Full wwPDB X-ray Structure Validation Report (i)

Oct 4, 2023 – 07:15 PM EDT

PDB ID	:	6UDL
Title	:	Structure of Human Cytochrome P450 1A1 with Duocarmycin Prodrug (S) ICT-2700
Authors	:	Bart, A.G.; Scott, E.E.
Deposited on	:	2019-09-19
Resolution	:	2.85 Å(reported)

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

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

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

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 2.85 Å.

There are no overall percentile quality scores available for this entry.

MolProbity and EDS failed to run properly - the sequence quality summary graphics cannot be shown.

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 30414 atoms, of which 15172 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 1A1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	469	7510	2406	3756	653	674	21	0	0	0
1	B	477	7627	2442	3811	664	689	21	0	0	0
1	C	466	7456	2390	3729	648	668	21	0	0	0
1	D	465	7438	2385	3720	646	666	21	0	0	0

There are 52 discrepancies between the modelled and reference sequences:

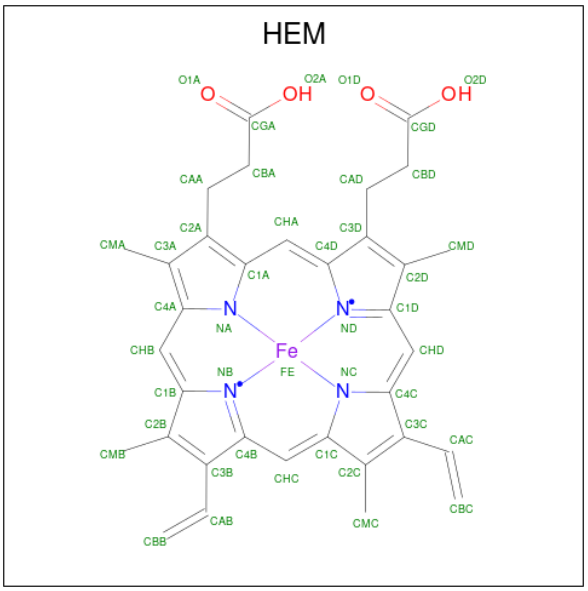
Chain	Residue	Modelled	Actual	Comment	Reference
A	28	MET	-	expression tag	UNP P04798
A	29	ALA	-	expression tag	UNP P04798
A	30	LYS	-	expression tag	UNP P04798
A	31	LYS	-	expression tag	UNP P04798
A	32	THR	-	expression tag	UNP P04798
A	33	SER	-	expression tag	UNP P04798
A	34	SER	-	expression tag	UNP P04798
A	513	HIS	-	expression tag	UNP P04798
A	514	HIS	-	expression tag	UNP P04798
A	515	HIS	-	expression tag	UNP P04798
A	516	HIS	-	expression tag	UNP P04798
A	517	HIS	-	expression tag	UNP P04798
A	518	HIS	-	expression tag	UNP P04798
B	28	MET	-	expression tag	UNP P04798
B	29	ALA	-	expression tag	UNP P04798
B	30	LYS	-	expression tag	UNP P04798
B	31	LYS	-	expression tag	UNP P04798
B	32	THR	-	expression tag	UNP P04798
B	33	SER	-	expression tag	UNP P04798
B	34	SER	-	expression tag	UNP P04798
B	513	HIS	-	expression tag	UNP P04798

Continued on next page...

Continued from previous page...

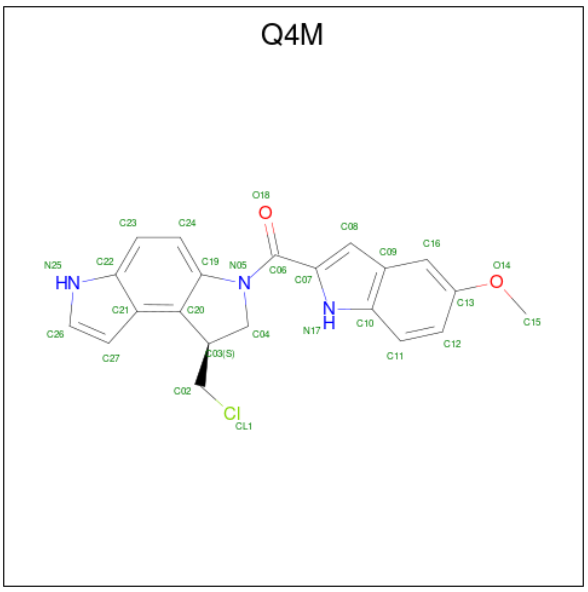
Chain	Residue	Modelled	Actual	Comment	Reference
B	514	HIS	-	expression tag	UNP P04798
B	515	HIS	-	expression tag	UNP P04798
B	516	HIS	-	expression tag	UNP P04798
B	517	HIS	-	expression tag	UNP P04798
B	518	HIS	-	expression tag	UNP P04798
C	28	MET	-	expression tag	UNP P04798
C	29	ALA	-	expression tag	UNP P04798
C	30	LYS	-	expression tag	UNP P04798
C	31	LYS	-	expression tag	UNP P04798
C	32	THR	-	expression tag	UNP P04798
C	33	SER	-	expression tag	UNP P04798
C	34	SER	-	expression tag	UNP P04798
C	513	HIS	-	expression tag	UNP P04798
C	514	HIS	-	expression tag	UNP P04798
C	515	HIS	-	expression tag	UNP P04798
C	516	HIS	-	expression tag	UNP P04798
C	517	HIS	-	expression tag	UNP P04798
C	518	HIS	-	expression tag	UNP P04798
D	28	MET	-	expression tag	UNP P04798
D	29	ALA	-	expression tag	UNP P04798
D	30	LYS	-	expression tag	UNP P04798
D	31	LYS	-	expression tag	UNP P04798
D	32	THR	-	expression tag	UNP P04798
D	33	SER	-	expression tag	UNP P04798
D	34	SER	-	expression tag	UNP P04798
D	513	HIS	-	expression tag	UNP P04798
D	514	HIS	-	expression tag	UNP P04798
D	515	HIS	-	expression tag	UNP P04798
D	516	HIS	-	expression tag	UNP P04798
D	517	HIS	-	expression tag	UNP P04798
D	518	HIS	-	expression tag	UNP P04798

- Molecule 2 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	H	N	O		
2	A	1	73	34	1	30	4	4	0	0
			Total	C	Fe	H	N	O		
2	B	1	73	34	1	30	4	4	0	0
			Total	C	Fe	H	N	O		
2	C	1	73	34	1	30	4	4	0	0
			Total	C	Fe	H	N	O		
2	D	1	73	34	1	30	4	4	0	0

- Molecule 3 is [(1S)-1-(chloromethyl)-1,6-dihydropyrrolo[3,2-e]indol-3(2H)-yl](5-methoxy-1H-indol-2-yl)methanone (three-letter code: Q4M) (formula: $C_{21}H_{18}ClN_3O_2$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
			Total	C	Cl	H	N	O		
3	A	1	45	21	1	18	3	2	0	0
3	B	1	45	21	1	18	3	2	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	B	1	1	1	0	0

MolProbity and EDS failed to run properly - this section is therefore empty.

3 Data and refinement statistics (i)

EDS failed to run properly - this section is therefore incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	65.37Å 196.15Å 237.22Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.80 – 2.85	Depositor
% Data completeness (in resolution range)	99.1 (48.80-2.85)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	0.68 (at 2.86Å)	Xtriage
Refinement program	PHENIX 1.16_3549	Depositor
R, R_{free}	0.238 , 0.287	Depositor
Wilson B-factor (Å^2)	55.9	Xtriage
Anisotropy	0.234	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	30414	wwPDB-VP
Average B, all atoms (Å^2)	75.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.71% 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.

4 Model quality (i)

4.1 Standard geometry (i)

MolProbity failed to run properly - this section is therefore empty.

4.2 Too-close contacts (i)

MolProbity failed to run properly - this section is therefore empty.

4.3 Torsion angles (i)

4.3.1 Protein backbone (i)

MolProbity failed to run properly - this section is therefore empty.

4.3.2 Protein sidechains (i)

MolProbity failed to run properly - this section is therefore empty.

4.3.3 RNA (i)

MolProbity failed to run properly - this section is therefore empty.

4.4 Non-standard residues in protein, DNA, RNA chains (i)

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

4.5 Carbohydrates (i)

There are no monosaccharides in this entry.

4.6 Ligand geometry (i)

6 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	$\#\|Z\| > 2$
2	HEM	D	601	-	41,50,50	1.44	3 (7%)	45,82,82	1.37	6 (13%)
2	HEM	C	601	-	41,50,50	1.45	4 (9%)	45,82,82	1.35	6 (13%)
3	Q4M	A	602	-	27,31,31	2.42	14 (51%)	31,46,46	2.91	12 (38%)
2	HEM	B	601	1	41,50,50	1.46	5 (12%)	45,82,82	1.38	6 (13%)
3	Q4M	B	602	-	27,31,31	2.01	9 (33%)	31,46,46	3.62	13 (41%)
2	HEM	A	601	-	41,50,50	1.44	3 (7%)	45,82,82	1.39	8 (17%)

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	HEM	D	601	-	-	1/12/54/54	-
2	HEM	C	601	-	-	4/12/54/54	-
3	Q4M	A	602	-	-	2/9/24/24	0/5/5/5
2	HEM	B	601	1	-	6/12/54/54	-
3	Q4M	B	602	-	-	4/9/24/24	0/5/5/5
2	HEM	A	601	-	-	6/12/54/54	-

All (38) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	602	Q4M	C20-C03	-5.22	1.46	1.52
3	A	602	Q4M	C07-C06	-4.10	1.43	1.50
2	D	601	HEM	C3C-C2C	-4.06	1.34	1.40
3	A	602	Q4M	C19-C20	4.06	1.41	1.37
2	B	601	HEM	C3C-C2C	-3.96	1.34	1.40
2	C	601	HEM	C3C-C2C	-3.95	1.34	1.40
2	A	601	HEM	C3C-C2C	-3.92	1.34	1.40
3	A	602	Q4M	C02-C03	-3.77	1.43	1.52
2	C	601	HEM	C3C-CAC	3.71	1.55	1.47
3	B	602	Q4M	C12-C13	-3.70	1.31	1.38
2	D	601	HEM	C3C-CAC	3.69	1.55	1.47
2	B	601	HEM	C3C-CAC	3.67	1.55	1.47
2	A	601	HEM	C3C-CAC	3.67	1.55	1.47

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	602	Q4M	C23-C22	-3.60	1.35	1.41
3	B	602	Q4M	O14-C13	-3.56	1.30	1.37
3	B	602	Q4M	C02-CL1	3.37	1.96	1.78
3	A	602	Q4M	O14-C13	-3.31	1.30	1.37
3	A	602	Q4M	C12-C13	-3.05	1.32	1.38
3	A	602	Q4M	C19-N05	2.99	1.45	1.39
2	D	601	HEM	CAB-C3B	2.93	1.55	1.47
2	C	601	HEM	CAB-C3B	2.92	1.55	1.47
2	A	601	HEM	CAB-C3B	2.86	1.55	1.47
3	B	602	Q4M	C11-C10	-2.84	1.36	1.41
2	B	601	HEM	CAB-C3B	2.83	1.55	1.47
3	A	602	Q4M	C04-C03	-2.77	1.47	1.53
3	B	602	Q4M	C22-N25	-2.71	1.30	1.38
3	A	602	Q4M	C16-C09	-2.68	1.35	1.42
3	B	602	Q4M	O14-C15	-2.45	1.35	1.42
3	A	602	Q4M	C21-C22	-2.38	1.36	1.42
3	A	602	Q4M	C20-C21	-2.38	1.39	1.43
3	B	602	Q4M	C19-C20	2.27	1.39	1.37
3	B	602	Q4M	C20-C21	-2.23	1.39	1.43
3	A	602	Q4M	C11-C10	-2.14	1.38	1.41
3	A	602	Q4M	C23-C24	2.12	1.41	1.36
2	B	601	HEM	CAA-C2A	2.06	1.55	1.52
2	B	601	HEM	FE-ND	2.06	2.07	1.96
3	A	602	Q4M	C10-N17	-2.03	1.32	1.38
2	C	601	HEM	CAA-C2A	2.02	1.55	1.52

All (51) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	B	602	Q4M	C07-C06-N05	12.50	133.93	118.39
3	A	602	Q4M	C07-C06-N05	10.23	131.11	118.39
3	B	602	Q4M	O18-C06-N05	-7.74	111.44	121.69
3	B	602	Q4M	C19-C20-C21	-6.95	115.10	120.48
3	B	602	Q4M	C24-C23-C22	-6.00	113.28	120.84
3	A	602	Q4M	O18-C06-N05	-5.68	114.17	121.69
3	A	602	Q4M	C03-C02-CL1	-5.31	101.39	111.47
3	A	602	Q4M	C24-C19-C20	-5.20	115.91	122.22
3	B	602	Q4M	C24-C19-C20	-4.19	117.14	122.22
3	B	602	Q4M	C23-C24-C19	3.95	127.66	119.95
3	B	602	Q4M	C03-C02-CL1	3.49	118.09	111.47
3	B	602	Q4M	C07-N17-C10	3.20	111.12	104.45
3	A	602	Q4M	C24-C23-C22	-3.05	117.00	120.84

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	B	602	Q4M	C07-C08-C09	3.01	110.31	106.55
3	A	602	Q4M	C07-N17-C10	3.01	110.73	104.45
2	B	601	HEM	C1B-NB-C4B	2.96	108.13	105.07
2	A	601	HEM	C1B-NB-C4B	2.84	108.00	105.07
3	A	602	Q4M	C12-C11-C10	-2.83	117.28	120.84
2	D	601	HEM	C4D-ND-C1D	2.75	107.91	105.07
2	C	601	HEM	C1B-NB-C4B	2.74	107.90	105.07
2	D	601	HEM	C1B-NB-C4B	2.71	107.88	105.07
2	D	601	HEM	C4B-CHC-C1C	2.68	126.10	122.56
2	C	601	HEM	C4D-ND-C1D	2.68	107.84	105.07
3	A	602	Q4M	C15-O14-C13	-2.65	111.77	117.51
2	C	601	HEM	C4B-CHC-C1C	2.60	125.99	122.56
2	B	601	HEM	C4C-CHD-C1D	2.58	125.97	122.56
2	A	601	HEM	C4C-CHD-C1D	2.52	125.89	122.56
2	B	601	HEM	C4D-ND-C1D	2.51	107.67	105.07
3	B	602	Q4M	C12-C11-C10	-2.50	117.69	120.84
2	B	601	HEM	C4B-CHC-C1C	2.50	125.86	122.56
2	A	601	HEM	C4D-ND-C1D	2.47	107.63	105.07
3	B	602	Q4M	C08-C07-C06	2.38	136.65	128.33
2	D	601	HEM	C4C-CHD-C1D	2.37	125.69	122.56
3	B	602	Q4M	O18-C06-C07	-2.32	114.52	119.00
2	B	601	HEM	CMA-C3A-C4A	-2.32	124.90	128.46
3	A	602	Q4M	O18-C06-C07	-2.30	114.56	119.00
2	A	601	HEM	C4B-CHC-C1C	2.30	125.59	122.56
2	A	601	HEM	CMA-C3A-C4A	-2.30	124.93	128.46
2	C	601	HEM	C4C-CHD-C1D	2.30	125.59	122.56
2	A	601	HEM	CAA-CBA-CGA	-2.24	107.49	113.76
2	A	601	HEM	C3B-C2B-C1B	2.19	108.11	106.49
2	B	601	HEM	C3B-C2B-C1B	2.19	108.11	106.49
3	A	602	Q4M	O14-C13-C16	-2.16	118.54	124.43
2	C	601	HEM	CMA-C3A-C4A	-2.15	125.17	128.46
3	B	602	Q4M	C03-C20-C21	2.14	135.18	129.84
3	A	602	Q4M	C23-C24-C19	2.14	124.13	119.95
2	C	601	HEM	C3B-C2B-C1B	2.13	108.07	106.49
3	A	602	Q4M	C04-N05-C19	-2.12	106.77	109.21
2	D	601	HEM	C3B-C2B-C1B	2.05	108.01	106.49
2	D	601	HEM	CMA-C3A-C4A	-2.04	125.33	128.46
2	A	601	HEM	CMC-C2C-C3C	2.03	128.47	124.68

There are no chirality outliers.

All (23) torsion outliers are listed below:

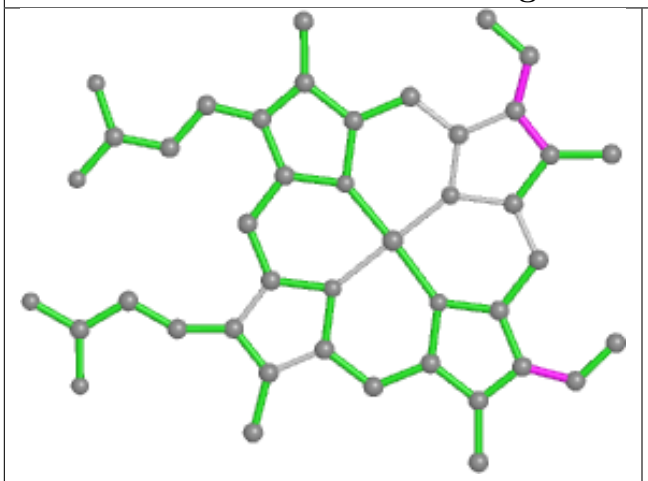
Mol	Chain	Res	Type	Atoms
2	A	601	HEM	C1A-C2A-CAA-CBA
2	A	601	HEM	C3A-C2A-CAA-CBA
2	B	601	HEM	C1A-C2A-CAA-CBA
2	B	601	HEM	C3A-C2A-CAA-CBA
3	B	602	Q4M	CL1-C02-C03-C04
3	A	602	Q4M	C16-C13-O14-C15
3	B	602	Q4M	C16-C13-O14-C15
3	A	602	Q4M	C12-C13-O14-C15
3	B	602	Q4M	C12-C13-O14-C15
3	B	602	Q4M	CL1-C02-C03-C20
2	C	601	HEM	C1A-C2A-CAA-CBA
2	C	601	HEM	C3A-C2A-CAA-CBA
2	C	601	HEM	CAA-CBA-CGA-O2A
2	C	601	HEM	CAA-CBA-CGA-O1A
2	A	601	HEM	CAA-CBA-CGA-O2A
2	A	601	HEM	CAD-CBD-CGD-O1D
2	B	601	HEM	CAD-CBD-CGD-O2D
2	B	601	HEM	CAA-CBA-CGA-O2A
2	B	601	HEM	CAA-CBA-CGA-O1A
2	D	601	HEM	C4D-C3D-CAD-CBD
2	A	601	HEM	CAA-CBA-CGA-O1A
2	A	601	HEM	CAD-CBD-CGD-O2D
2	B	601	HEM	CAD-CBD-CGD-O1D

There are no ring outliers.

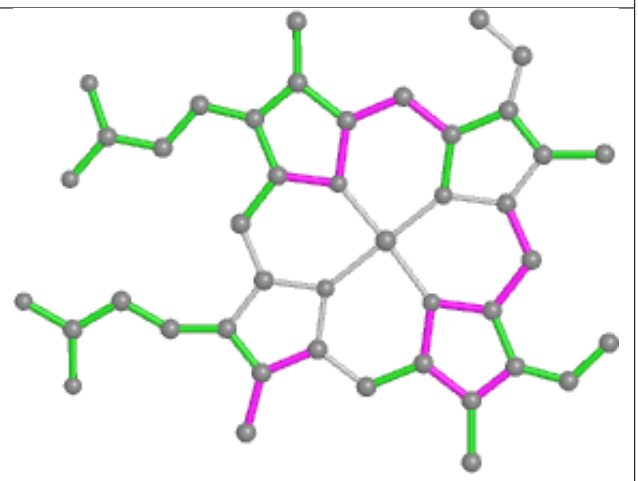
No monomer is involved in short contacts.

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.

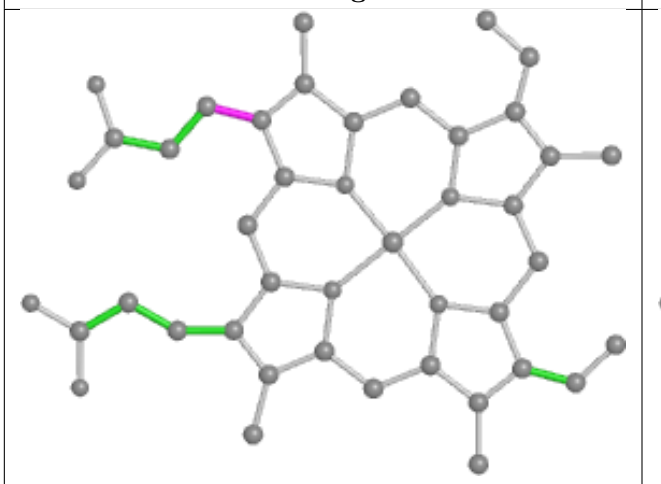
Ligand HEM D 601



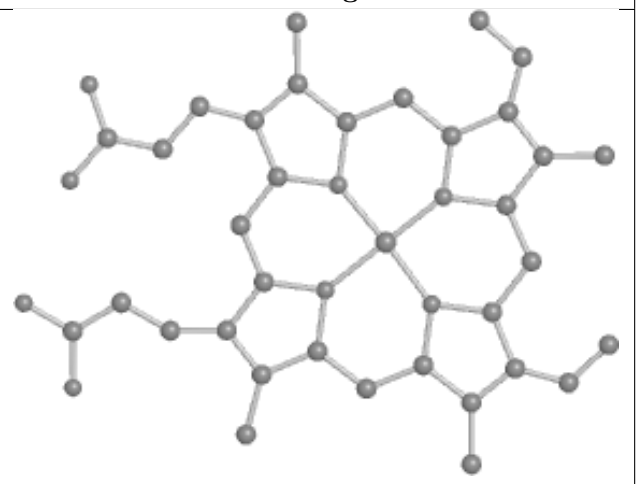
Bond lengths



Bond angles

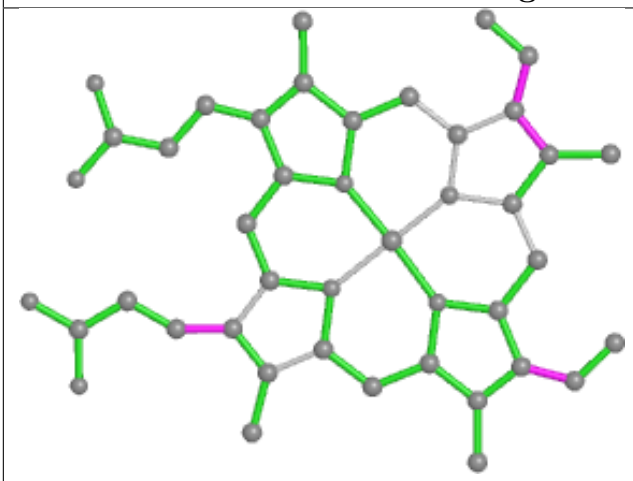


Torsions

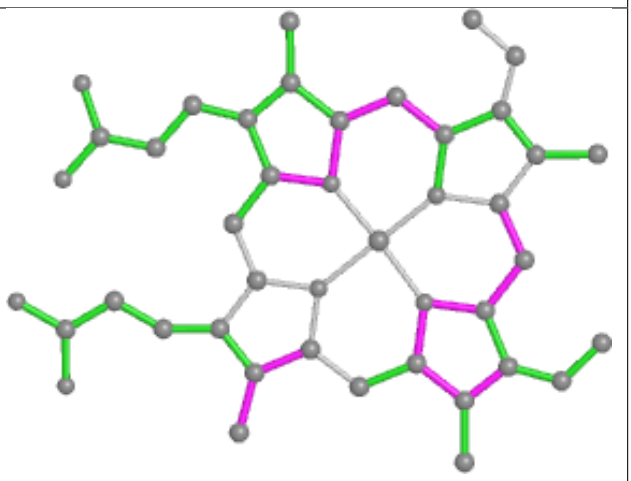


Rings

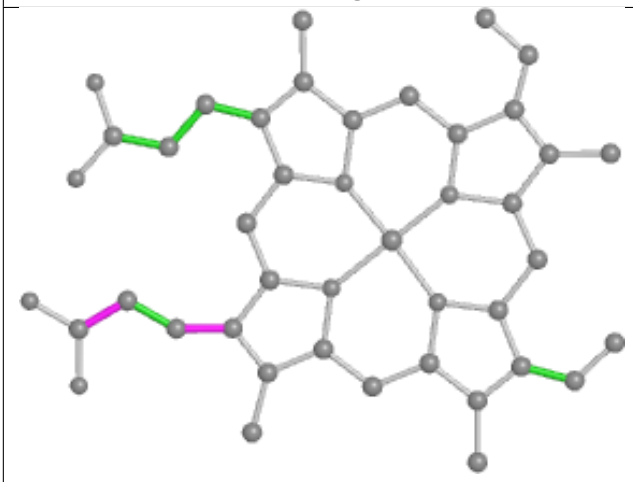
Ligand HEM C 601



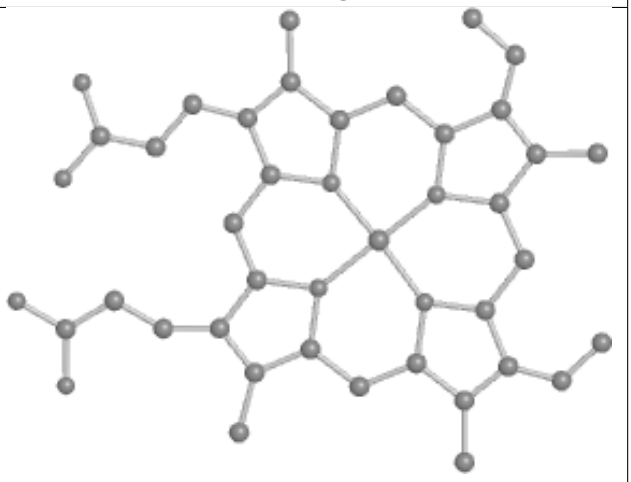
Bond lengths



Bond angles

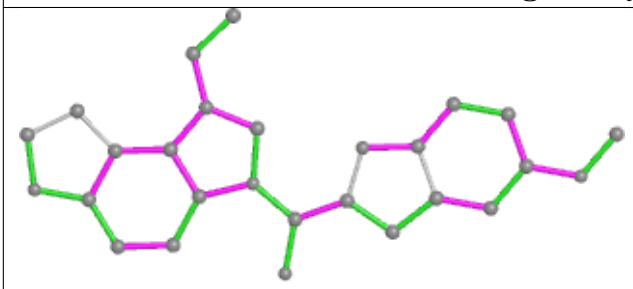


Torsions

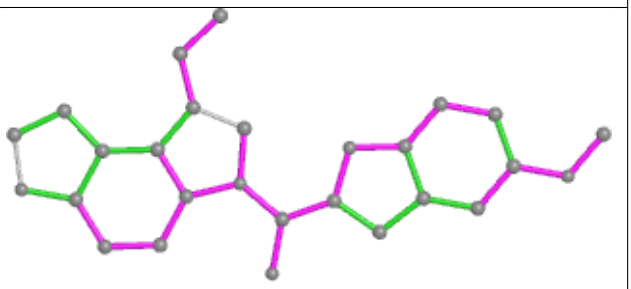


Rings

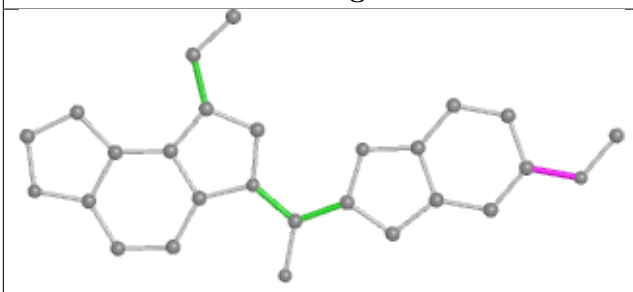
Ligand Q4M A 602



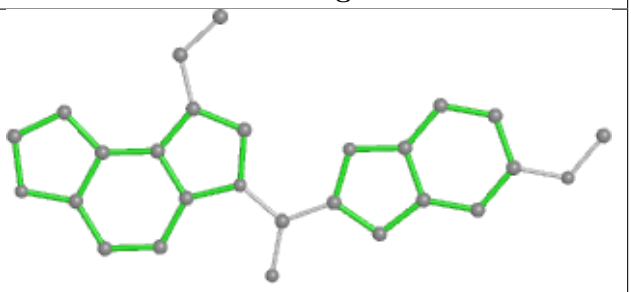
Bond lengths



Bond angles

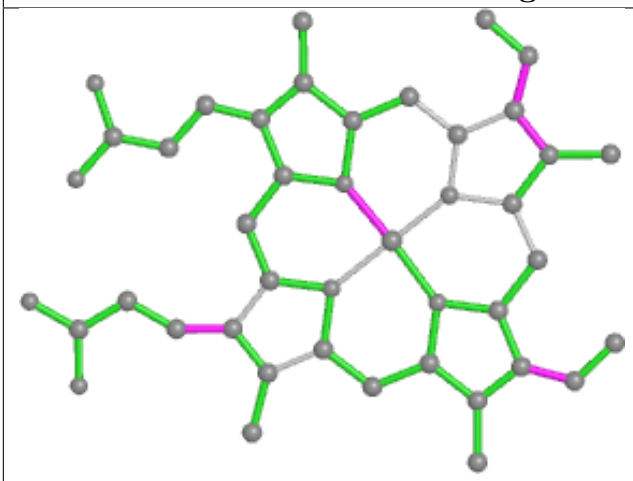


Torsions

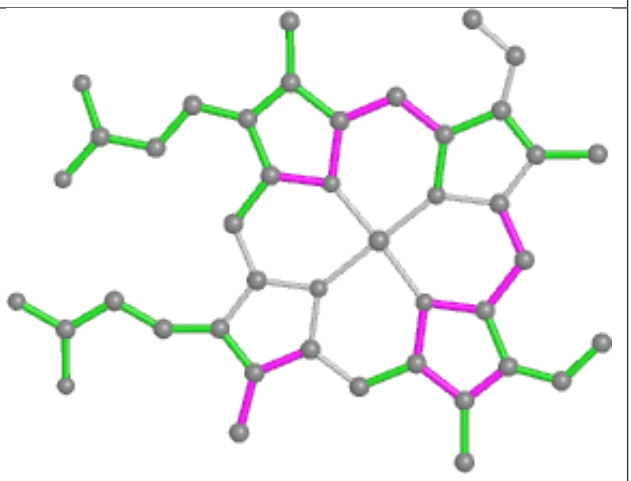


Rings

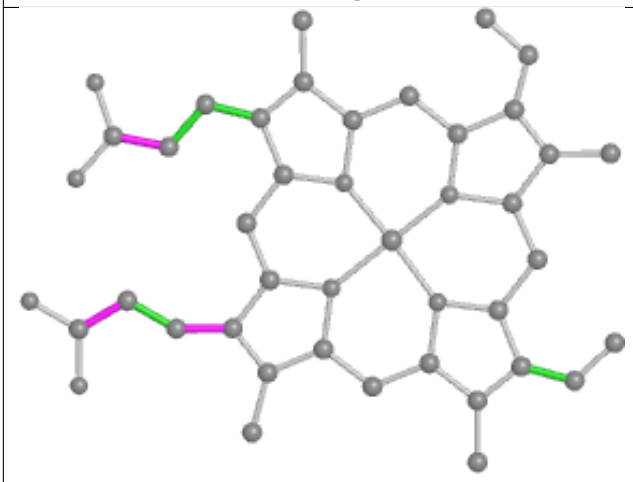
Ligand HEM B 601



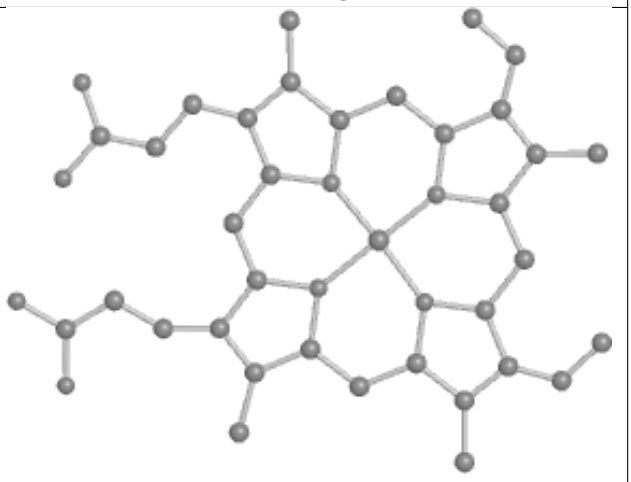
Bond lengths



Bond angles

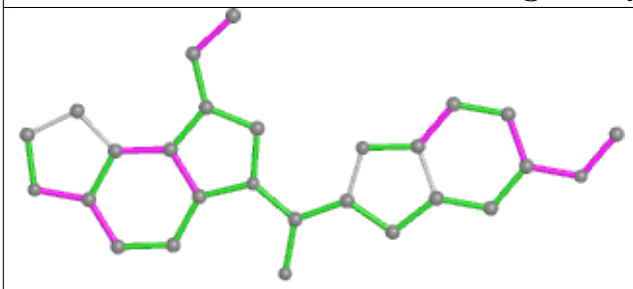


Torsions

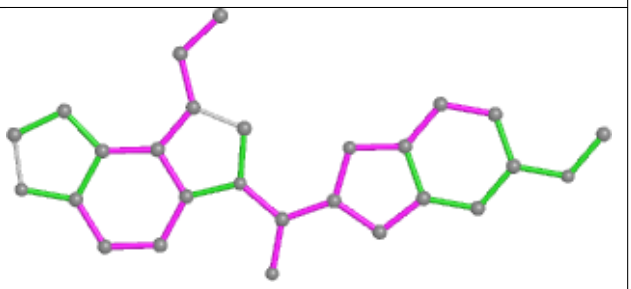


Rings

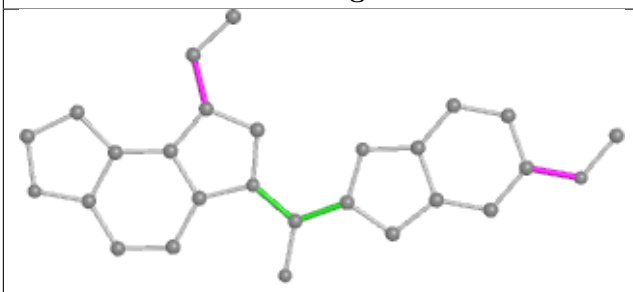
Ligand Q4M B 602



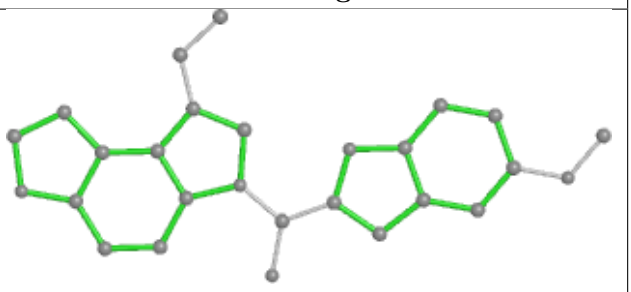
Bond lengths



Bond angles

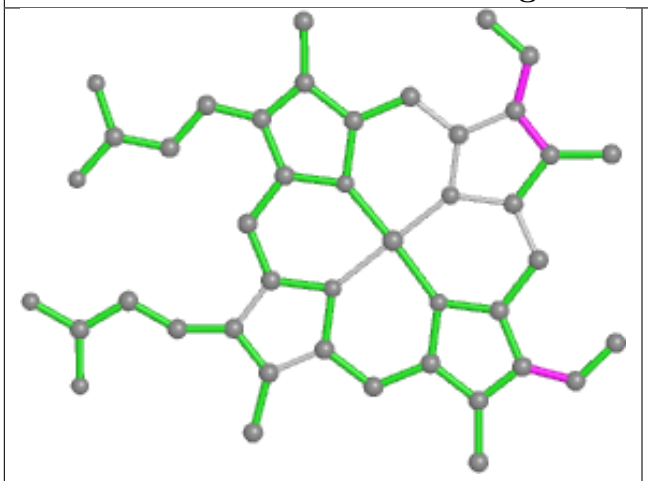


Torsions

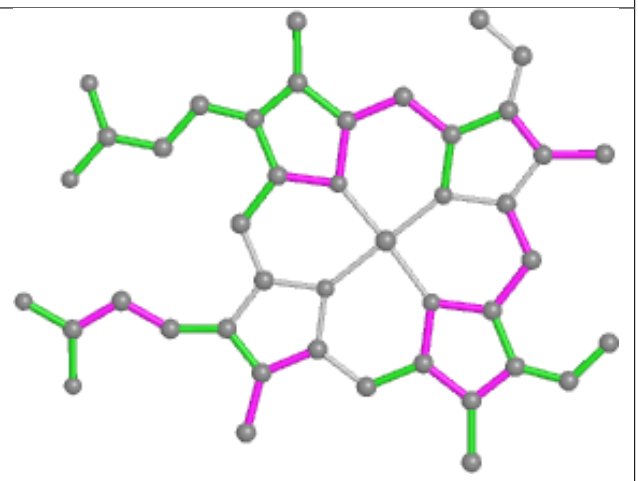


Rings

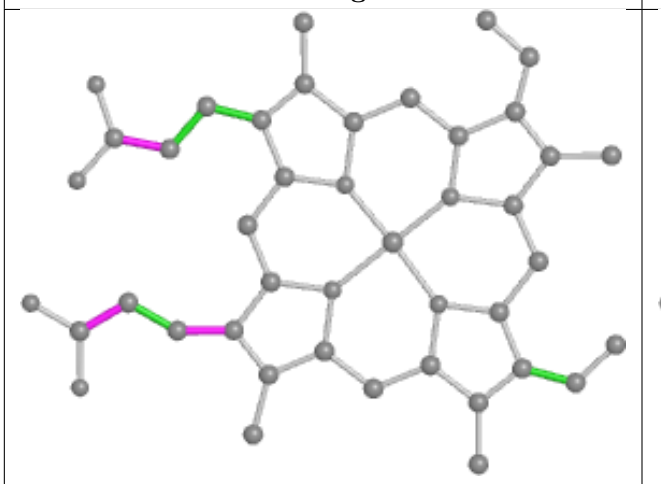
Ligand HEM A 601



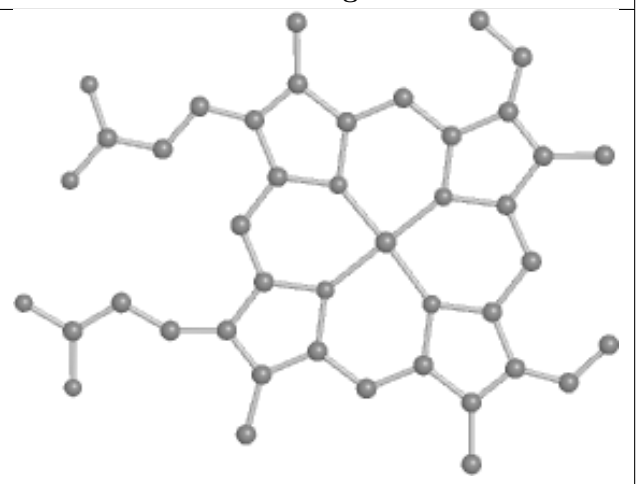
Bond lengths



Bond angles



Torsions



Rings

4.7 Other polymers (i)

There are no such residues in this entry.

4.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

5 Fit of model and data (i)

5.1 Protein, DNA and RNA chains (i)

EDS failed to run properly - this section is therefore empty.

5.2 Non-standard residues in protein, DNA, RNA chains (i)

EDS failed to run properly - this section is therefore empty.

5.3 Carbohydrates (i)

EDS failed to run properly - this section is therefore empty.

5.4 Ligands (i)

EDS failed to run properly - this section is therefore empty.

5.5 Other polymers (i)

EDS failed to run properly - this section is therefore empty.