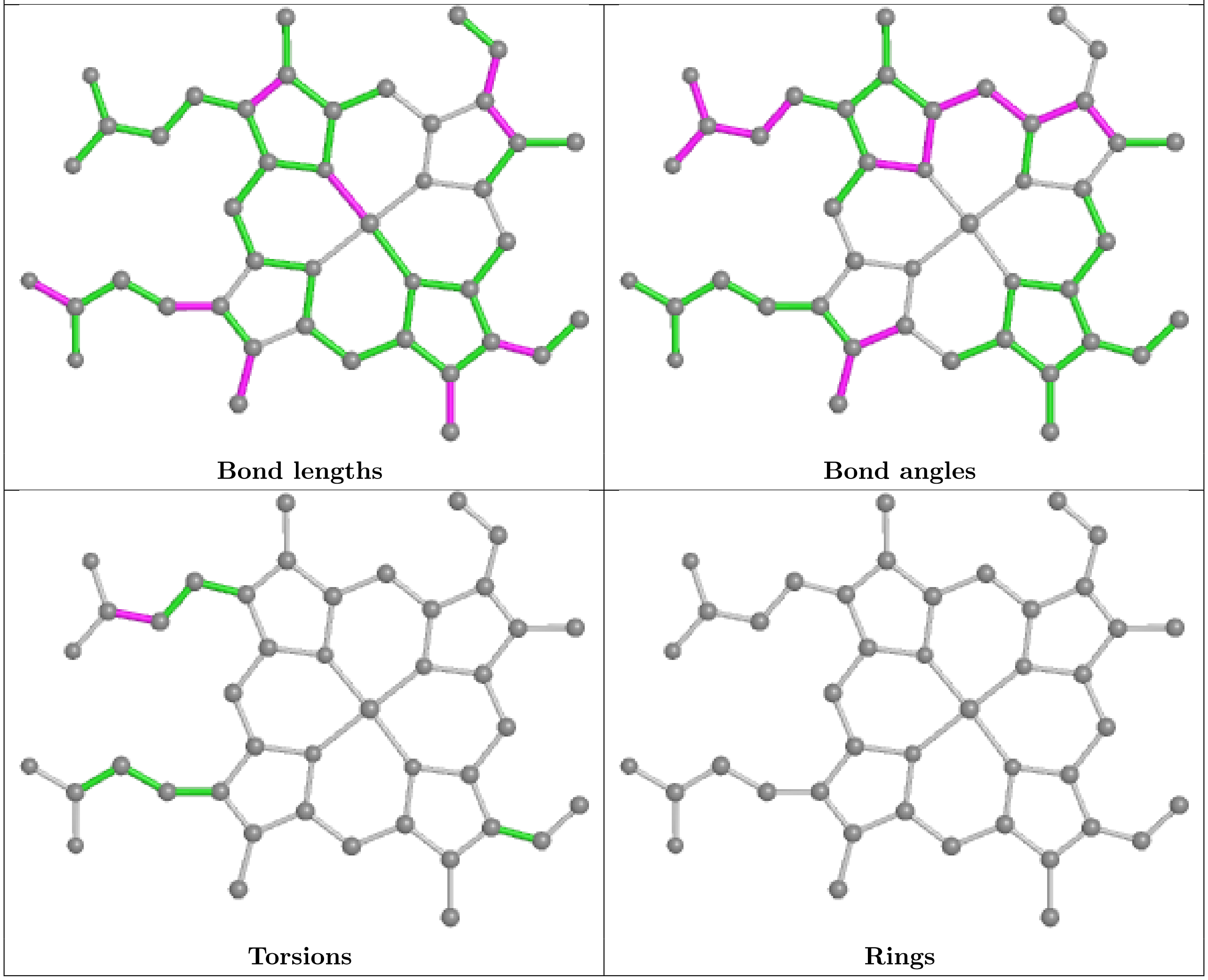
There are no ring outliers.

20 monomers are involved in 78 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	C	502	CPZ	4	0
4	E	503	CPZ	3	0
4	A	503	CPZ	1	0
4	C	503	CPZ	3	0
3	C	501	HEM	7	0
4	A	502	CPZ	4	0
4	E	502	CPZ	6	0
4	B	502	CPZ	5	0
3	A	501	HEM	9	0
4	B	503	CPZ	5	0
3	E	501	HEM	8	0
3	F	501	HEM	7	0
3	B	501	HEM	9	0
4	D	502	CPZ	4	0
4	F	502	CPZ	5	0
4	D	503	CPZ	2	0
4	F	503	CPZ	6	0
5	E	504	CM5	2	0
3	D	501	HEM	2	0
6	B	504	GOL	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 HEM C 501



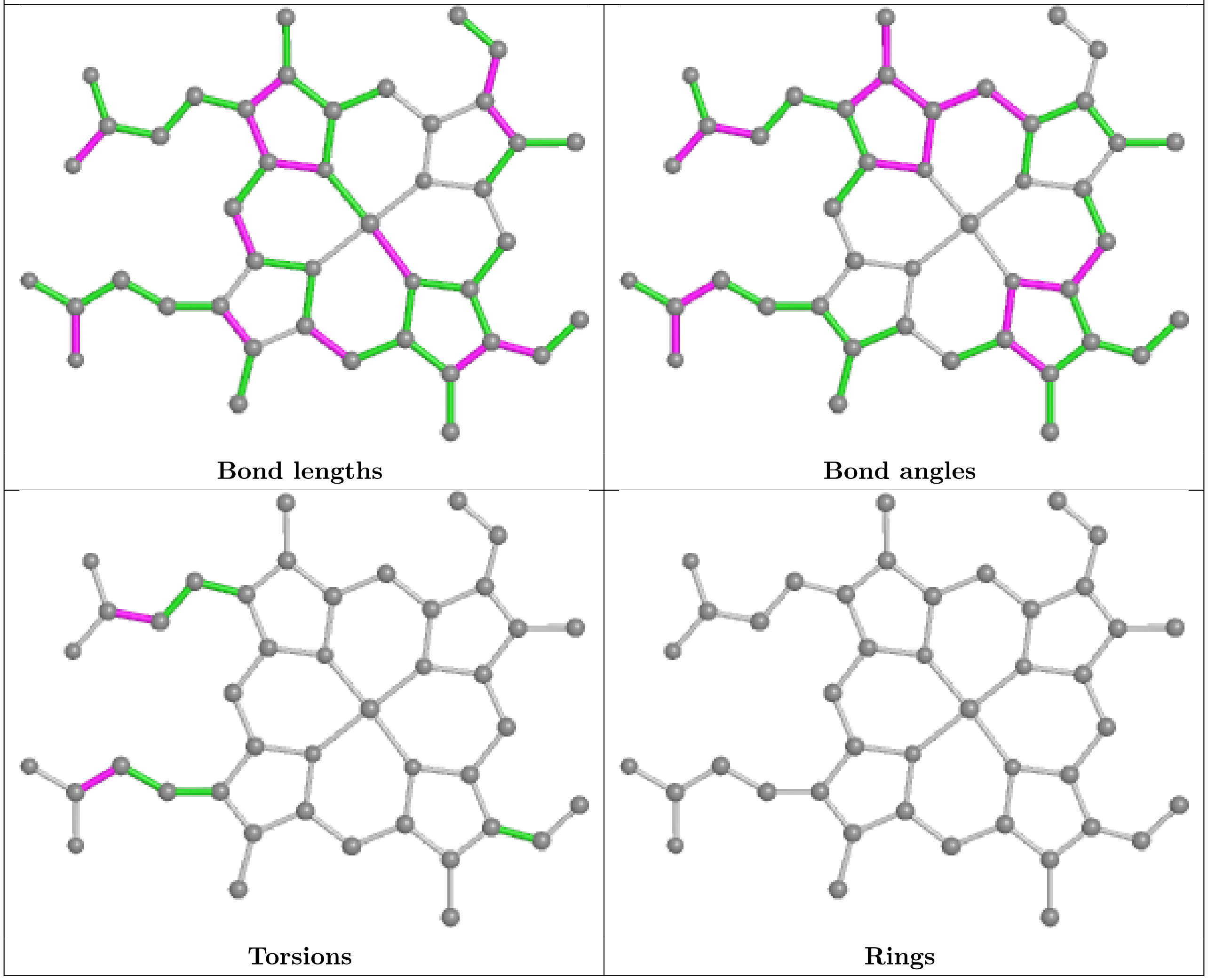
Bond lengths

Bond angles

Torsions

Rings

Ligand HEM A 501



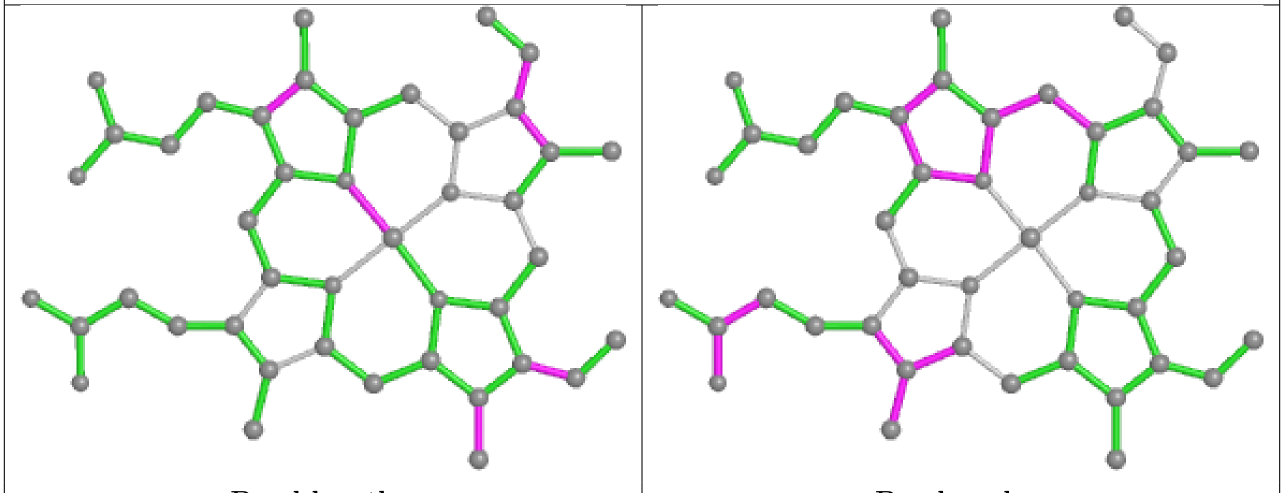
Bond lengths

Bond angles

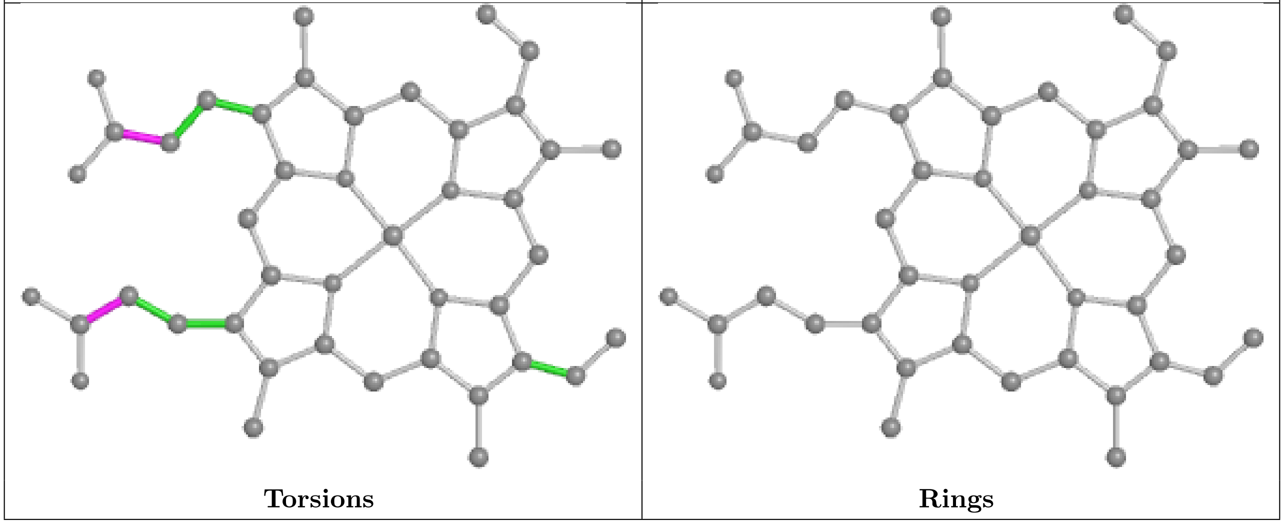
Torsions

Rings

Ligand HEM E 501

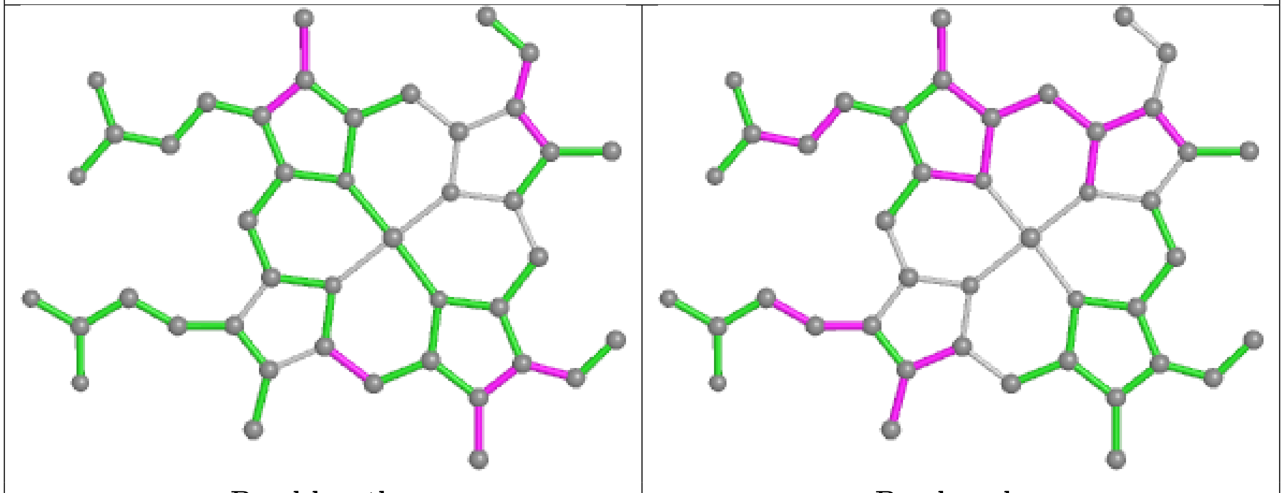


Bond lengths **Bond angles**

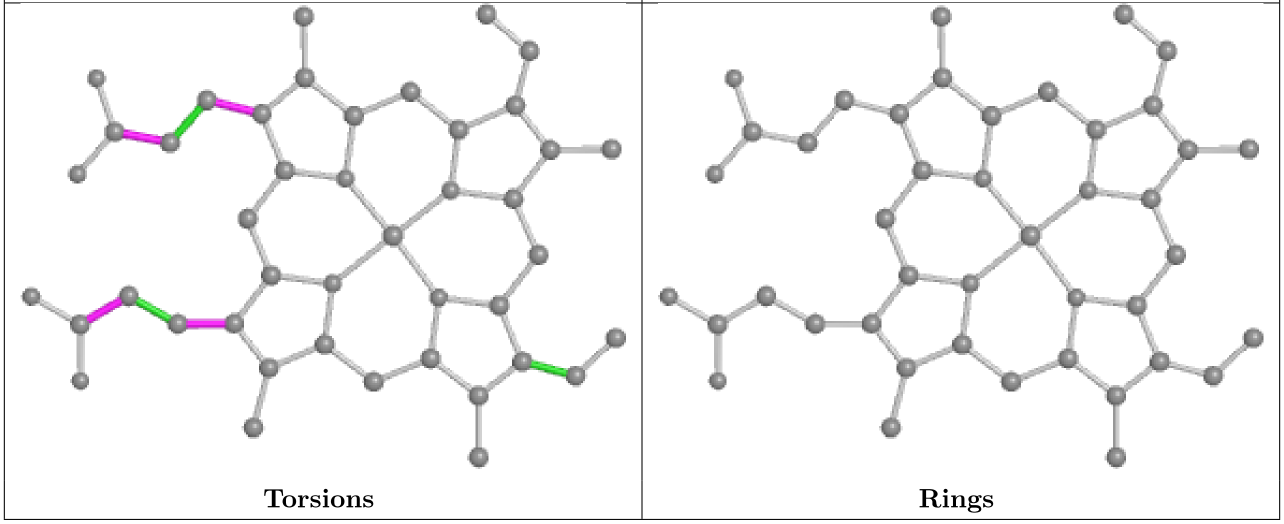


Torsions **Rings**

Ligand HEM F 501

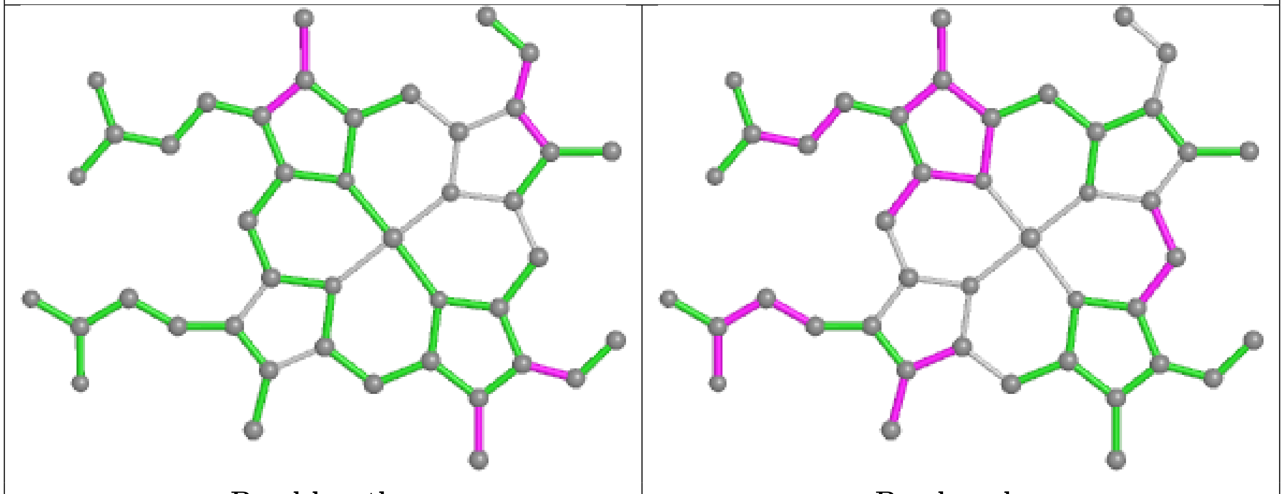


Bond lengths **Bond angles**

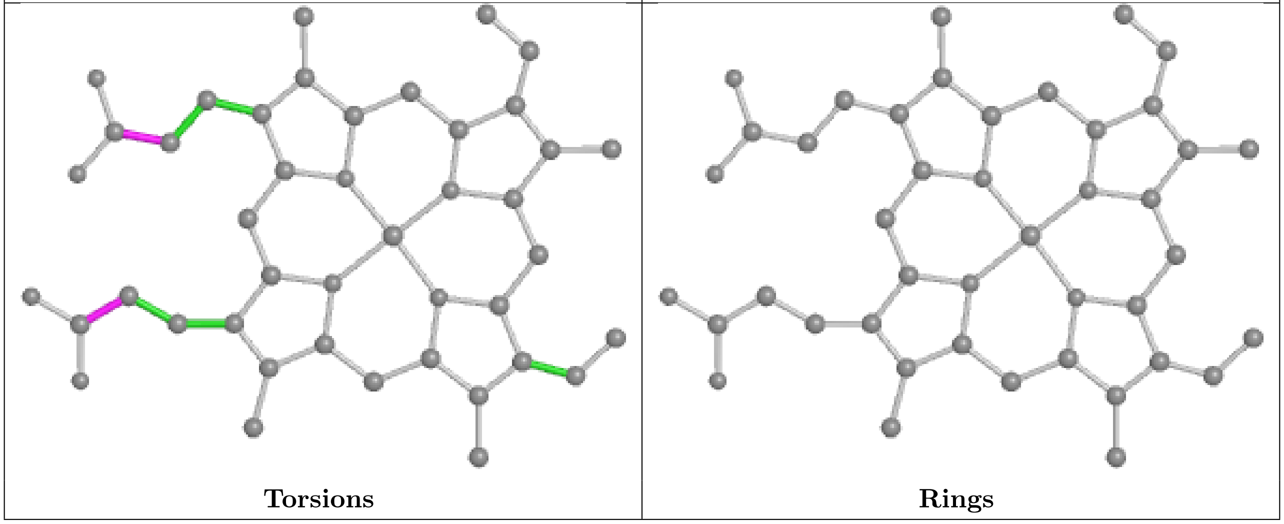


Torsions **Rings**

Ligand HEM B 501

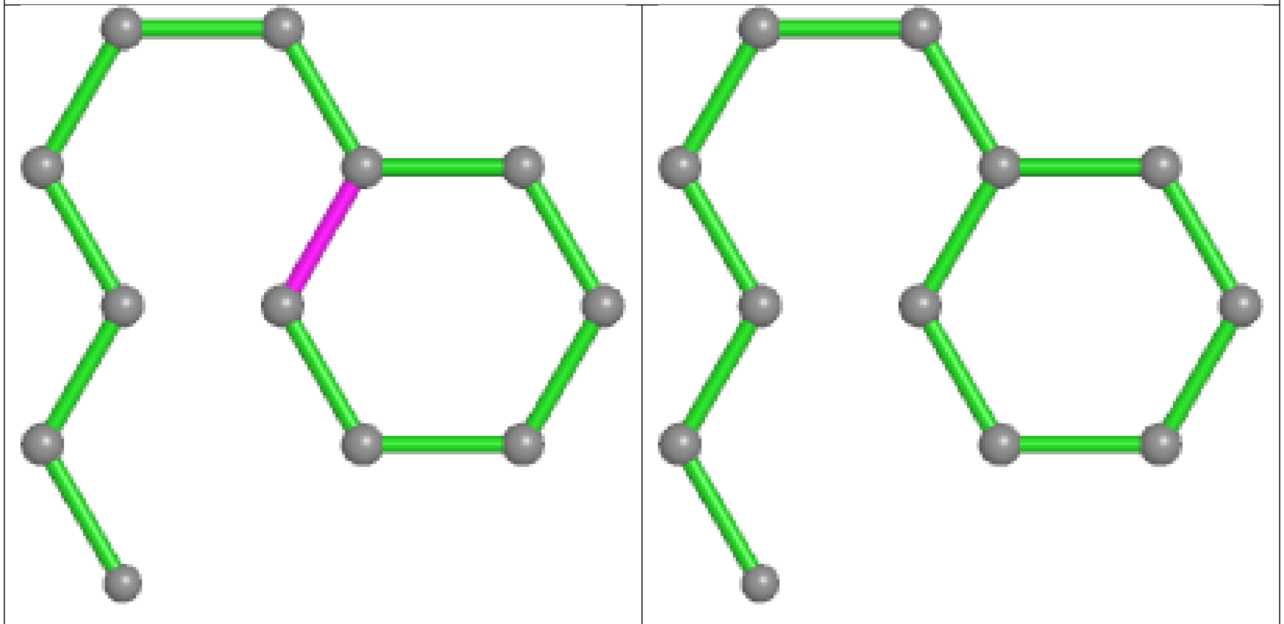


Bond lengths **Bond angles**



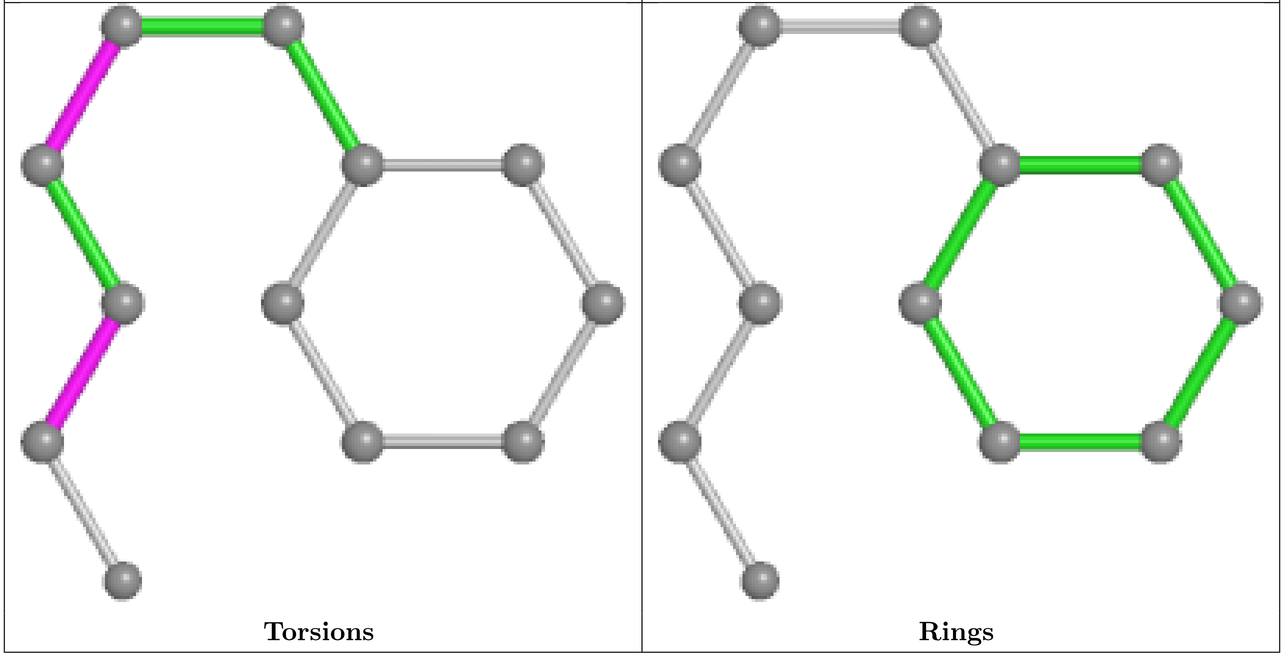
Torsions **Rings**

Ligand CM5 E 504



Bond lengths

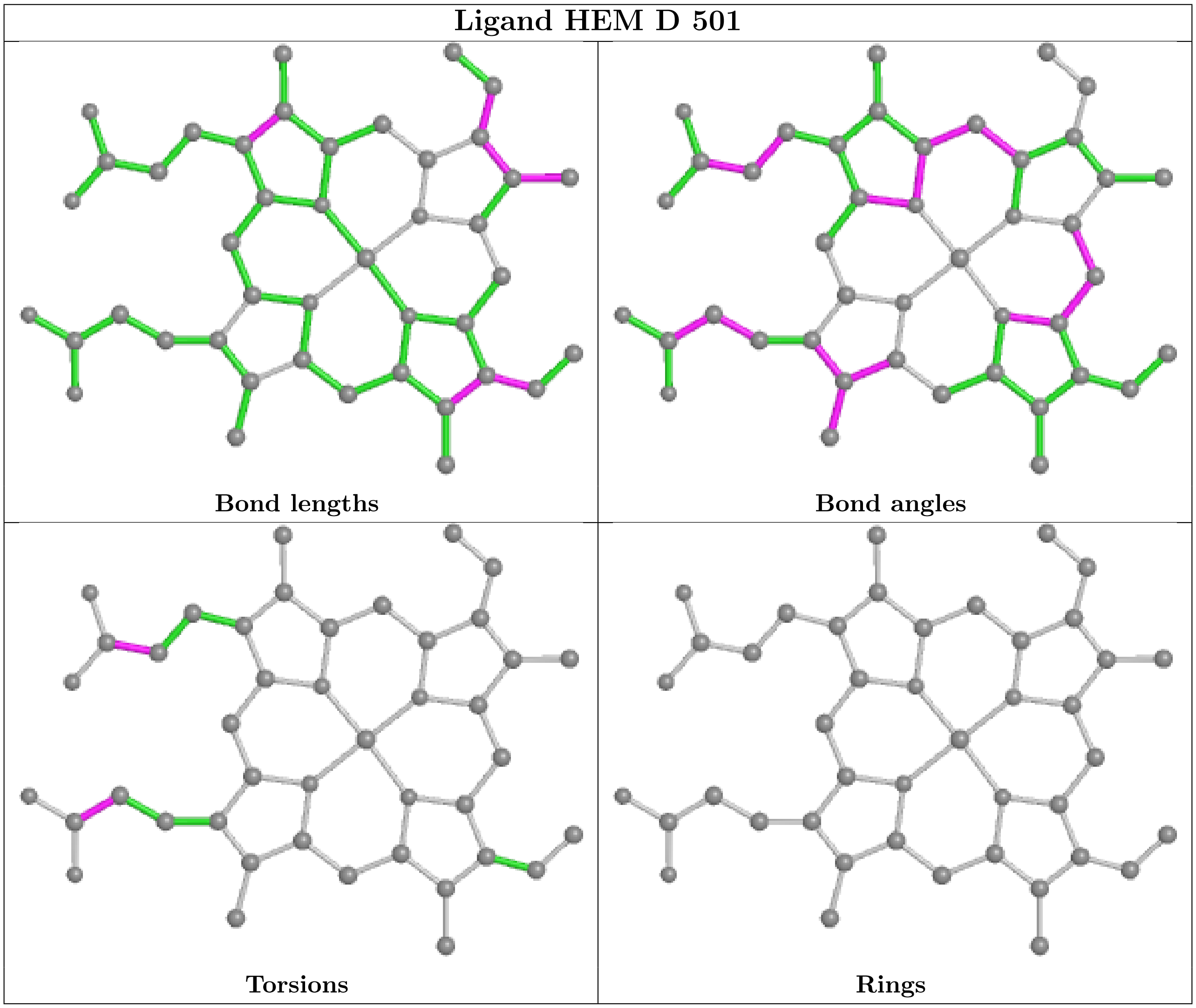
Bond angles



Torsions

Rings

Ligand HEM D 501



Bond lengths

Bond angles

Torsions

Rings

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, 95th 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	464/493 (94%)	-0.21	11 (2%)	59	57	19, 35, 60, 90	0
1	B	458/493 (92%)	0.13	22 (4%)	30	29	30, 50, 73, 86	0
1	C	465/493 (94%)	-0.17	5 (1%)	80	79	22, 43, 71, 90	0
1	D	463/493 (93%)	-0.10	14 (3%)	50	49	24, 45, 63, 84	0
1	E	462/493 (93%)	-0.03	15 (3%)	47	46	27, 48, 71, 79	0
1	F	458/493 (92%)	0.40	38 (8%)	11	10	27, 57, 75, 95	0
All	All	2770/2958 (93%)	0.00	105 (3%)	40	39	19, 47, 71, 95	0

The worst 5 of 105 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	492	HIS	5.1
1	F	490	ALA	5.0
1	F	259	SER	4.8
1	F	260	ASN	4.7
1	F	280	HIS	4.5

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

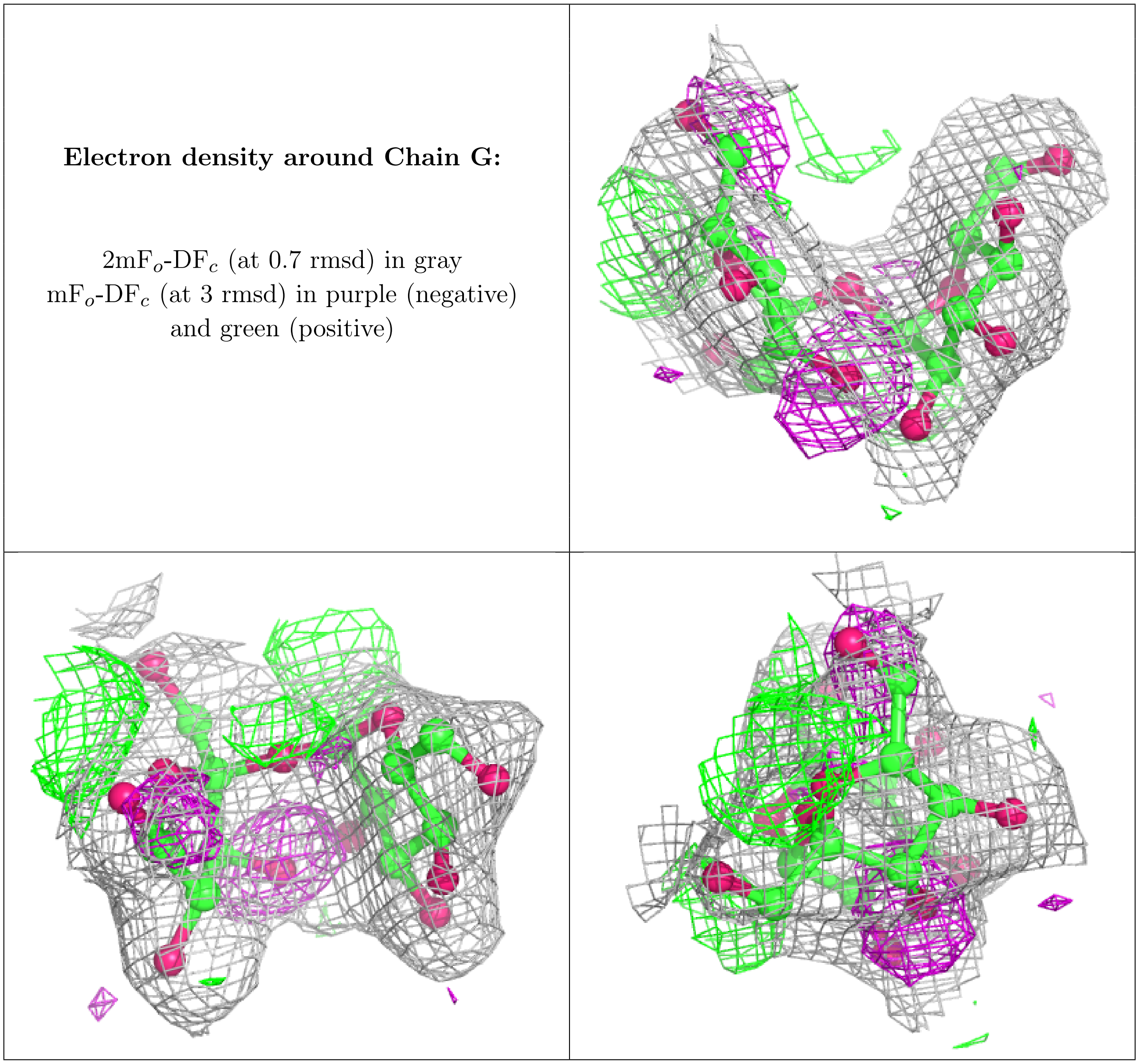
There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

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, 95th 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
2	FRU	G	2	12/12	0.70	0.24	35,45,54,55	0
2	GLC	G	1	11/12	0.88	0.11	59,60,62,63	0

The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.



Electron density around Chain G:

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

6.4 Ligands ⓘ

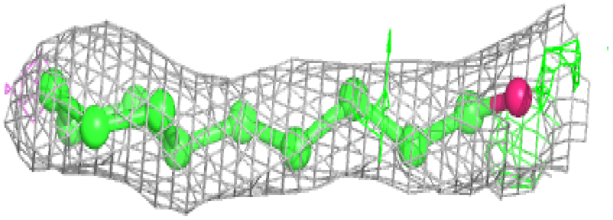
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

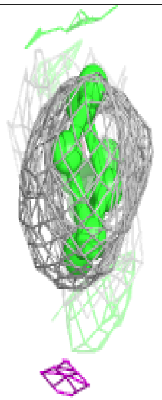
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($Å^2$)	Q<0.9
6	GOL	A	506	6/6	0.74	0.09	47,52,53,54	0
6	GOL	C	504	6/6	0.74	0.14	65,69,69,71	0
6	GOL	E	505	6/6	0.83	0.13	65,67,68,69	0
4	CPZ	D	502	12/12	0.84	0.22	57,58,61,65	0
6	GOL	B	504	6/6	0.84	0.15	54,59,60,61	0
4	CPZ	A	503	12/12	0.85	0.23	49,52,56,60	0
4	CPZ	B	503	12/12	0.86	0.27	68,70,74,76	0
4	CPZ	F	503	12/12	0.88	0.23	68,69,71,71	0
5	CM5	E	504	12/34	0.88	0.18	63,66,67,67	0
4	CPZ	E	503	12/12	0.90	0.23	69,71,72,76	0
4	CPZ	F	502	12/12	0.92	0.25	54,65,69,72	0
5	CM5	A	505	6/34	0.93	0.16	55,55,56,56	0
4	CPZ	C	503	12/12	0.93	0.16	49,51,53,60	0
4	CPZ	D	503	12/12	0.95	0.17	35,42,49,52	0
3	HEM	E	501	43/43	0.96	0.20	33,38,42,47	0
4	CPZ	B	502	12/12	0.96	0.20	39,41,43,44	0
4	CPZ	E	502	12/12	0.96	0.21	42,46,46,47	0
3	HEM	F	501	43/43	0.97	0.18	40,44,49,53	0
4	CPZ	A	502	12/12	0.97	0.18	31,33,35,37	0
4	CPZ	C	502	12/12	0.97	0.18	40,42,44,45	0
3	HEM	B	501	43/43	0.97	0.20	29,35,37,39	0
3	HEM	D	501	43/43	0.98	0.17	23,30,34,39	0
3	HEM	A	501	43/43	0.98	0.20	12,17,21,26	0
3	HEM	C	501	43/43	0.98	0.16	16,24,29,37	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 CM5 E 504:

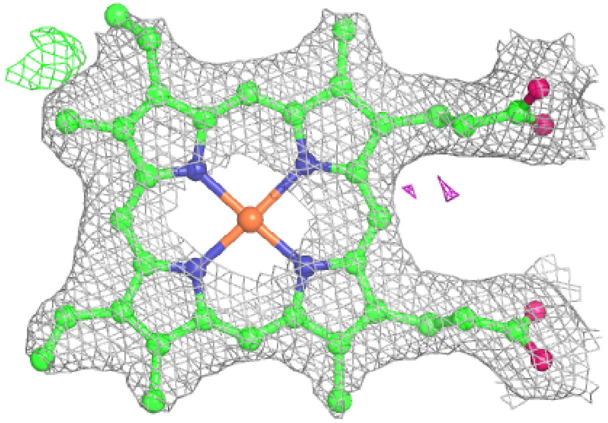
2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative)
and green (positive)

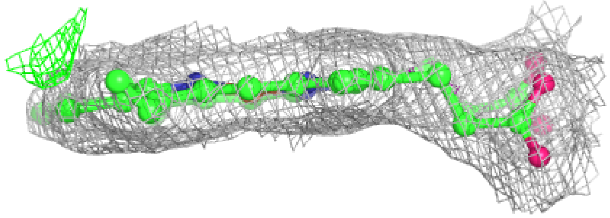


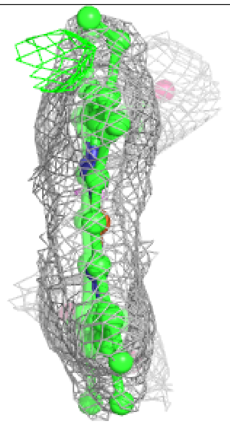


Electron density around HEM E 501:

2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative)
and green (positive)

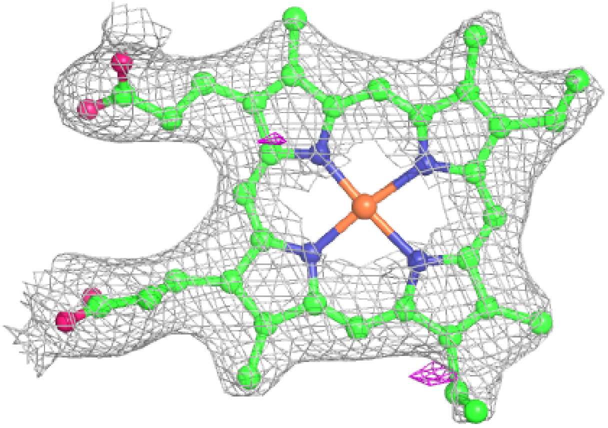


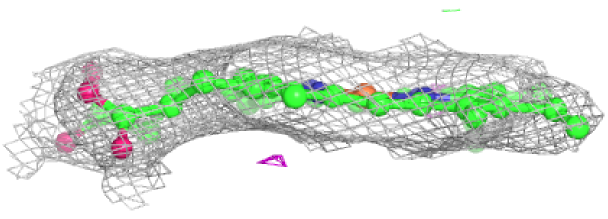


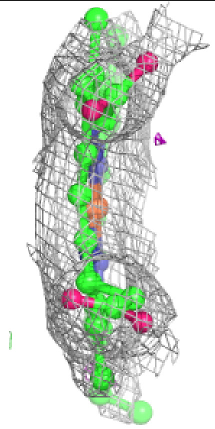


Electron density around HEM F 501:

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

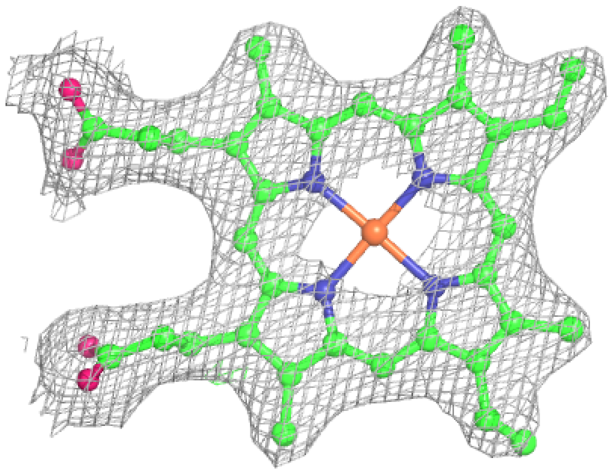


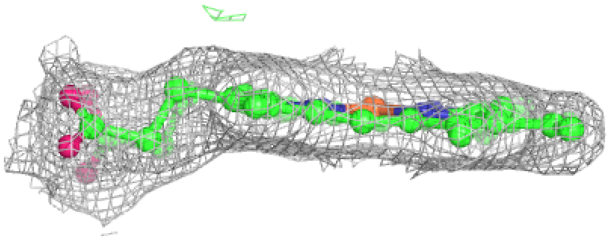


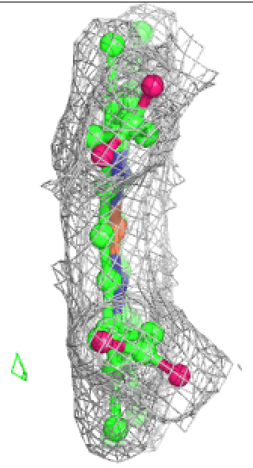


Electron density around HEM B 501:

2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative)
and green (positive)

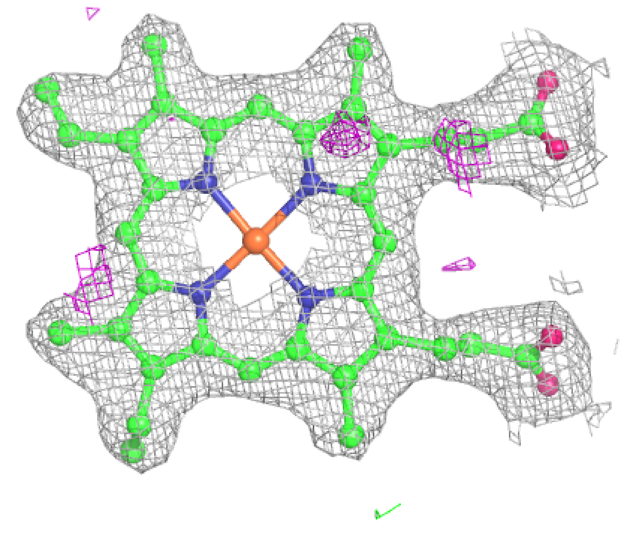


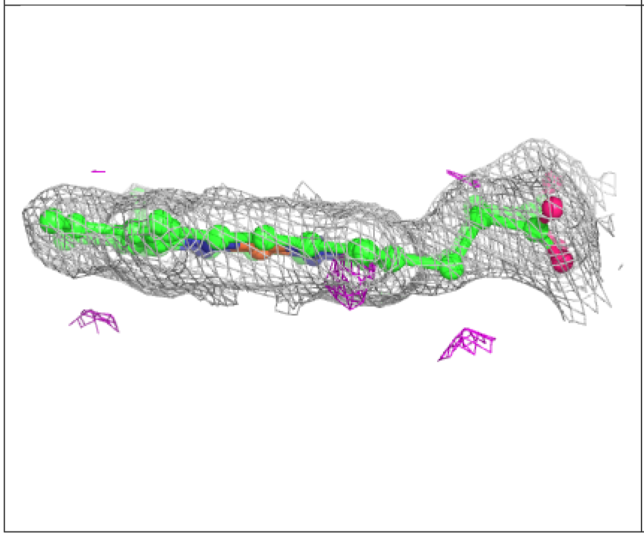


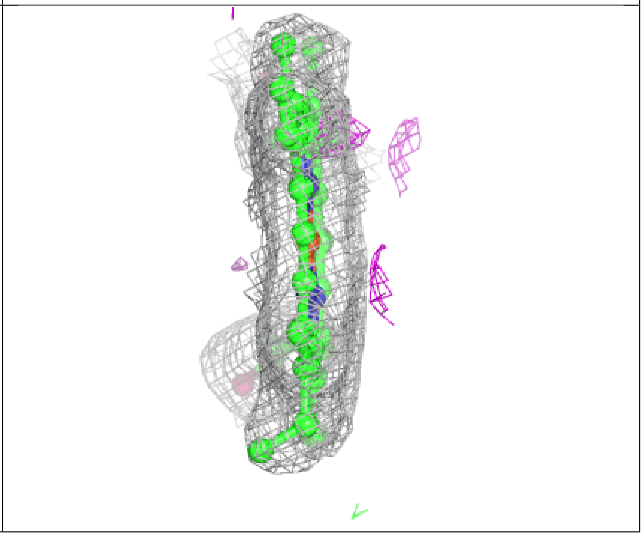


Electron density around HEM D 501:

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

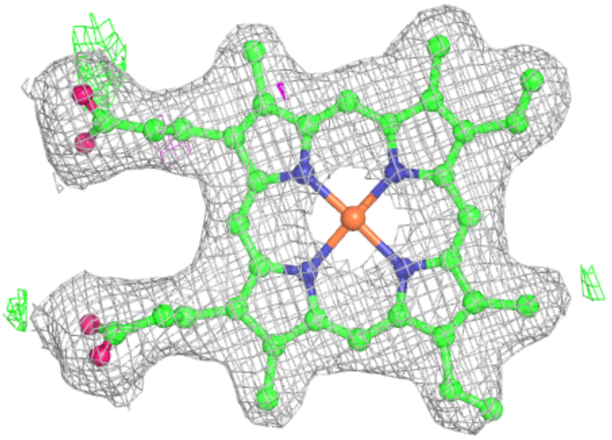


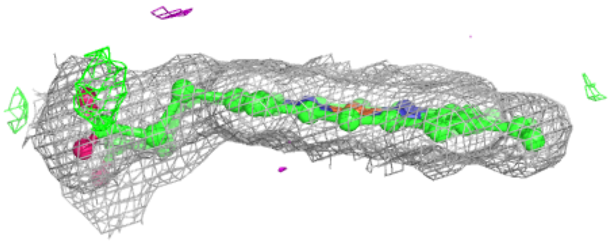


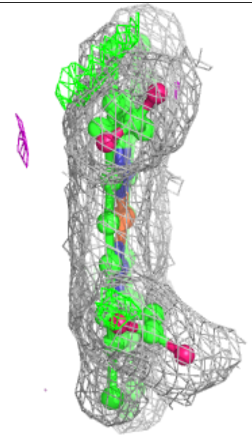


Electron density around HEM A 501:

2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative)
and green (positive)

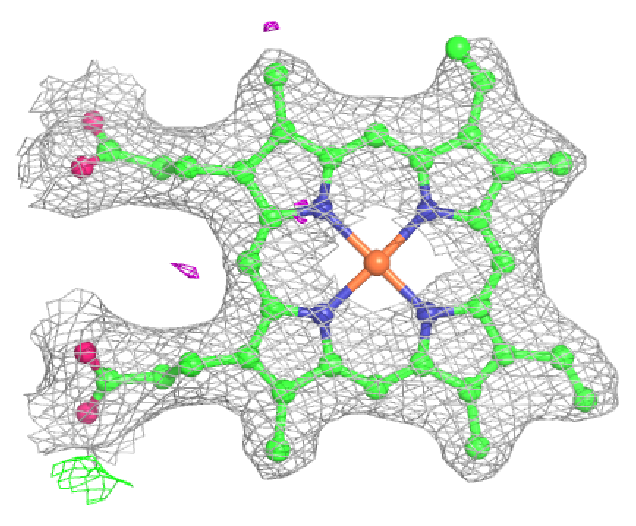


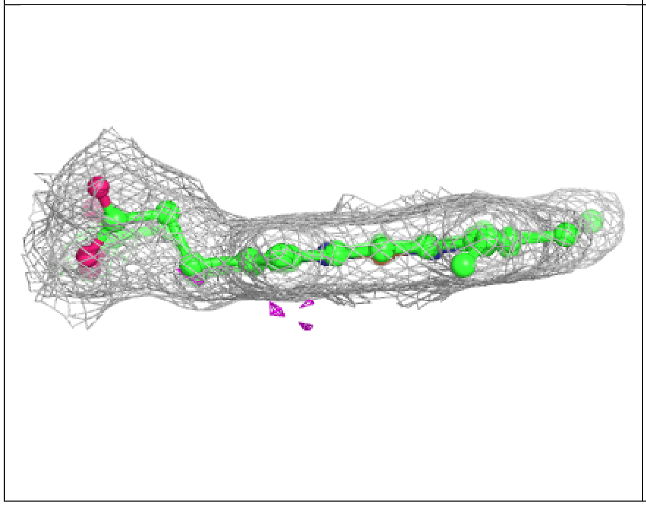


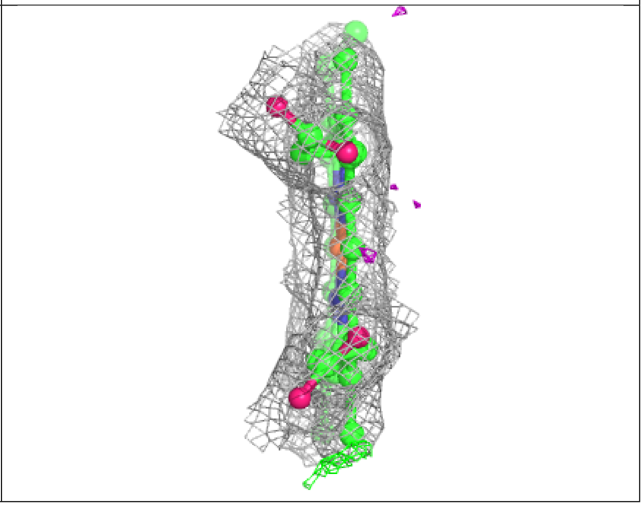


Electron density around HEM C 501:

$2mF_o$-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 ⓘ

There are no such residues in this entry.