wwPDB EM Validation Summary Report

Mar 18, 2024 – 04:27 PM JST

PDB ID	:	6JNU
Title	:	Catalase structure determined by eEFD (dataset 2)
Authors	:	Yonekura, K.; Maki-Yonekura, S.
Deposited on	:	2019-03-18
Resolution	:	3.00 Å(reported)
Based on initial model	:	3NWL

This is a wwPDB EM Validation Summary Report for a publicly released PDB/EMDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/EMValidationReportHelp
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 (i)) were used in the production of this report:

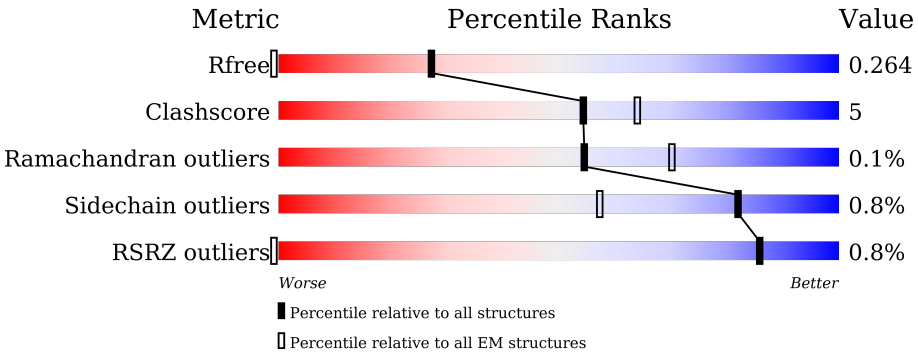
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance

The following experimental techniques were used to determine the structure:
ELECTRON CRYSTALLOGRAPHY

The reported resolution of this entry is 3.00 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)	EM structures (#Entries)
R_{free}	130704	0
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826
RSRZ outliers	127900	0

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the 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%

Mol	Chain	Length	Quality of chain
1	A	527	81% 13% 5%
1	B	527	84% 10% 5%
1	C	527	83% 12% 5%
1	D	527	83% 11% 5%

2 Entry composition

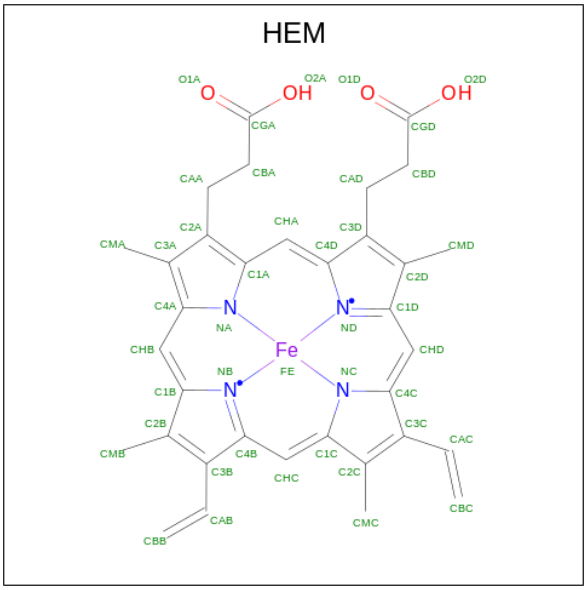
There are 4 unique types of molecules in this entry. The entry contains 32072 atoms, of which 15603 are hydrogens and 0 are deuteriums.

In the tables below, 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 Catalase.

Mol	Chain	Residues	Atoms						AltConf	Trace
1	A	499	Total	C	H	N	O	S	0	0
			7863	2548	3846	717	738	14		
1	B	499	Total	C	H	N	O	S	0	0
			7861	2548	3844	717	738	14		
1	C	499	Total	C	H	N	O	S	0	0
			7861	2548	3844	717	738	14		
1	D	499	Total	C	H	N	O	S	0	0
			7862	2548	3845	717	738	14		

- Molecule 2 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (formula: $C_{34}H_{32}FeN_4O_4$).



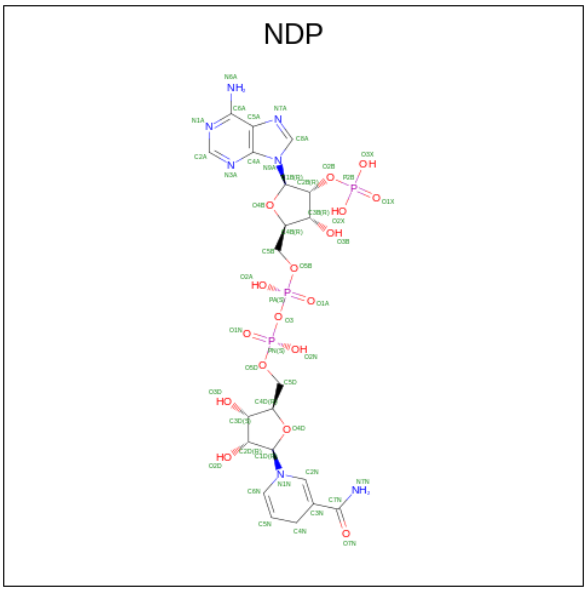
Mol	Chain	Residues	Atoms						AltConf
2	A	1	Total	C	Fe	H	N	O	0
			73	34	1	30	4	4	
2	B	1	Total	C	Fe	H	N	O	0
			73	34	1	30	4	4	

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms						AltConf
2	C	1	Total	C	Fe	H	N	O	0
			73	34	1	30	4	4	
2	D	1	Total	C	Fe	H	N	O	0
			73	34	1	30	4	4	

- Molecule 3 is NADPH DIHYDRO-NICOTINAMIDE-ADENINE-DINUCLEOTIDE PHOSPHATE (three-letter code: NDP) (formula: $C_{21}H_{30}N_7O_{17}P_3$).



Mol	Chain	Residues	Atoms						AltConf
3	A	1	Total	C	H	N	O	P	0
			74	21	26	7	17	3	
3	B	1	Total	C	H	N	O	P	0
			74	21	26	7	17	3	
3	C	1	Total	C	H	N	O	P	0
			74	21	26	7	17	3	
3	D	1	Total	C	H	N	O	P	0
			74	21	26	7	17	3	

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		AltConf
4	A	6	Total	O	0
			6	6	
4	B	8	Total	O	0
			8	8	
4	C	10	Total	O	0
			10	10	

Continued on next page...

Continued from previous page...

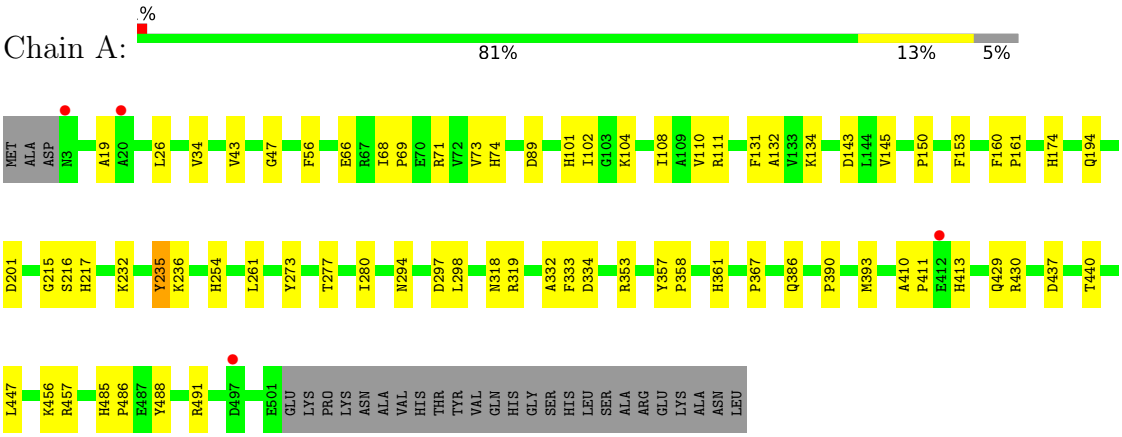
Mol	Chain	Residues	Atoms		AltConf
4	D	13	Total	O	0
			13	13	

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. 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. 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.

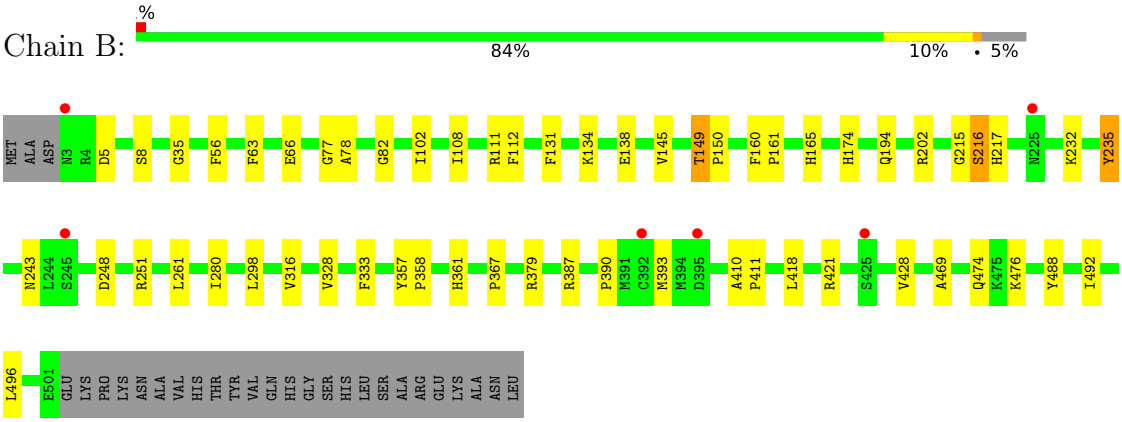
- Molecule 1: Catalase

Chain A:



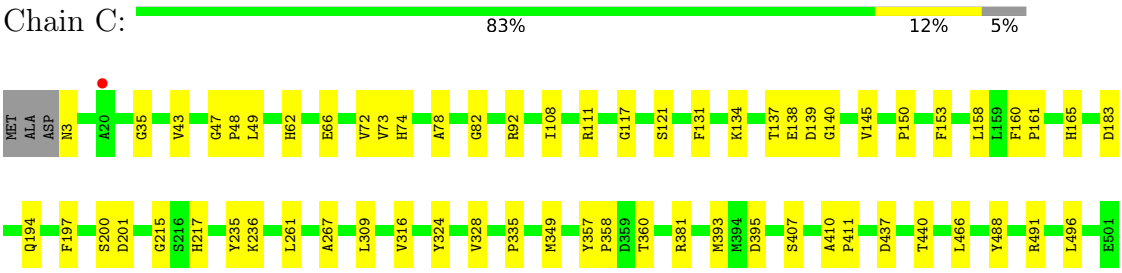
- Molecule 1: Catalase

Chain B:



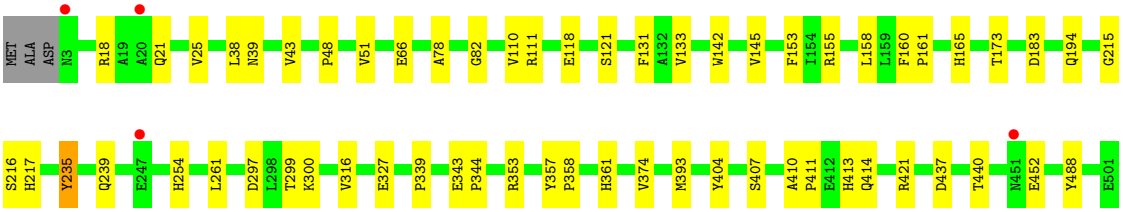
- Molecule 1: Catalase

Chain C:



• Molecule 1: Catalase

Chain D:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	68.94Å 172.14Å 201.38Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	43.62 – 3.00 43.62 – 2.99	Depositor EDS
% Data completeness (in resolution range)	72.0 (43.62-3.00) 65.6 (43.62-2.99)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.74 (at 3.01Å)	Xtriage
Refinement program	PHENIX (1.14_3260: ???)	Depositor
R, R_{free}	0.207 , 0.251 0.227 , 0.264	Depositor DCC
R_{free} test set	2000 reflections (5.79%)	wwPDB-VP
Wilson B-factor (Å^2)	19.3	Xtriage
Anisotropy	1.347	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.23 , 4.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.44$, $< L^2 > = 0.26$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.84	EDS
Total number of atoms	32072	wwPDB-VP
Average B, all atoms (Å^2)	21.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.07% 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: HEM, NDP

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	0.38	0/4137	0.55	0/5619
1	B	0.37	0/4137	0.54	0/5619
1	C	0.38	0/4137	0.54	2/5619 (0.0%)
1	D	0.39	1/4137 (0.0%)	0.55	0/5619
All	All	0.38	1/16548 (0.0%)	0.55	2/22476 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	452	GLU	CB-CG	-5.59	1.41	1.52

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	491	ARG	NE-CZ-NH1	-8.63	115.99	120.30
1	C	111	ARG	NE-CZ-NH2	-5.18	117.71	120.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

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	4017	3846	3843	51	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	4017	3844	3843	42	0
1	C	4017	3844	3843	49	0
1	D	4017	3845	3843	43	0
2	A	43	30	30	1	0
2	B	43	30	30	5	0
2	C	43	30	30	6	0
2	D	43	30	30	4	0
3	A	48	26	26	2	0
3	B	48	26	25	3	0
3	C	48	26	26	2	0
3	D	48	26	26	0	0
4	A	6	0	0	7	0
4	B	8	0	0	4	0
4	C	10	0	0	6	0
4	D	13	0	0	6	0
All	All	16469	15603	15595	173	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 5.

The worst 5 of 173 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)
3:B:602:NDP:O2B	4:B:701:HOH:O	1.82	0.98
1:A:19:ALA:O	4:A:701:HOH:O	1.88	0.91
1:A:429:GLN:OE1	4:A:702:HOH:O	1.91	0.86
1:D:38:LEU:O	4:D:701:HOH:O	1.95	0.85
3:B:602:NDP:P2B	4:B:701:HOH:O	2.34	0.83

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 PDB entries followed by that with respect to all EM entries.

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	497/527 (94%)	468 (94%)	29 (6%)	0	100 100
1	B	497/527 (94%)	467 (94%)	29 (6%)	1 (0%)	47 82
1	C	497/527 (94%)	470 (95%)	27 (5%)	0	100 100
1	D	497/527 (94%)	475 (96%)	21 (4%)	1 (0%)	47 82
All	All	1988/2108 (94%)	1880 (95%)	106 (5%)	2 (0%)	54 85

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	216	SER
1	B	216	SER

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all EM entries.

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	431/454 (95%)	428 (99%)	3 (1%)	84 94
1	B	431/454 (95%)	426 (99%)	5 (1%)	71 90
1	C	431/454 (95%)	428 (99%)	3 (1%)	84 94
1	D	431/454 (95%)	428 (99%)	3 (1%)	84 94
All	All	1724/1816 (95%)	1710 (99%)	14 (1%)	82 93

5 of 14 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	488	TYR
1	C	131	PHE
1	D	488	TYR
1	D	131	PHE
1	D	235	TYR

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

Mol	Chain	Res	Type
1	B	304	HIS
1	D	254	HIS

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

8 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
3	NDP	D	602	-	45,52,52	2.72	11 (24%)	53,80,80	1.84	14 (26%)
2	HEM	C	601	1	41,50,50	1.48	4 (9%)	45,82,82	1.59	9 (20%)
2	HEM	D	601	1	41,50,50	1.47	5 (12%)	45,82,82	1.28	5 (11%)
3	NDP	B	602	-	45,52,52	2.86	6 (13%)	53,80,80	2.03	18 (33%)
3	NDP	C	602	-	45,52,52	2.68	7 (15%)	53,80,80	1.85	11 (20%)
2	HEM	B	601	1	41,50,50	1.44	5 (12%)	45,82,82	1.47	9 (20%)
2	HEM	A	601	1	41,50,50	1.50	6 (14%)	45,82,82	1.55	6 (13%)
3	NDP	A	602	-	45,52,52	2.34	7 (15%)	53,80,80	1.89	14 (26%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the

Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	NDP	D	602	-	-	6/30/77/77	0/5/5/5
2	HEM	C	601	1	-	5/12/54/54	-
2	HEM	D	601	1	-	4/12/54/54	-
3	NDP	B	602	-	-	10/30/77/77	0/5/5/5
3	NDP	C	602	-	-	4/30/77/77	0/5/5/5
2	HEM	B	601	1	-	3/12/54/54	-
2	HEM	A	601	1	-	2/12/54/54	-
3	NDP	A	602	-	-	9/30/77/77	0/5/5/5

The worst 5 of 51 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	602	NDP	P2B-O2B	16.83	1.91	1.59
3	C	602	NDP	P2B-O2B	15.26	1.88	1.59
3	D	602	NDP	P2B-O2B	15.20	1.88	1.59
3	A	602	NDP	P2B-O2B	11.96	1.81	1.59
3	A	602	NDP	PN-O5D	5.32	1.80	1.59

The worst 5 of 86 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	B	602	NDP	PN-O3-PA	-7.64	106.61	132.83
3	A	602	NDP	PN-O3-PA	-7.02	108.75	132.83
3	C	602	NDP	PN-O3-PA	-6.91	109.11	132.83
3	D	602	NDP	PN-O3-PA	-6.31	111.18	132.83
2	A	601	HEM	CBA-CAA-C2A	-3.87	106.02	112.62

There are no chirality outliers.

5 of 43 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	602	NDP	C5B-O5B-PA-O1A
3	B	602	NDP	PN-O3-PA-O5B
3	D	602	NDP	C2N-C3N-C7N-N7N
3	B	602	NDP	O4B-C4B-C5B-O5B
3	B	602	NDP	C3B-C4B-C5B-O5B

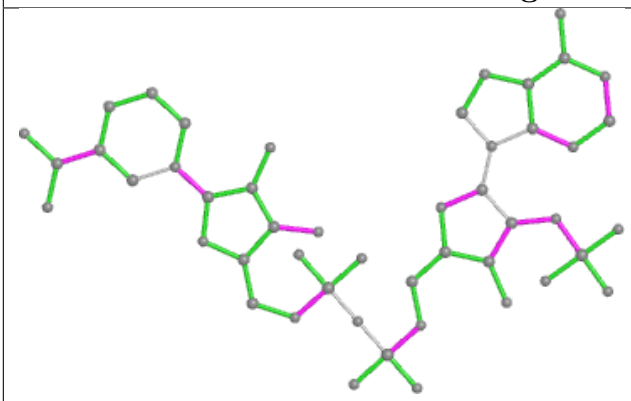
There are no ring outliers.

7 monomers are involved in 23 short contacts:

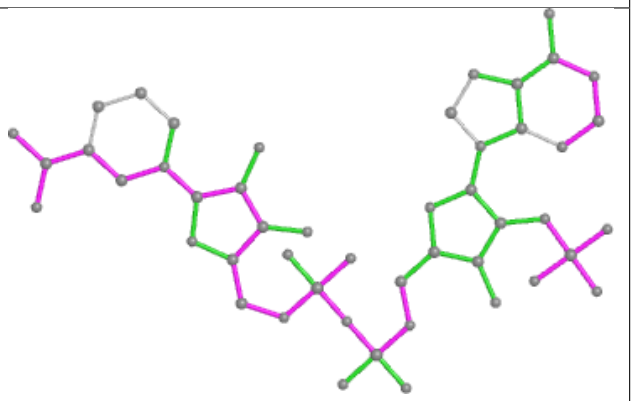
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	601	HEM	6	0
2	D	601	HEM	4	0
3	B	602	NDP	3	0
3	C	602	NDP	2	0
2	B	601	HEM	5	0
2	A	601	HEM	1	0
3	A	602	NDP	2	0

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.

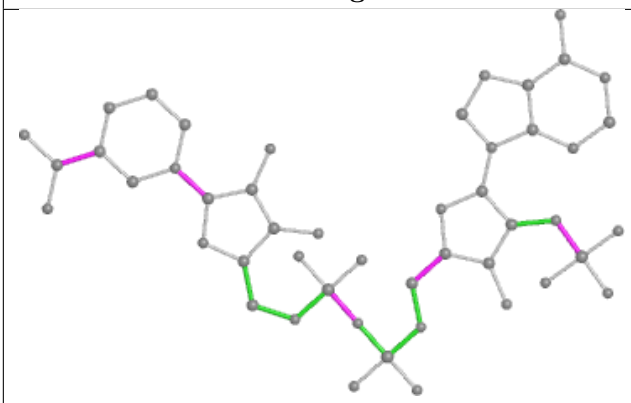
Ligand NDP D 602



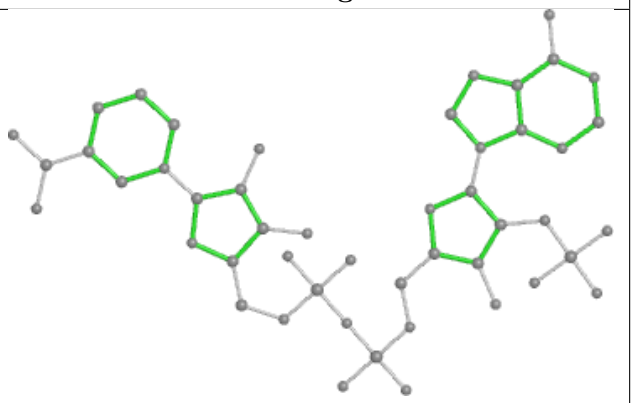
Bond lengths



Bond angles

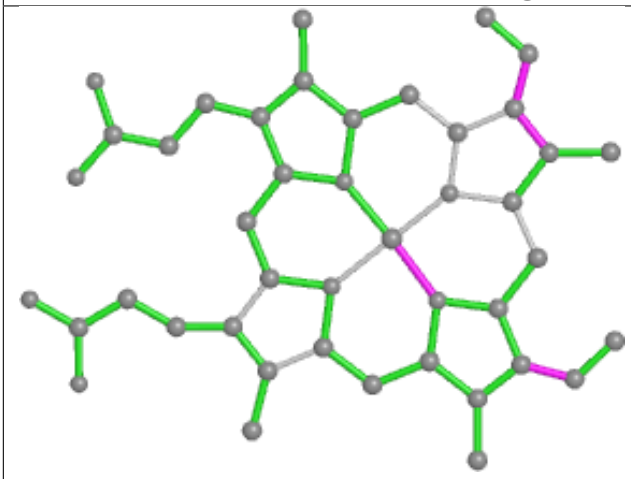


Torsions

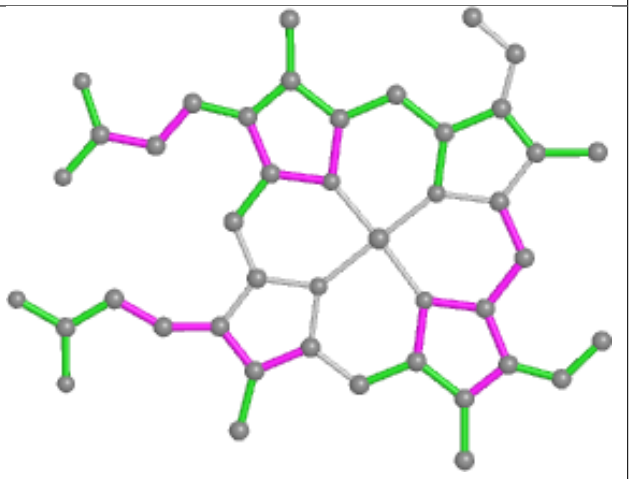


Rings

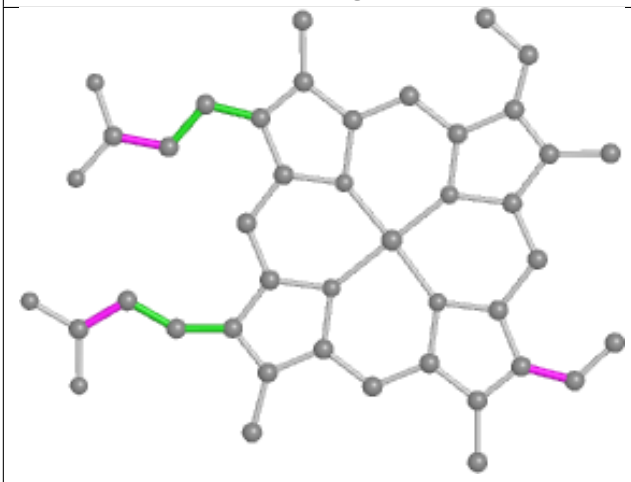
Ligand HEM C 601



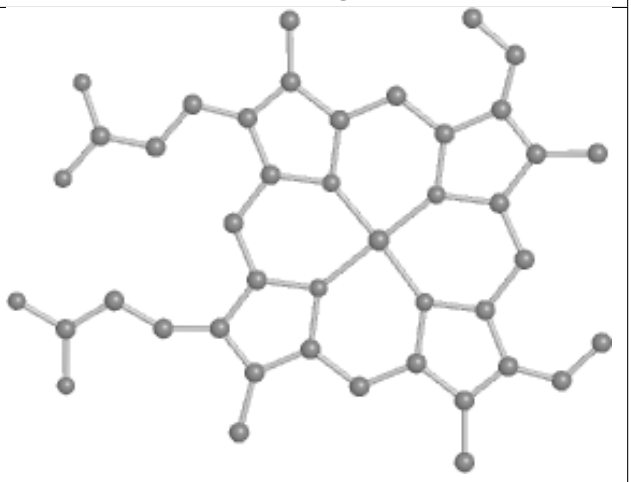
Bond lengths



Bond angles

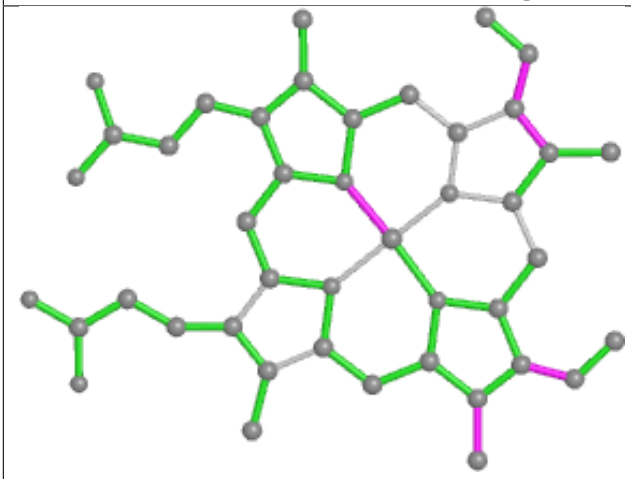


Torsions

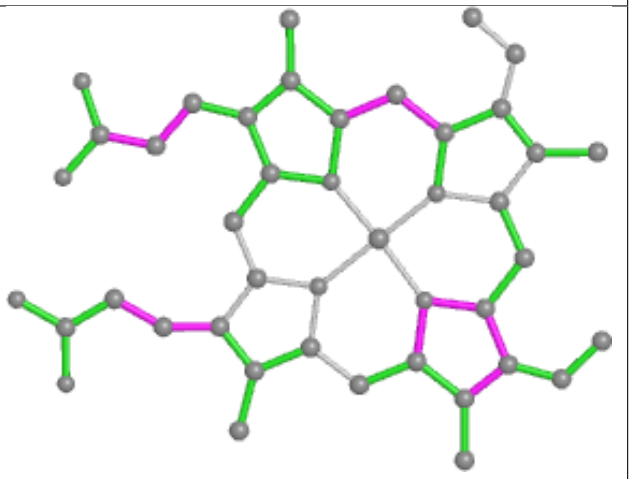


Rings

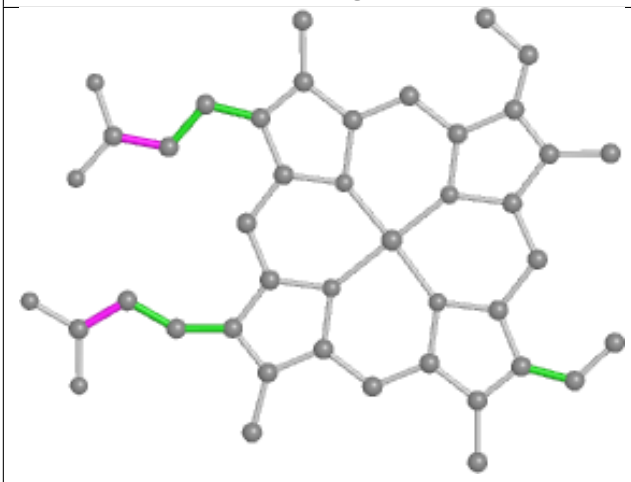
Ligand HEM D 601



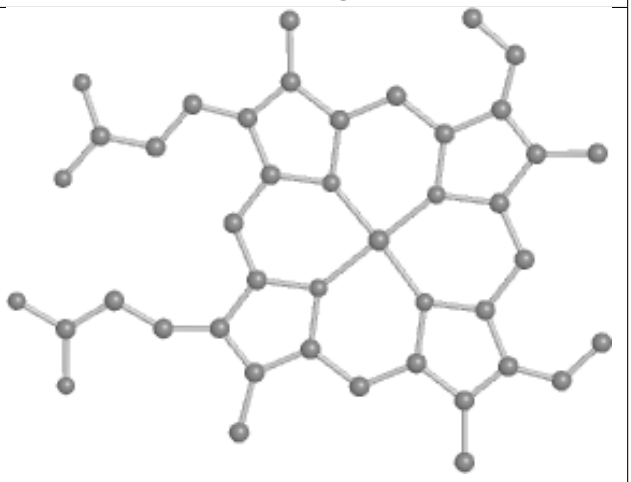
Bond lengths



Bond angles

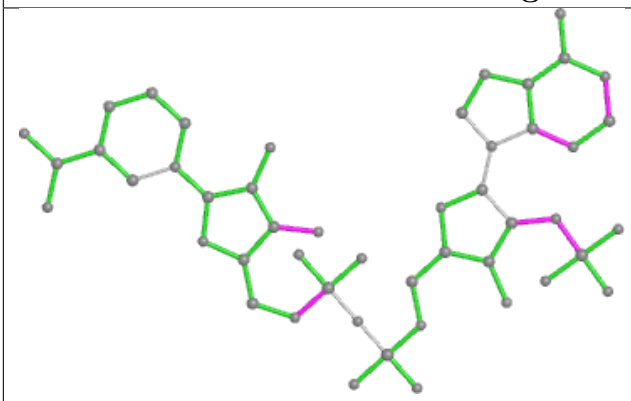


Torsions

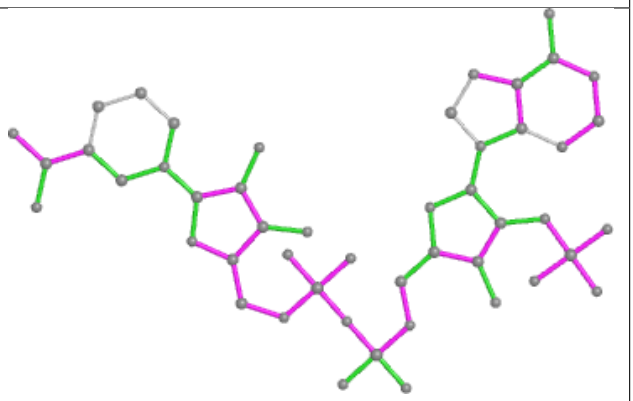


Rings

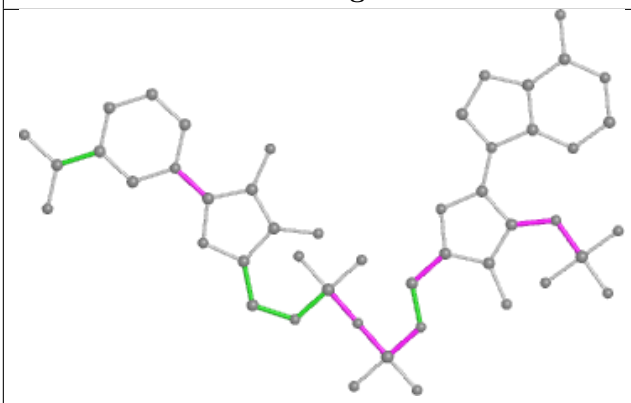
Ligand NDP B 602



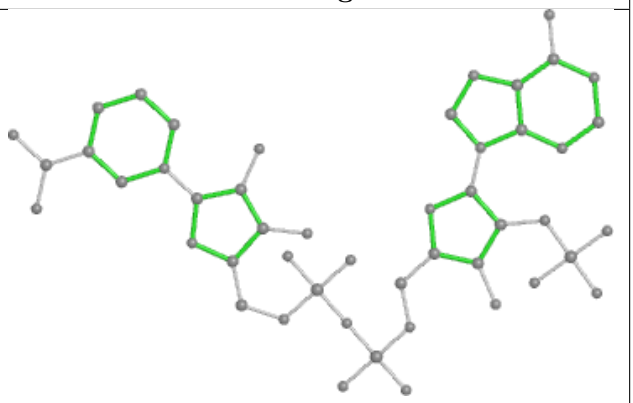
Bond lengths



Bond angles

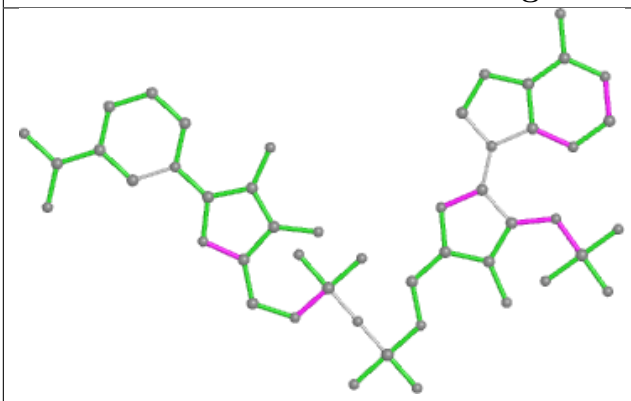


Torsions

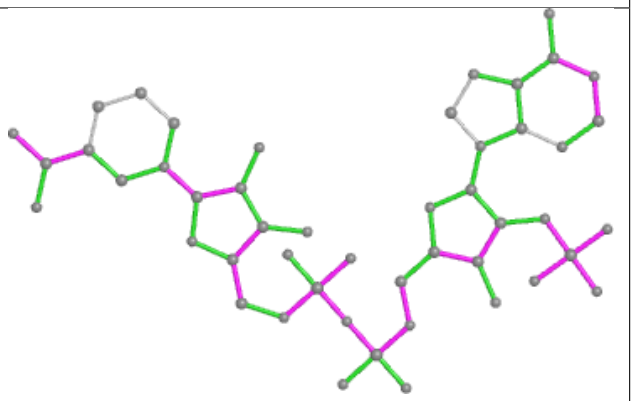


Rings

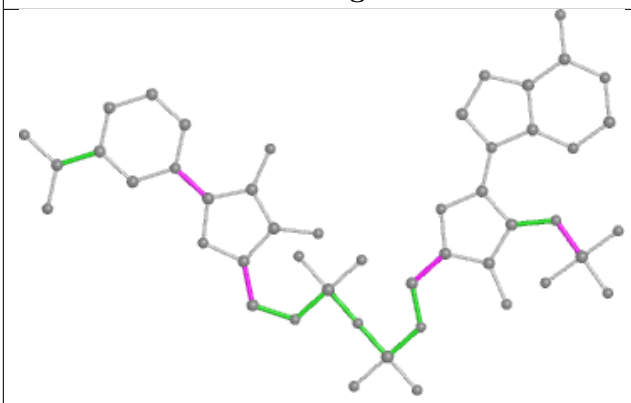
Ligand NDP C 602



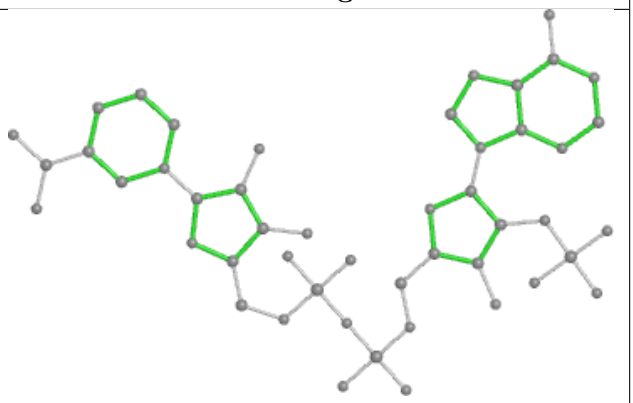
Bond lengths



Bond angles

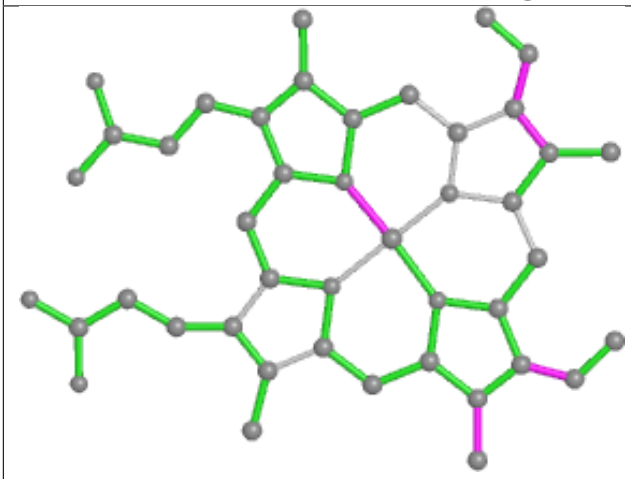


Torsions

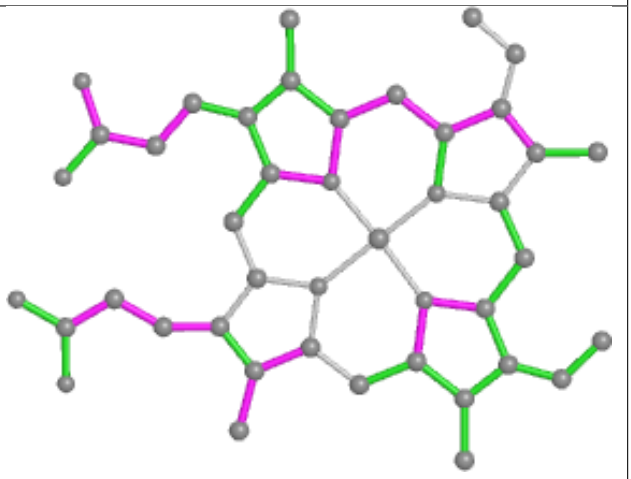


Rings

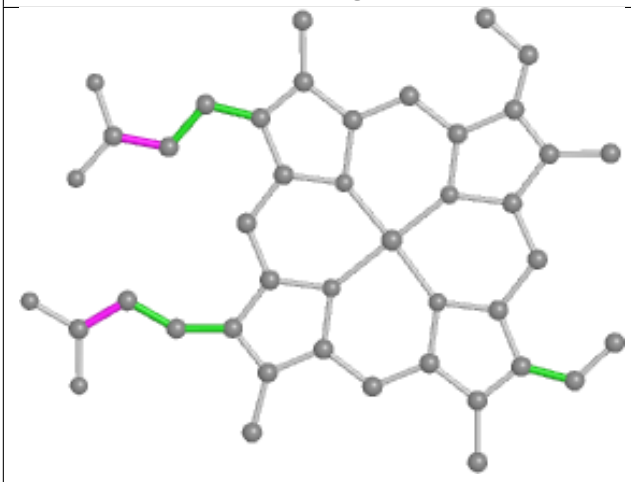
Ligand HEM B 601



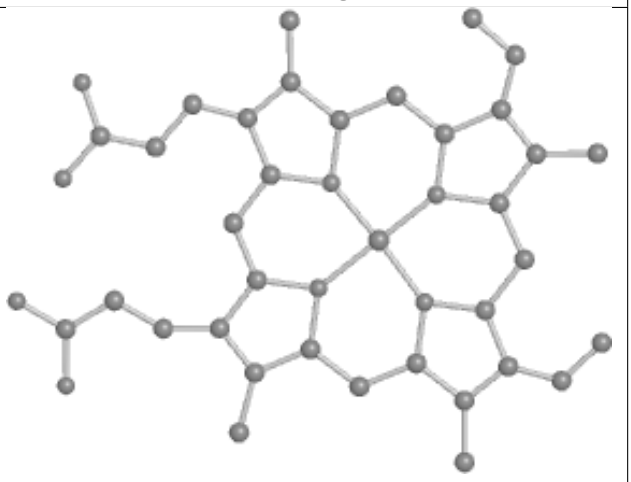
Bond lengths



Bond angles

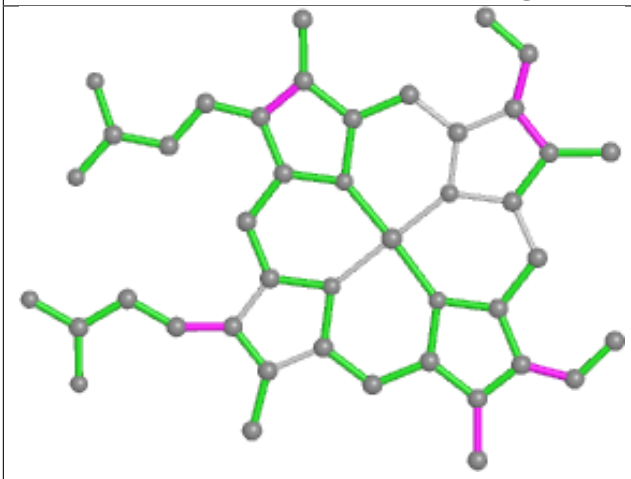


Torsions

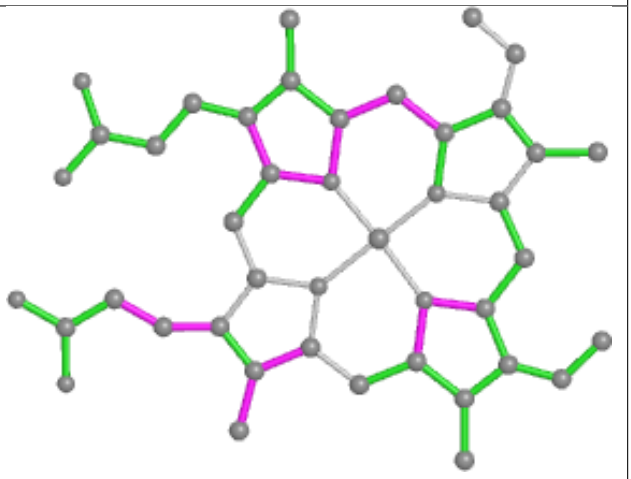


Rings

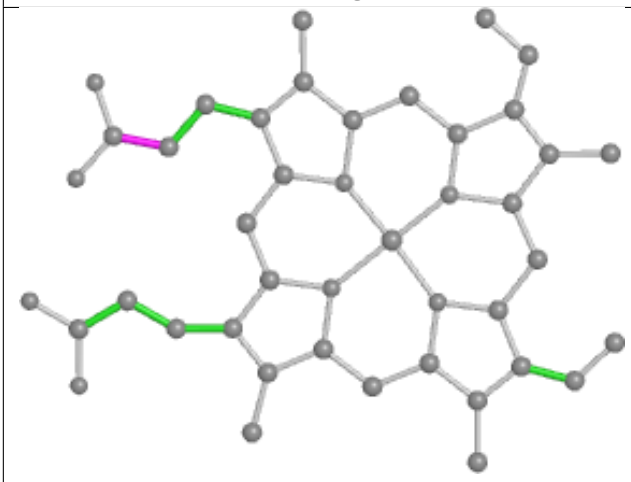
Ligand HEM A 601



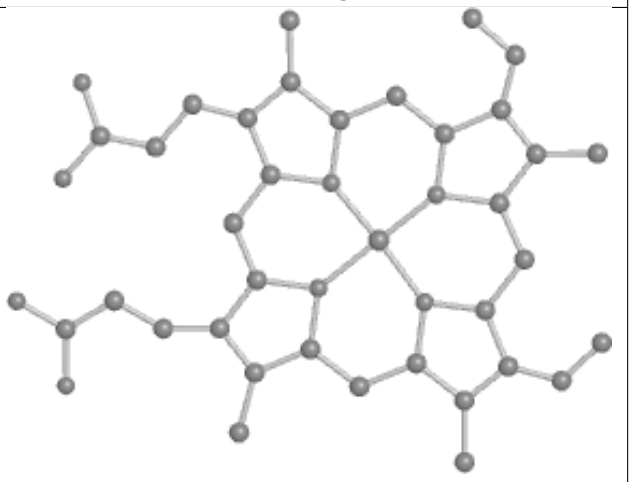
Bond lengths



Bond angles

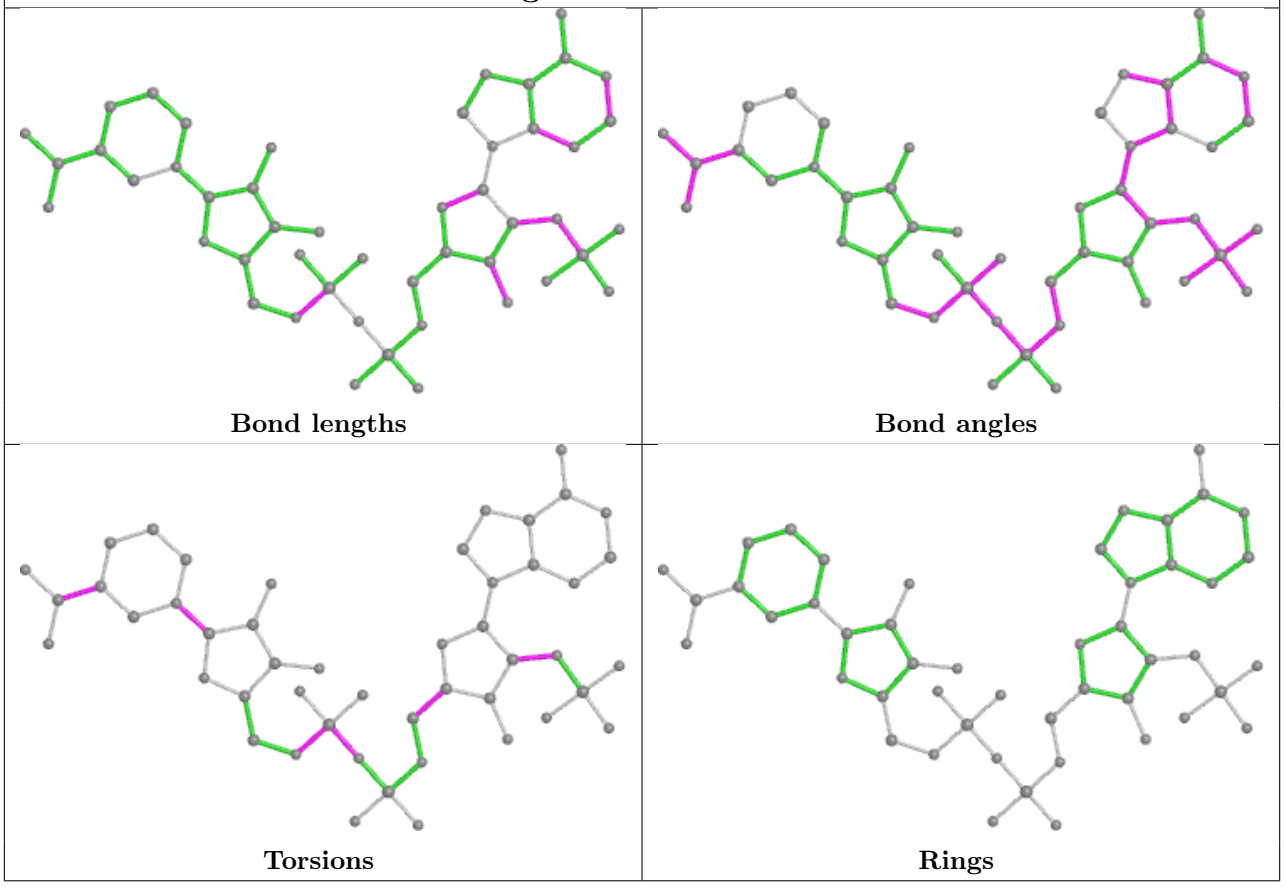


Torsions



Rings

Ligand NDP A 602



5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.