wwPDB X-ray Structure Validation Summary Report (i)

Oct 16, 2023 – 01:46 AM EDT

PDB ID	:	8ED4
Title	:	Structure of the complex between the arsenite oxidase and its native electron acceptor cytochrome c552 from Pseudorhizobium sp. str. NT-26
Authors	:	Maher, M.J.; Poddar, N.
Deposited on	:	2022-09-03
Resolution	:	2.25 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 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:

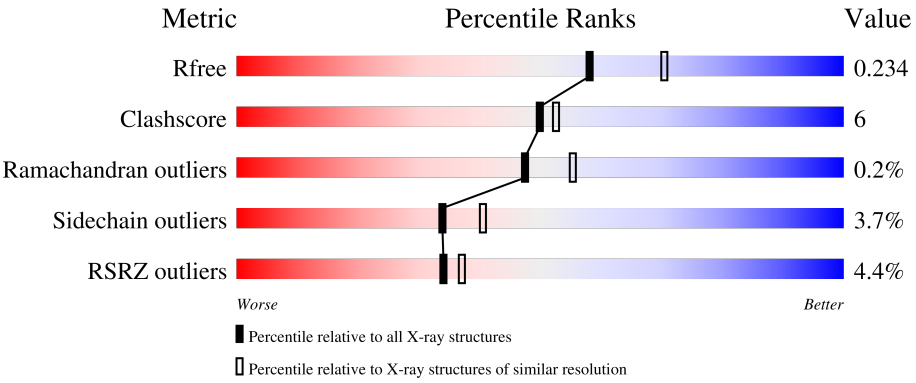
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.25 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	1377 (2.26-2.26)
Clashscore	141614	1487 (2.26-2.26)
Ramachandran outliers	138981	1449 (2.26-2.26)
Sidechain outliers	138945	1450 (2.26-2.26)
RSRZ outliers	127900	1356 (2.26-2.26)

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	844	89% 10%
1	C	844	% 88% 11%
1	E	844	88% 11%
1	G	844	90% 9%
2	B	162	3% 77% 19%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	D	162	9% 69% 12% 19%
2	F	162	% 72% 8% 19%
2	H	162	72% 9% 19%
3	I	117	21% 74% 16% 10%
3	J	117	51% 69% 15% 13%
3	K	117	31% 70% 17% 12%
3	L	117	32% 68% 16% 13%

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	PG4	G	2006	-	-	X	-

2 Entry composition (i)

There are 13 unique types of molecules in this entry. The entry contains 35834 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 AroA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	843	Total	C	N	O	S	0	5	0
			6577	4110	1174	1255	38			
1	C	843	Total	C	N	O	S	0	5	0
			6573	4108	1168	1258	39			
1	E	843	Total	C	N	O	S	0	12	0
			6624	4140	1181	1265	38			
1	G	843	Total	C	N	O	S	0	9	0
			6597	4124	1175	1261	37			

- Molecule 2 is a protein called AroB.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	132	Total	C	N	O	S	0	0	0
			995	629	166	198	2			
2	D	132	Total	C	N	O	S	0	0	0
			995	629	166	198	2			
2	F	132	Total	C	N	O	S	0	0	0
			995	629	166	198	2			
2	H	131	Total	C	N	O	S	0	0	0
			990	626	165	197	2			

There are 108 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	14	HIS	-	expression tag	UNP Q6VAL9
B	15	HIS	-	expression tag	UNP Q6VAL9
B	16	HIS	-	expression tag	UNP Q6VAL9
B	17	HIS	-	expression tag	UNP Q6VAL9
B	18	HIS	-	expression tag	UNP Q6VAL9
B	19	HIS	-	expression tag	UNP Q6VAL9
B	20	ASP	-	expression tag	UNP Q6VAL9
B	21	TYR	-	expression tag	UNP Q6VAL9

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	22	ASP	-	expression tag	UNP Q6VAL9
B	23	ILE	-	expression tag	UNP Q6VAL9
B	24	PRO	-	expression tag	UNP Q6VAL9
B	25	THR	-	expression tag	UNP Q6VAL9
B	26	THR	-	expression tag	UNP Q6VAL9
B	27	GLU	-	expression tag	UNP Q6VAL9
B	28	ASN	-	expression tag	UNP Q6VAL9
B	29	LEU	-	expression tag	UNP Q6VAL9
B	30	TYR	-	expression tag	UNP Q6VAL9
B	31	PHE	-	expression tag	UNP Q6VAL9
B	32	GLN	-	expression tag	UNP Q6VAL9
B	33	GLY	-	expression tag	UNP Q6VAL9
B	34	ALA	-	expression tag	UNP Q6VAL9
B	35	MET	-	expression tag	UNP Q6VAL9
B	36	GLY	-	expression tag	UNP Q6VAL9
B	37	SER	-	expression tag	UNP Q6VAL9
B	38	GLY	-	expression tag	UNP Q6VAL9
B	39	ILE	-	expression tag	UNP Q6VAL9
B	40	GLN	-	expression tag	UNP Q6VAL9
D	14	HIS	-	expression tag	UNP Q6VAL9
D	15	HIS	-	expression tag	UNP Q6VAL9
D	16	HIS	-	expression tag	UNP Q6VAL9
D	17	HIS	-	expression tag	UNP Q6VAL9
D	18	HIS	-	expression tag	UNP Q6VAL9
D	19	HIS	-	expression tag	UNP Q6VAL9
D	20	ASP	-	expression tag	UNP Q6VAL9
D	21	TYR	-	expression tag	UNP Q6VAL9
D	22	ASP	-	expression tag	UNP Q6VAL9
D	23	ILE	-	expression tag	UNP Q6VAL9
D	24	PRO	-	expression tag	UNP Q6VAL9
D	25	THR	-	expression tag	UNP Q6VAL9
D	26	THR	-	expression tag	UNP Q6VAL9
D	27	GLU	-	expression tag	UNP Q6VAL9
D	28	ASN	-	expression tag	UNP Q6VAL9
D	29	LEU	-	expression tag	UNP Q6VAL9
D	30	TYR	-	expression tag	UNP Q6VAL9
D	31	PHE	-	expression tag	UNP Q6VAL9
D	32	GLN	-	expression tag	UNP Q6VAL9
D	33	GLY	-	expression tag	UNP Q6VAL9
D	34	ALA	-	expression tag	UNP Q6VAL9
D	35	MET	-	expression tag	UNP Q6VAL9
D	36	GLY	-	expression tag	UNP Q6VAL9

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	37	SER	-	expression tag	UNP Q6VAL9
D	38	GLY	-	expression tag	UNP Q6VAL9
D	39	ILE	-	expression tag	UNP Q6VAL9
D	40	GLN	-	expression tag	UNP Q6VAL9
F	14	HIS	-	expression tag	UNP Q6VAL9
F	15	HIS	-	expression tag	UNP Q6VAL9
F	16	HIS	-	expression tag	UNP Q6VAL9
F	17	HIS	-	expression tag	UNP Q6VAL9
F	18	HIS	-	expression tag	UNP Q6VAL9
F	19	HIS	-	expression tag	UNP Q6VAL9
F	20	ASP	-	expression tag	UNP Q6VAL9
F	21	TYR	-	expression tag	UNP Q6VAL9
F	22	ASP	-	expression tag	UNP Q6VAL9
F	23	ILE	-	expression tag	UNP Q6VAL9
F	24	PRO	-	expression tag	UNP Q6VAL9
F	25	THR	-	expression tag	UNP Q6VAL9
F	26	THR	-	expression tag	UNP Q6VAL9
F	27	GLU	-	expression tag	UNP Q6VAL9
F	28	ASN	-	expression tag	UNP Q6VAL9
F	29	LEU	-	expression tag	UNP Q6VAL9
F	30	TYR	-	expression tag	UNP Q6VAL9
F	31	PHE	-	expression tag	UNP Q6VAL9
F	32	GLN	-	expression tag	UNP Q6VAL9
F	33	GLY	-	expression tag	UNP Q6VAL9
F	34	ALA	-	expression tag	UNP Q6VAL9
F	35	MET	-	expression tag	UNP Q6VAL9
F	36	GLY	-	expression tag	UNP Q6VAL9
F	37	SER	-	expression tag	UNP Q6VAL9
F	38	GLY	-	expression tag	UNP Q6VAL9
F	39	ILE	-	expression tag	UNP Q6VAL9
F	40	GLN	-	expression tag	UNP Q6VAL9
H	14	HIS	-	expression tag	UNP Q6VAL9
H	15	HIS	-	expression tag	UNP Q6VAL9
H	16	HIS	-	expression tag	UNP Q6VAL9
H	17	HIS	-	expression tag	UNP Q6VAL9
H	18	HIS	-	expression tag	UNP Q6VAL9
H	19	HIS	-	expression tag	UNP Q6VAL9
H	20	ASP	-	expression tag	UNP Q6VAL9
H	21	TYR	-	expression tag	UNP Q6VAL9
H	22	ASP	-	expression tag	UNP Q6VAL9
H	23	ILE	-	expression tag	UNP Q6VAL9
H	24	PRO	-	expression tag	UNP Q6VAL9

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
H	25	THR	-	expression tag	UNP Q6VAL9
H	26	THR	-	expression tag	UNP Q6VAL9
H	27	GLU	-	expression tag	UNP Q6VAL9
H	28	ASN	-	expression tag	UNP Q6VAL9
H	29	LEU	-	expression tag	UNP Q6VAL9
H	30	TYR	-	expression tag	UNP Q6VAL9
H	31	PHE	-	expression tag	UNP Q6VAL9
H	32	GLN	-	expression tag	UNP Q6VAL9
H	33	GLY	-	expression tag	UNP Q6VAL9
H	34	ALA	-	expression tag	UNP Q6VAL9
H	35	MET	-	expression tag	UNP Q6VAL9
H	36	GLY	-	expression tag	UNP Q6VAL9
H	37	SER	-	expression tag	UNP Q6VAL9
H	38	GLY	-	expression tag	UNP Q6VAL9
H	39	ILE	-	expression tag	UNP Q6VAL9
H	40	GLN	-	expression tag	UNP Q6VAL9

- Molecule 3 is a protein called C-type cytochrome c552.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	I	105	Total	C	N	O	S	0	0	0
			788	508	132	145	3			
3	J	102	Total	C	N	O	S	0	0	0
			769	498	129	139	3			
3	K	103	Total	C	N	O	S	0	0	0
			775	501	130	141	3			
3	L	102	Total	C	N	O	S	0	1	0
			775	502	129	141	3			

There are 40 discrepancies between the modelled and reference sequences:

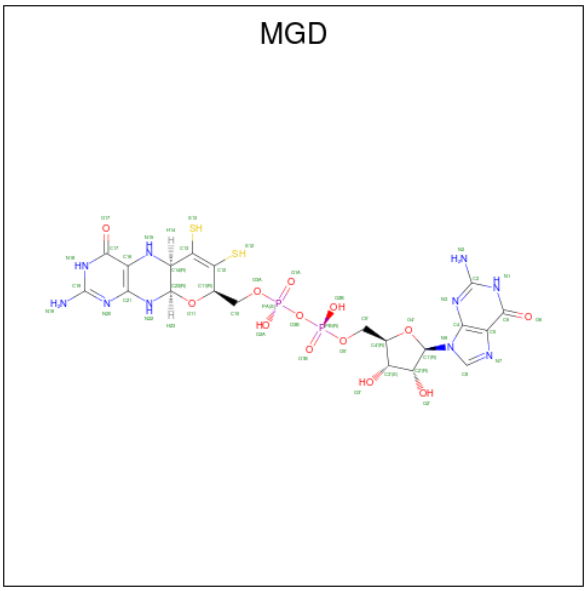
Chain	Residue	Modelled	Actual	Comment	Reference
I	19	MET	-	initiating methionine	UNP Q2TV05
I	20	ASP	-	expression tag	UNP Q2TV05
I	128	LEU	-	expression tag	UNP Q2TV05
I	129	GLU	-	expression tag	UNP Q2TV05
I	130	HIS	-	expression tag	UNP Q2TV05
I	131	HIS	-	expression tag	UNP Q2TV05
I	132	HIS	-	expression tag	UNP Q2TV05
I	133	HIS	-	expression tag	UNP Q2TV05
I	134	HIS	-	expression tag	UNP Q2TV05
I	135	HIS	-	expression tag	UNP Q2TV05

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
J	19	MET	-	initiating methionine	UNP Q2TV05
J	20	ASP	-	expression tag	UNP Q2TV05
J	128	LEU	-	expression tag	UNP Q2TV05
J	129	GLU	-	expression tag	UNP Q2TV05
J	130	HIS	-	expression tag	UNP Q2TV05
J	131	HIS	-	expression tag	UNP Q2TV05
J	132	HIS	-	expression tag	UNP Q2TV05
J	133	HIS	-	expression tag	UNP Q2TV05
J	134	HIS	-	expression tag	UNP Q2TV05
J	135	HIS	-	expression tag	UNP Q2TV05
K	19	MET	-	initiating methionine	UNP Q2TV05
K	20	ASP	-	expression tag	UNP Q2TV05
K	128	LEU	-	expression tag	UNP Q2TV05
K	129	GLU	-	expression tag	UNP Q2TV05
K	130	HIS	-	expression tag	UNP Q2TV05
K	131	HIS	-	expression tag	UNP Q2TV05
K	132	HIS	-	expression tag	UNP Q2TV05
K	133	HIS	-	expression tag	UNP Q2TV05
K	134	HIS	-	expression tag	UNP Q2TV05
K	135	HIS	-	expression tag	UNP Q2TV05
L	19	MET	-	initiating methionine	UNP Q2TV05
L	20	ASP	-	expression tag	UNP Q2TV05
L	128	LEU	-	expression tag	UNP Q2TV05
L	129	GLU	-	expression tag	UNP Q2TV05
L	130	HIS	-	expression tag	UNP Q2TV05
L	131	HIS	-	expression tag	UNP Q2TV05
L	132	HIS	-	expression tag	UNP Q2TV05
L	133	HIS	-	expression tag	UNP Q2TV05
L	134	HIS	-	expression tag	UNP Q2TV05
L	135	HIS	-	expression tag	UNP Q2TV05

- Molecule 4 is 2-AMINO-5,6-DIMERCAPTO-7-METHYL-3,7,8A,9-TETRAHYDRO-8-OXA-1,3,9,10-TETRAAZA-ANTHRACEN-4-ONE GUANOSINE DINUCLEOTIDE (three-letter code: MGD) (formula: $C_{20}H_{26}N_{10}O_{13}P_2S_2$).

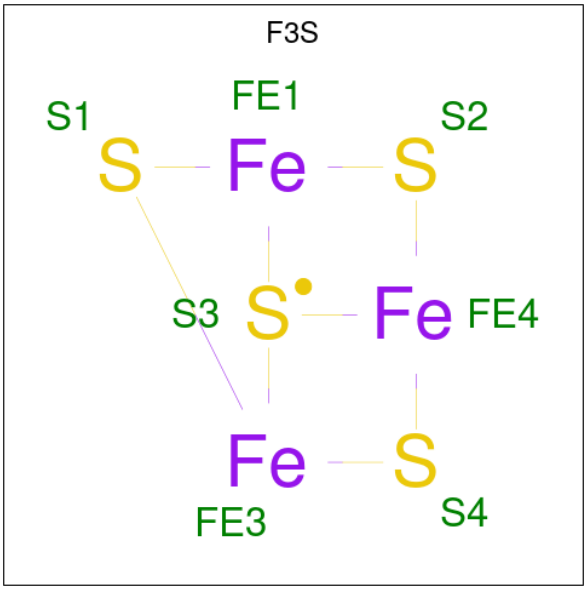


Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
4	A	1	Total	C	N	O	P	S	0	0
			47	20	10	13	2	2		
4	A	1	Total	C	N	O	P	S	0	0
			47	20	10	13	2	2		
4	C	1	Total	C	N	O	P	S	0	0
			47	20	10	13	2	2		
4	C	1	Total	C	N	O	P	S	0	0
			47	20	10	13	2	2		
4	E	1	Total	C	N	O	P	S	0	0
			47	20	10	13	2	2		
4	E	1	Total	C	N	O	P	S	0	0
			47	20	10	13	2	2		
4	G	1	Total	C	N	O	P	S	0	0
			47	20	10	13	2	2		
4	G	1	Total	C	N	O	P	S	0	0
			47	20	10	13	2	2		

- Molecule 5 is MOLYBDENUM(IV) ION (three-letter code: 4MO) (formula: Mo).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	1	Total	Mo	0	0
			1	1		
5	C	1	Total	Mo	0	0
			1	1		
5	E	1	Total	Mo	0	0
			1	1		
5	G	1	Total	Mo	0	0
			1	1		

- Molecule 6 is FE3-S4 CLUSTER (three-letter code: F3S) (formula: Fe_3S_4).

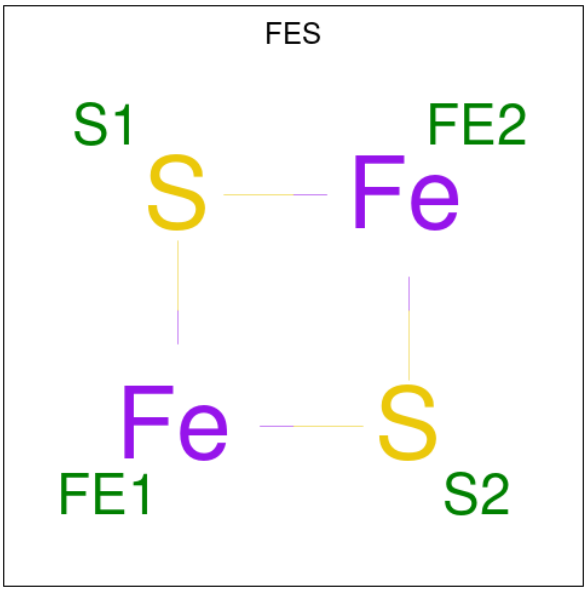


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total Fe S 7 3 4	0	0
6	C	1	Total Fe S 7 3 4	0	0
6	E	1	Total Fe S 7 3 4	0	0
6	G	1	Total Fe S 7 3 4	0	0

- Molecule 7 is OXYGEN ATOM (three-letter code: O) (formula: O).

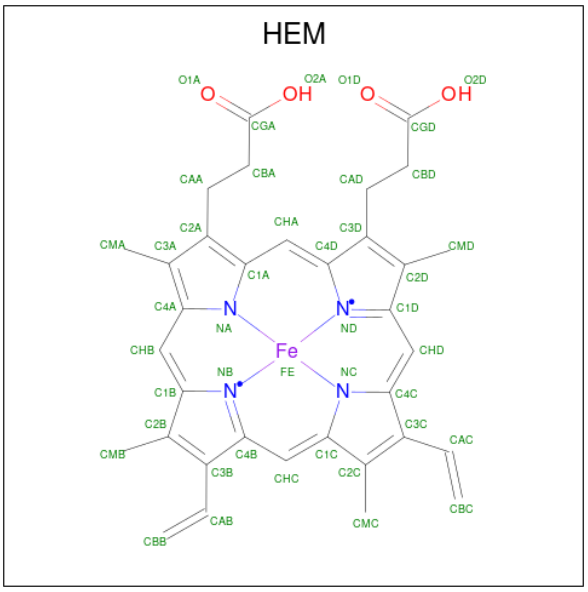
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	1	Total O 1 1	0	0
7	C	1	Total O 1 1	0	0
7	E	1	Total O 1 1	0	0
7	G	1	Total O 1 1	0	0

- Molecule 8 is FE2/S2 (INORGANIC) CLUSTER (three-letter code: FES) (formula: Fe_2S_2).



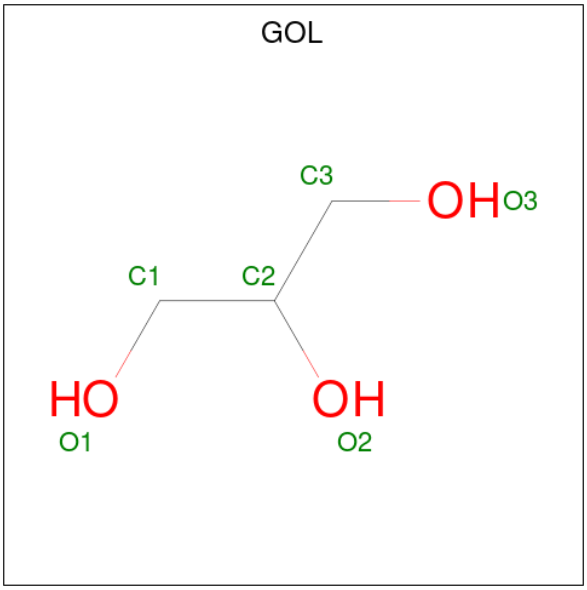
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	B	1	Total Fe S 4 2 2	0	0
8	D	1	Total Fe S 4 2 2	0	0
8	F	1	Total Fe S 4 2 2	0	0
8	H	1	Total Fe S 4 2 2	0	0

- Molecule 9 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (formula: $C_{34}H_{32}FeN_4O_4$).



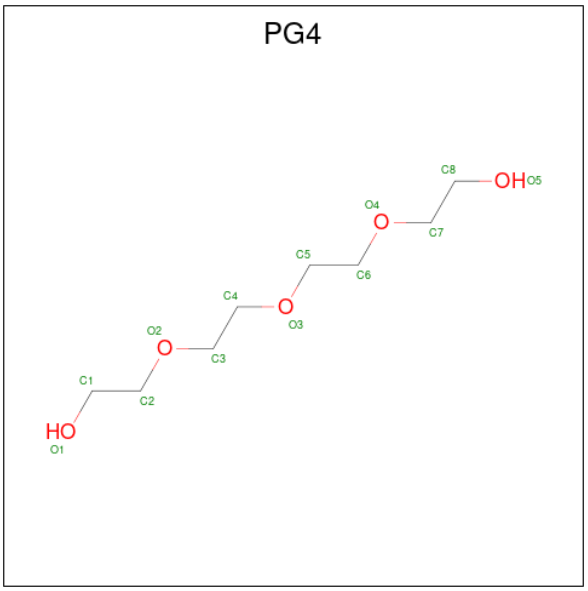
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
9	I	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
9	J	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
9	L	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		

- Molecule 10 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



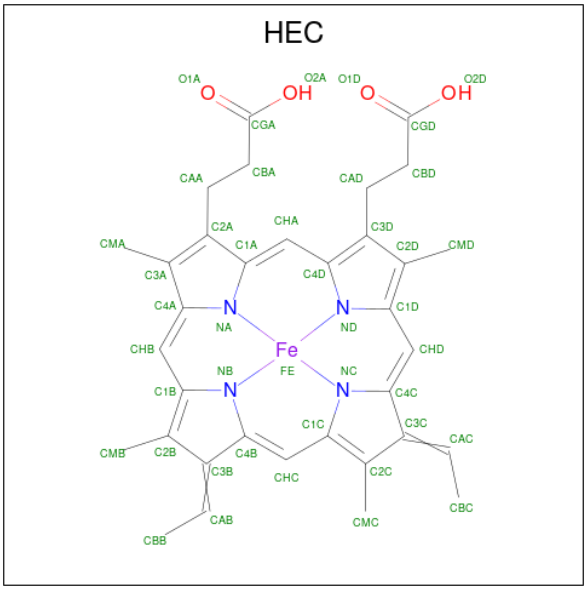
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
10	E	1	Total	C	O	0	0
			6	3	3		

- Molecule 11 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula: $C_8H_{18}O_5$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
11	E	1	Total C O 13 8 5	0	0
11	G	1	Total C O 13 8 5	0	0

- Molecule 12 is HEME C (three-letter code: HEC) (formula: $C_{34}H_{34}FeN_4O_4$).



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

- Molecule 13 is water.

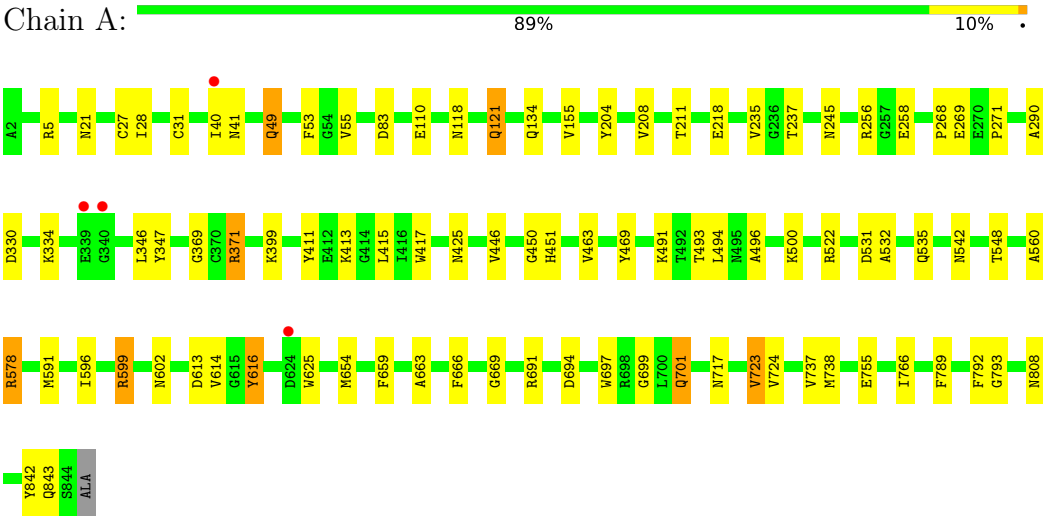
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
13	A	341	Total O 341 341	0	0
13	B	21	Total O 21 21	0	0
13	I	8	Total O 8 8	0	0
13	C	281	Total O 281 281	0	0
13	D	24	Total O 24 24	0	0
13	J	13	Total O 13 13	0	0
13	E	481	Total O 481 481	0	0
13	F	73	Total O 73 73	0	0
13	K	19	Total O 19 19	0	0
13	G	420	Total O 420 420	0	0
13	H	54	Total O 54 54	0	0
13	L	14	Total O 14 14	0	0

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 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.

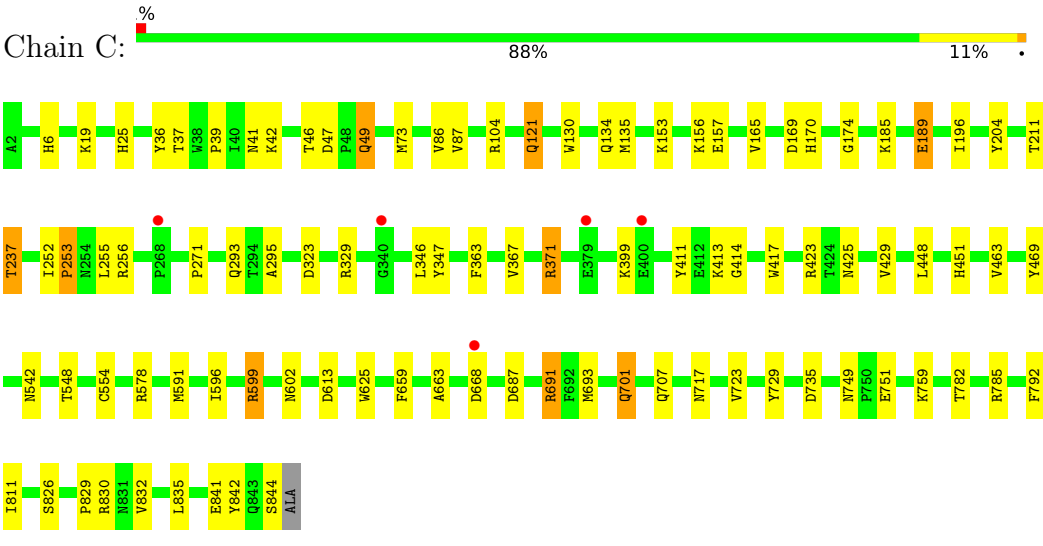
• Molecule 1: AroA

Chain A:



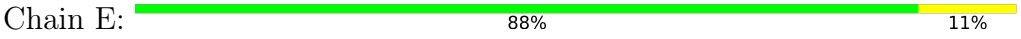
• Molecule 1: AroA

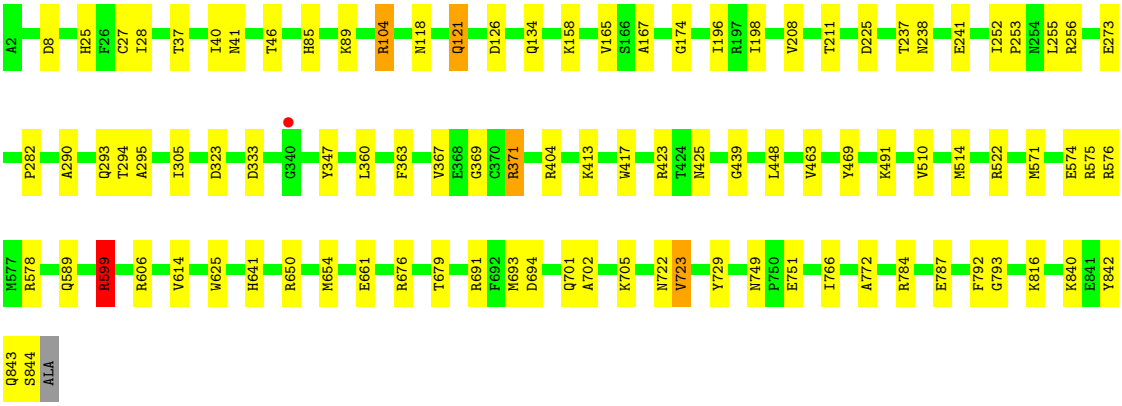
Chain C:



• Molecule 1: AroA

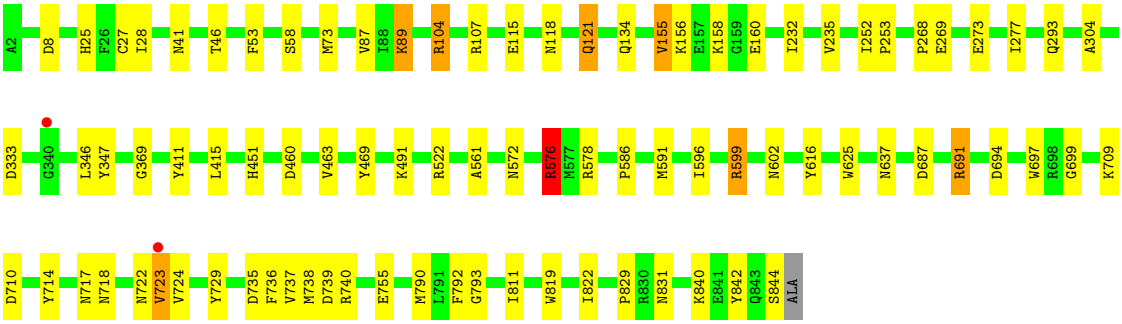
Chain E:





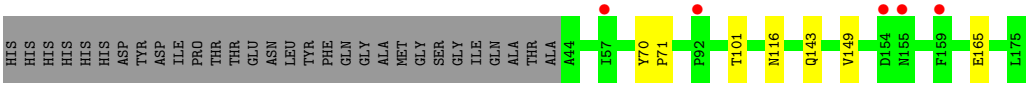
• Molecule 1: AroA

Chain G:



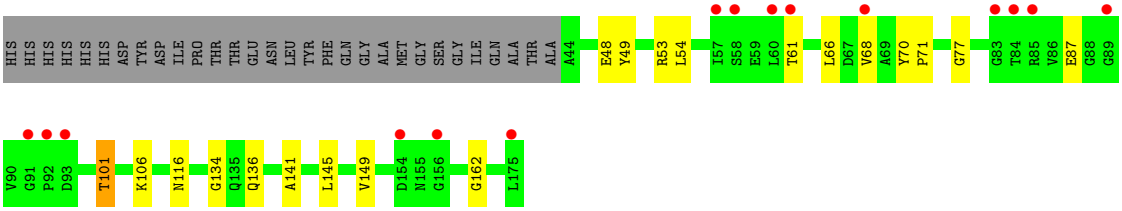
• Molecule 2: AroB

Chain B:



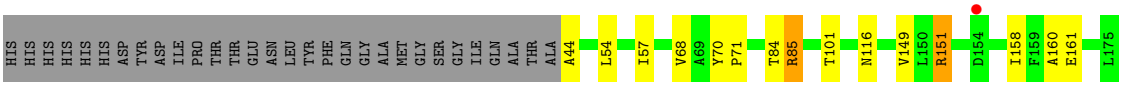
• Molecule 2: AroB

Chain D:



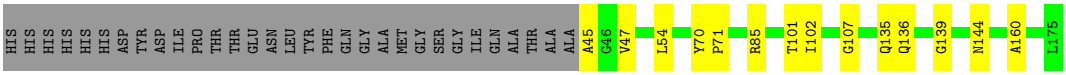
• Molecule 2: AroB

Chain F:



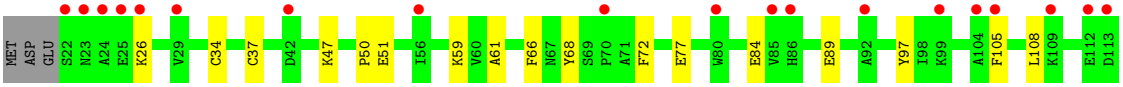
• Molecule 2: AroB

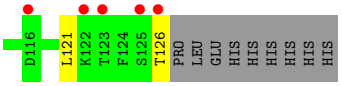
Chain H:



• Molecule 3: C-type cytochrome c552

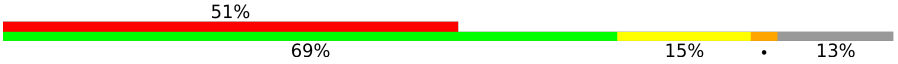
Chain I:

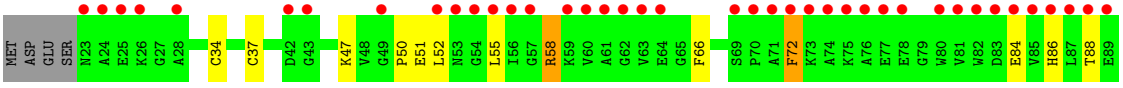


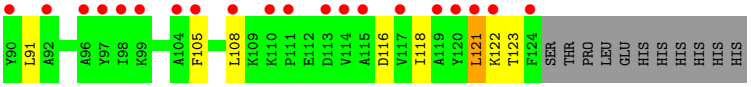


• Molecule 3: C-type cytochrome c552

Chain J:

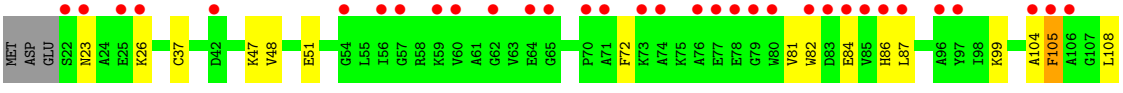


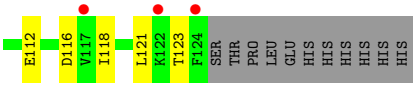




• Molecule 3: C-type cytochrome c552

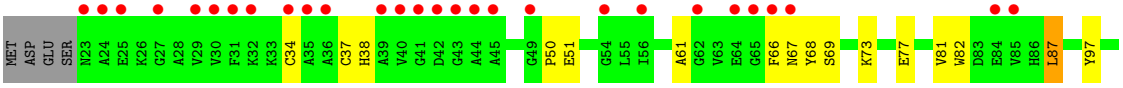
Chain K:

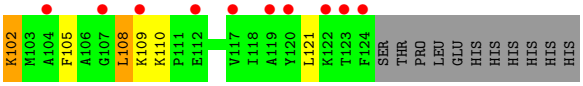




• Molecule 3: C-type cytochrome c552

Chain L:




K102
M103
A104
F105
A106
G107
L108
K109
K110
P111
E112
V117
I118
A119
Y120
L121
K122
T123
F124
SER
THR
PRO
LEU
GLU
HIS
HIS
HIS
HIS
HIS
HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	129.40Å 126.56Å 148.02Å 90.00° 107.81° 90.00°	Depositor
Resolution (Å)	49.23 – 2.25 49.19 – 2.25	Depositor EDS
% Data completeness (in resolution range)	92.2 (49.23-2.25) 92.2 (49.19-2.25)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	4.55 (at 2.24Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.181 , 0.230 0.189 , 0.234	Depositor DCC
R_{free} test set	10062 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å^2)	3.0	Xtriage
Anisotropy	0.533	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 31.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.35$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	35834	wwPDB-VP
Average B, all atoms (Å^2)	12.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 33.14 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 8.3893e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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, PG4, 4MO, F3S, FES, HEM, HEC, O, MGD

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	0.69	0/6738	0.81	2/9128 (0.0%)
1	C	0.68	1/6733 (0.0%)	0.80	0/9122
1	E	0.67	0/6806	0.83	4/9217 (0.0%)
1	G	0.68	1/6769 (0.0%)	0.84	7/9169 (0.1%)
2	B	0.64	0/1018	0.82	0/1387
2	D	0.69	0/1018	0.84	0/1387
2	F	0.66	0/1018	0.84	0/1387
2	H	0.65	0/1013	0.84	0/1380
3	I	0.71	0/806	0.78	0/1088
3	J	0.72	0/787	0.80	0/1062
3	K	0.70	0/793	0.79	0/1070
3	L	0.75	0/796	0.81	0/1074
All	All	0.68	2/34295 (0.0%)	0.82	13/46471 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	G	115	GLU	CD-OE1	6.99	1.33	1.25
1	C	189	GLU	CD-OE2	5.15	1.31	1.25

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	E	104	ARG	NE-CZ-NH2	-8.47	116.06	120.30
1	G	576	ARG	NE-CZ-NH2	-7.88	116.36	120.30
1	G	576	ARG	NE-CZ-NH1	7.49	124.05	120.30
1	G	104	ARG	NE-CZ-NH2	-7.22	116.69	120.30
1	G	599	ARG	NE-CZ-NH2	-6.92	116.84	120.30

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	6577	0	6339	62	0
1	C	6573	0	6332	63	0
1	E	6624	0	6397	68	0
1	G	6597	0	6372	80	0
2	B	995	0	947	6	0
2	D	995	0	947	12	0
2	F	995	0	947	10	0
2	H	990	0	942	19	0
3	I	788	0	787	14	0
3	J	769	0	770	20	0
3	K	775	0	774	13	0
3	L	775	0	776	19	0
4	A	94	0	44	2	0
4	C	94	0	44	3	0
4	E	94	0	44	2	0
4	G	94	0	44	2	0
5	A	1	0	0	0	0
5	C	1	0	0	0	0
5	E	1	0	0	0	0
5	G	1	0	0	0	0
6	A	7	0	0	0	0
6	C	7	0	0	1	0
6	E	7	0	0	0	0
6	G	7	0	0	0	0
7	A	1	0	0	1	0
7	C	1	0	0	1	0
7	E	1	0	0	1	0
7	G	1	0	0	0	0
8	B	4	0	0	0	0
8	D	4	0	0	1	0
8	F	4	0	0	0	0
8	H	4	0	0	0	0
9	I	43	0	30	9	0
9	J	43	0	30	11	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
9	L	43	0	30	12	0
10	E	6	0	8	0	0
11	E	13	0	17	0	0
11	G	13	0	17	40	0
12	K	43	0	32	7	0
13	A	341	0	0	9	0
13	B	21	0	0	1	0
13	C	281	0	0	7	0
13	D	24	0	0	3	0
13	E	481	0	0	23	0
13	F	73	0	0	5	0
13	G	420	0	0	15	0
13	H	54	0	0	4	0
13	I	8	0	0	0	0
13	J	13	0	0	4	0
13	K	19	0	0	2	0
13	L	14	0	0	0	0
All	All	35834	0	32670	387	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 6.

The worst 5 of 387 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:G:723[B]:VAL:HG12	11:G:2006:PG4:C2	1.55	1.33
3:I:34:CYS:SG	9:I:200:HEM:CAB	2.25	1.25
1:G:723[B]:VAL:HG13	11:G:2006:PG4:C1	1.67	1.22
3:K:37:CYS:SG	12:K:200:HEC:CAC	2.28	1.22
1:G:723[B]:VAL:CG1	11:G:2006:PG4:C2	2.23	1.15

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	846/844 (100%)	793 (94%)	51 (6%)	2 (0%)	47 55
1	C	846/844 (100%)	803 (95%)	41 (5%)	2 (0%)	47 55
1	E	853/844 (101%)	813 (95%)	40 (5%)	0	100 100
1	G	850/844 (101%)	811 (95%)	37 (4%)	2 (0%)	47 55
2	B	130/162 (80%)	120 (92%)	10 (8%)	0	100 100
2	D	130/162 (80%)	120 (92%)	10 (8%)	0	100 100
2	F	130/162 (80%)	117 (90%)	13 (10%)	0	100 100
2	H	129/162 (80%)	122 (95%)	7 (5%)	0	100 100
3	I	103/117 (88%)	97 (94%)	5 (5%)	1 (1%)	15 13
3	J	100/117 (86%)	84 (84%)	15 (15%)	1 (1%)	15 13
3	K	101/117 (86%)	91 (90%)	9 (9%)	1 (1%)	15 13
3	L	101/117 (86%)	95 (94%)	6 (6%)	0	100 100
All	All	4319/4492 (96%)	4066 (94%)	244 (6%)	9 (0%)	47 55

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	I	105	PHE
3	J	105	PHE
3	K	105	PHE
1	A	613	ASP
1	C	668	ASP

5.3.2 Protein sidechains (i)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	682/677 (101%)	659 (97%)	23 (3%)	37 45
1	C	682/677 (101%)	661 (97%)	21 (3%)	40 49

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	E	689/677 (102%)	670 (97%)	19 (3%)	43 52
1	G	686/677 (101%)	664 (97%)	22 (3%)	39 47
2	B	105/129 (81%)	104 (99%)	1 (1%)	76 84
2	D	105/129 (81%)	99 (94%)	6 (6%)	20 20
2	F	105/129 (81%)	101 (96%)	4 (4%)	33 39
2	H	105/129 (81%)	104 (99%)	1 (1%)	76 84
3	I	77/89 (86%)	69 (90%)	8 (10%)	7 5
3	J	74/89 (83%)	66 (89%)	8 (11%)	6 4
3	K	75/89 (84%)	66 (88%)	9 (12%)	5 3
3	L	75/89 (84%)	65 (87%)	10 (13%)	4 2
All	All	3460/3580 (97%)	3328 (96%)	132 (4%)	34 39

5 of 132 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	616	TYR
1	G	755	GLU
3	L	110	LYS
1	C	707	GLN
1	C	701	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 50 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	49	GLN
2	F	116	ASN
2	H	143	GLN
1	E	118	ASN
1	E	425	ASN

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)

Of 31 ligands modelled in this entry, 8 are monoatomic - leaving 23 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	$\#\|Z\| > 2$
6	F3S	G	2004	1	0,9,9	-	-	-		
11	PG4	G	2006	-	12,12,12	0.37	0	11,11,11	0.42	0
12	HEC	K	200	3	32,50,50	1.50	4 (12%)	24,82,82	1.70	5 (20%)
8	FES	D	201	2	0,4,4	-	-	-		
4	MGD	G	2002	5	41,52,52	0.86	1 (2%)	40,81,81	1.13	4 (10%)
8	FES	B	201	2	0,4,4	-	-	-		
4	MGD	C	2002	5	41,52,52	0.89	3 (7%)	40,81,81	1.02	2 (5%)
9	HEM	J	200	3	41,50,50	1.34	4 (9%)	45,82,82	1.79	10 (22%)
9	HEM	L	200	3	41,50,50	1.33	5 (12%)	45,82,82	1.75	10 (22%)
6	F3S	A	2004	1	0,9,9	-	-	-		
9	HEM	I	200	3	41,50,50	1.35	5 (12%)	45,82,82	1.85	8 (17%)
4	MGD	G	2001	5	41,52,52	0.84	1 (2%)	40,81,81	1.00	3 (7%)
8	FES	F	201	2	0,4,4	-	-	-		
4	MGD	E	2002	5	41,52,52	0.88	2 (4%)	40,81,81	1.18	2 (5%)
4	MGD	A	2002	5	41,52,52	0.79	3 (7%)	40,81,81	1.13	3 (7%)
4	MGD	E	2001	5	41,52,52	0.91	3 (7%)	40,81,81	1.02	2 (5%)
4	MGD	A	2001	5	41,52,52	0.96	3 (7%)	40,81,81	1.14	3 (7%)
6	F3S	E	2004	1	0,9,9	-	-	-		
6	F3S	C	2004	1	0,9,9	-	-	-		
8	FES	H	201	2	0,4,4	-	-	-		

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
10	GOL	E	2006	-	5,5,5	0.14	0	5,5,5	0.19	0
11	PG4	E	2007	-	12,12,12	0.39	0	11,11,11	0.20	0
4	MGD	C	2001	5	41,52,52	0.82	3 (7%)	40,81,81	1.10	3 (7%)

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
6	F3S	G	2004	1	-	-	0/3/3/3
11	PG4	G	2006	-	-	9/10/10/10	-
12	HEC	K	200	3	-	3/10/54/54	-
8	FES	D	201	2	-	-	0/1/1/1
4	MGD	G	2002	5	-	4/18/66/66	0/6/6/6
8	FES	B	201	2	-	-	0/1/1/1
4	MGD	C	2002	5	-	4/18/66/66	0/6/6/6
9	HEM	J	200	3	-	3/12/54/54	-
9	HEM	L	200	3	-	2/12/54/54	-
6	F3S	A	2004	1	-	-	0/3/3/3
9	HEM	I	200	3	-	0/12/54/54	-
4	MGD	G	2001	5	-	5/18/66/66	0/6/6/6
8	FES	F	201	2	-	-	0/1/1/1
4	MGD	E	2002	5	-	4/18/66/66	0/6/6/6
4	MGD	A	2002	5	-	4/18/66/66	0/6/6/6
4	MGD	E	2001	5	-	4/18/66/66	0/6/6/6
4	MGD	A	2001	5	-	5/18/66/66	0/6/6/6
6	F3S	E	2004	1	-	-	0/3/3/3
6	F3S	C	2004	1	-	-	0/3/3/3
8	FES	H	201	2	-	-	0/1/1/1
10	GOL	E	2006	-	-	4/4/4/4	-
11	PG4	E	2007	-	-	5/10/10/10	-
4	MGD	C	2001	5	-	4/18/66/66	0/6/6/6

The worst 5 of 37 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
12	K	200	HEC	C2B-C3B	-4.34	1.36	1.40
12	K	200	HEC	CBB-CAB	-4.16	1.33	1.49

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
12	K	200	HEC	CBC-CAC	-4.07	1.34	1.49
9	J	200	HEM	C1B-NB	-3.63	1.34	1.40
4	A	2001	MGD	C23-C14	3.31	1.56	1.53

The worst 5 of 55 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
9	L	200	HEM	C1B-NB-C4B	5.02	110.26	105.07
9	J	200	HEM	CHC-C4B-NB	4.93	129.79	124.43
9	I	200	HEM	CHC-C4B-NB	4.89	129.75	124.43
9	L	200	HEM	CHC-C4B-NB	4.75	129.59	124.43
9	I	200	HEM	C1B-NB-C4B	4.62	109.84	105.07

There are no chirality outliers.

5 of 60 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	2001	MGD	C5'-O5'-PB-O3B
4	A	2002	MGD	PA-O3B-PB-O5'
4	A	2002	MGD	C5'-O5'-PB-O1B
4	A	2002	MGD	C5'-O5'-PB-O3B
4	C	2001	MGD	C5'-O5'-PB-O3B

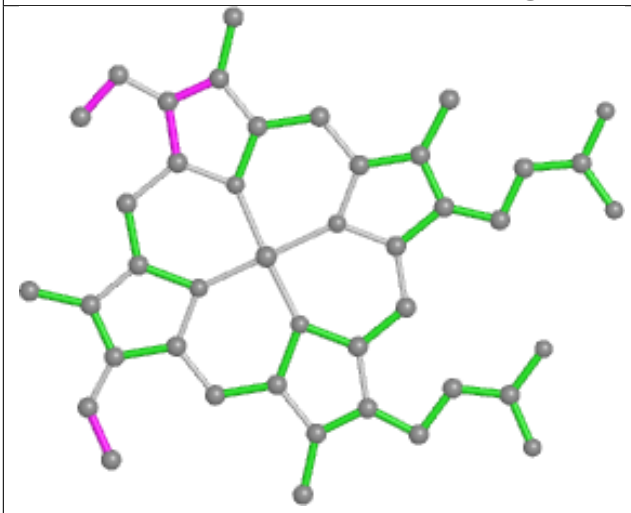
There are no ring outliers.

13 monomers are involved in 90 short contacts:

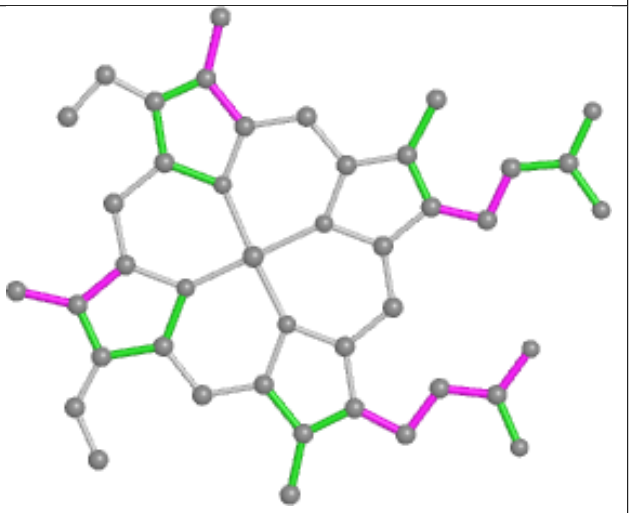
Mol	Chain	Res	Type	Clashes	Symm-Clashes
11	G	2006	PG4	40	0
12	K	200	HEC	7	0
8	D	201	FES	1	0
4	G	2002	MGD	1	0
4	C	2002	MGD	1	0
9	J	200	HEM	11	0
9	L	200	HEM	12	0
9	I	200	HEM	9	0
4	G	2001	MGD	1	0
4	A	2002	MGD	2	0
4	E	2001	MGD	2	0
6	C	2004	F3S	1	0
4	C	2001	MGD	2	0

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.

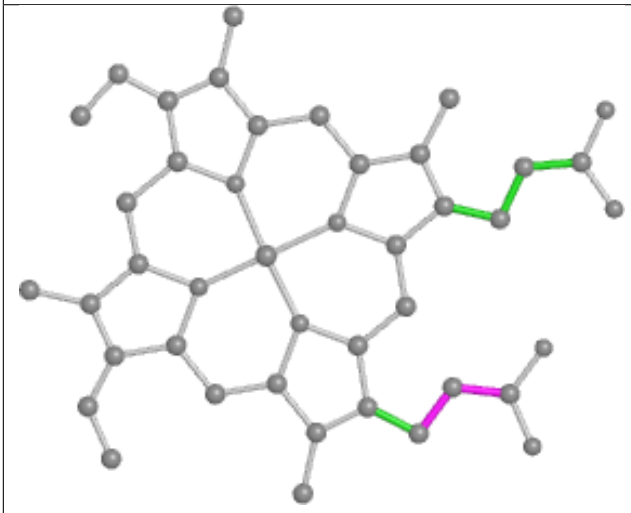
Ligand HEC K 200



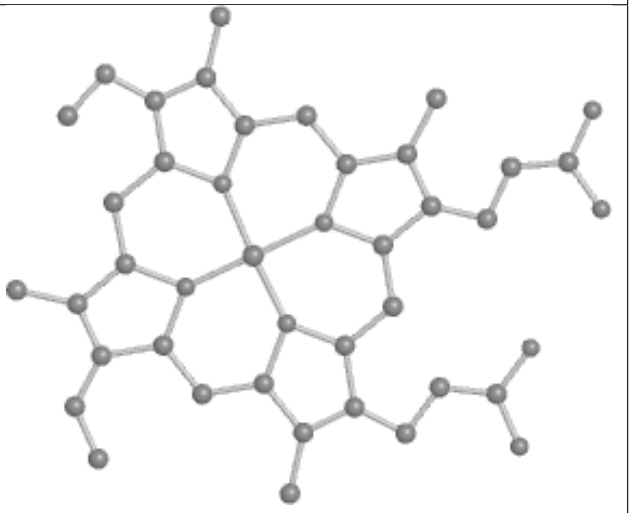
Bond lengths



Bond angles

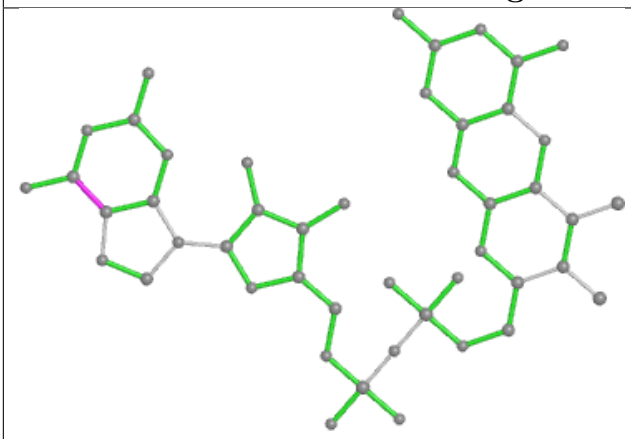


Torsions

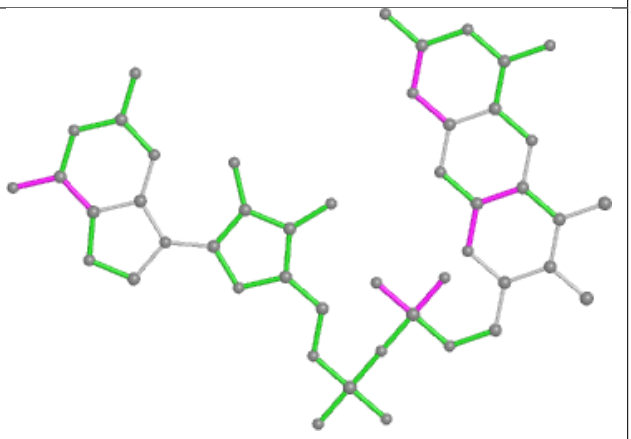


Rings

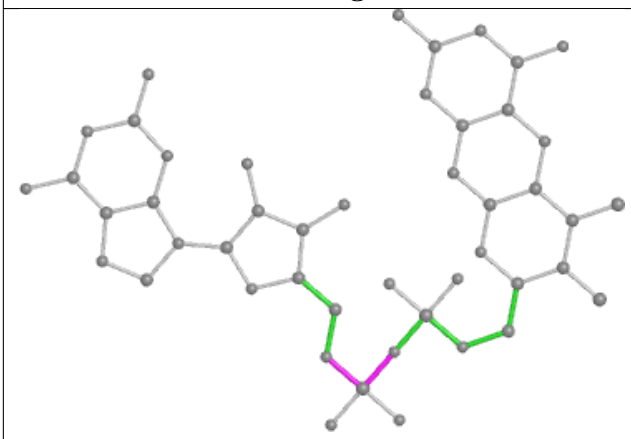
Ligand MGD G 2002



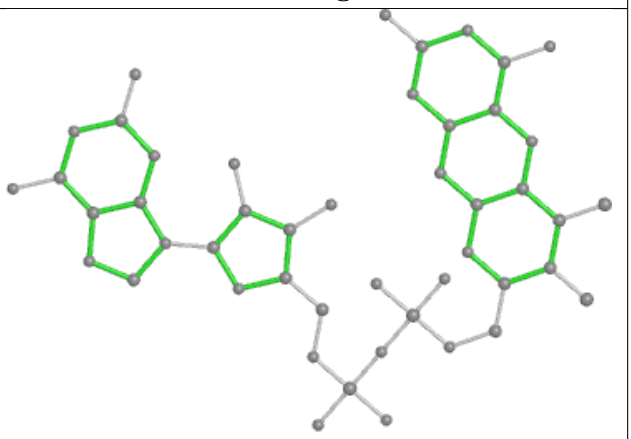
Bond lengths



Bond angles

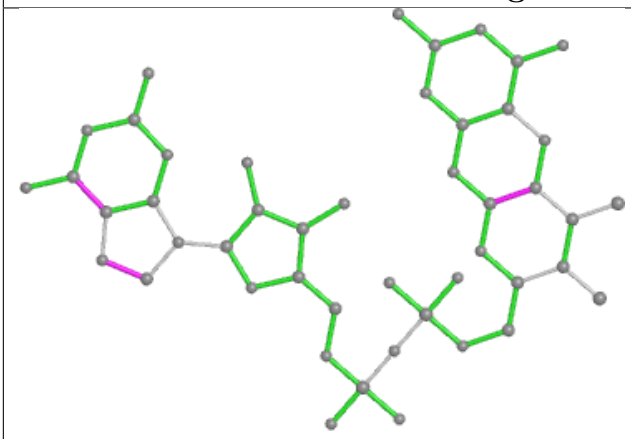


Torsions

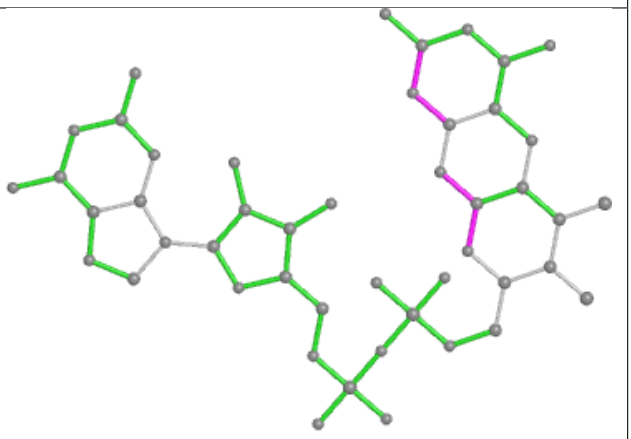


Rings

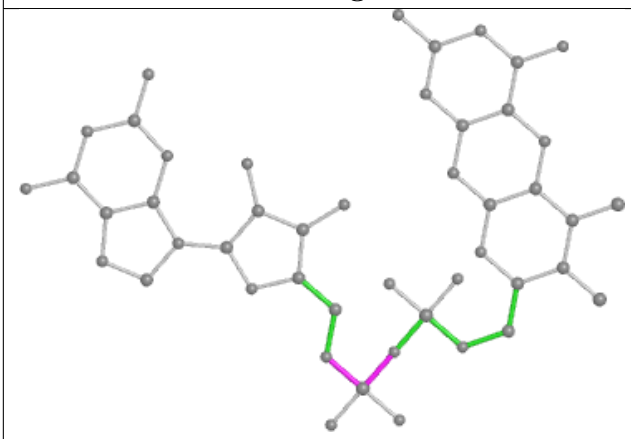
Ligand MGD C 2002



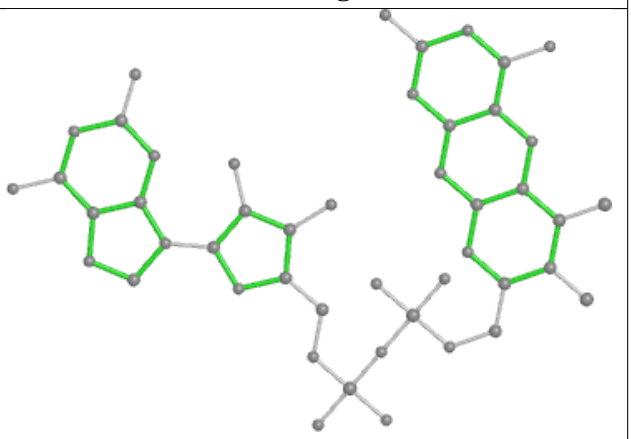
Bond lengths



Bond angles

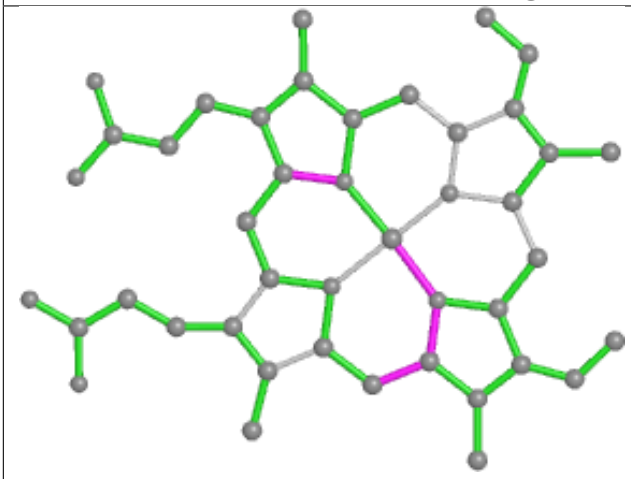


Torsions

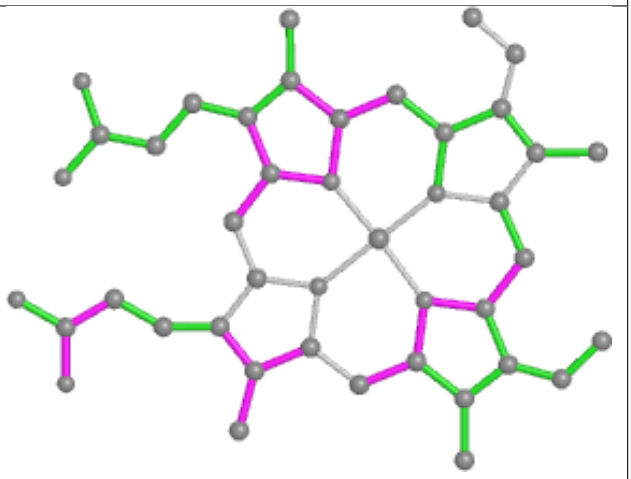


Rings

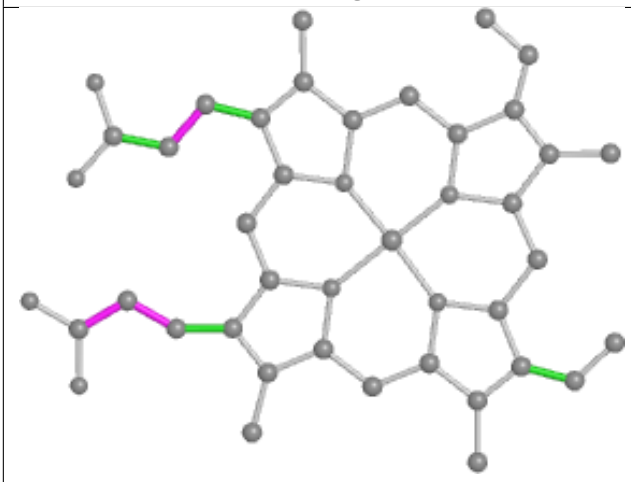
Ligand HEM J 200



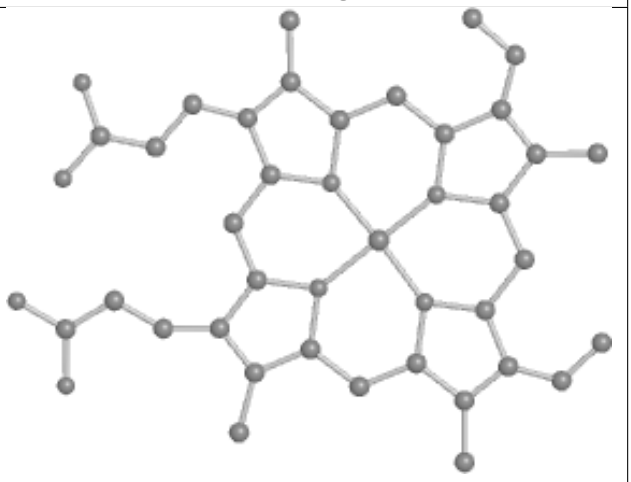
Bond lengths



Bond angles

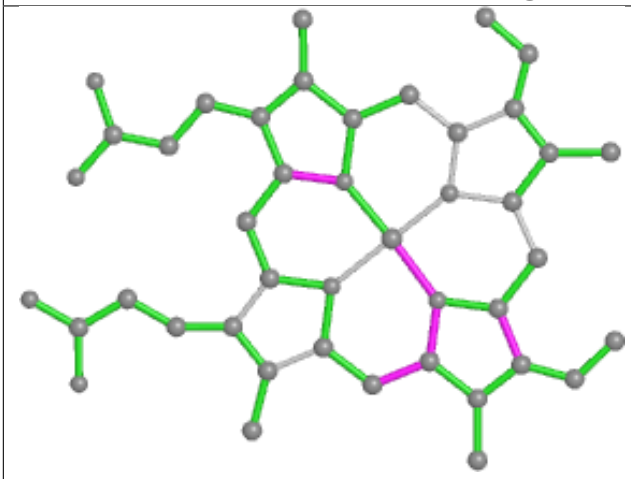


Torsions

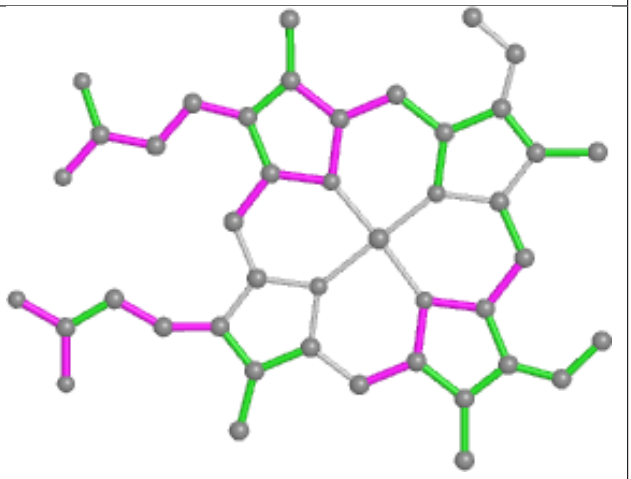


Rings

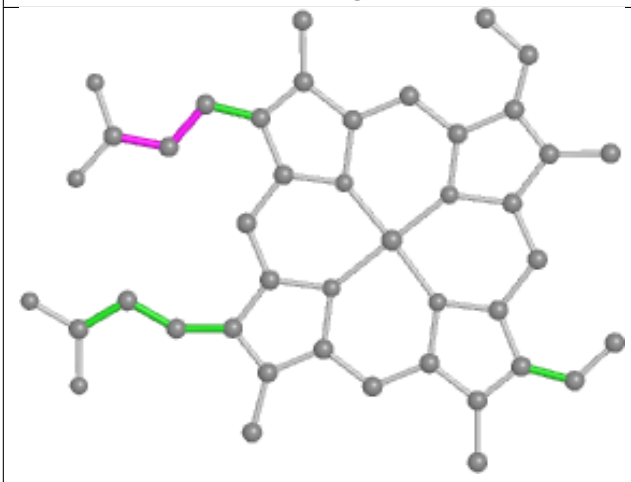
Ligand HEM L 200



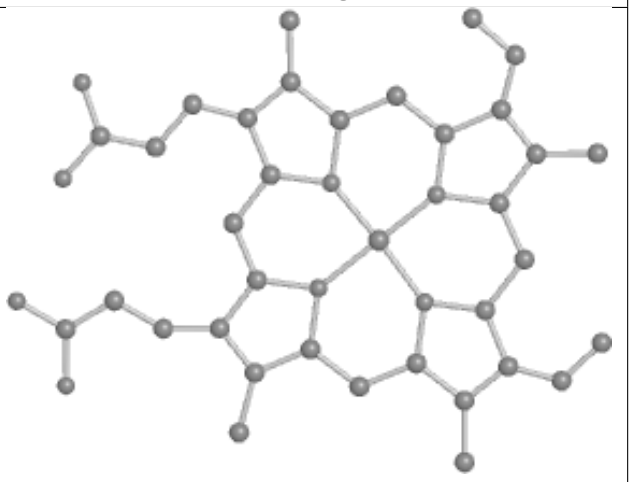
Bond lengths



Bond angles

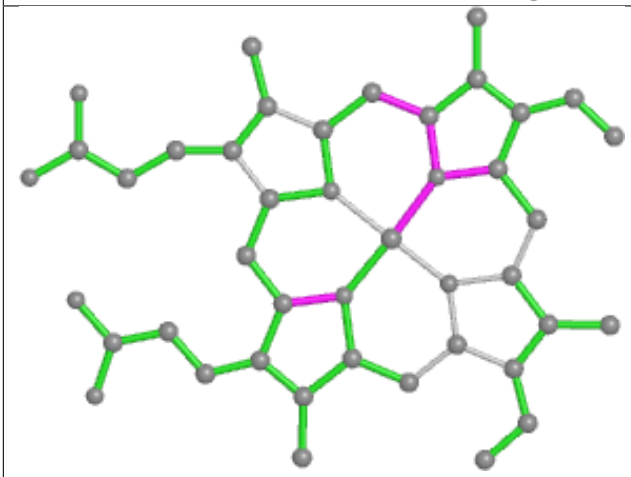


Torsions

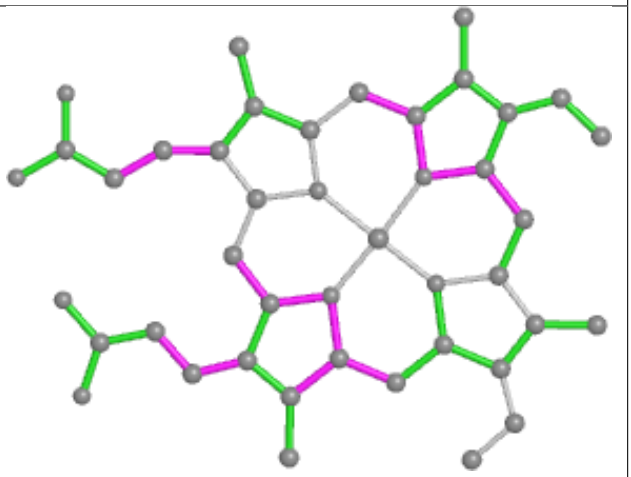


Rings

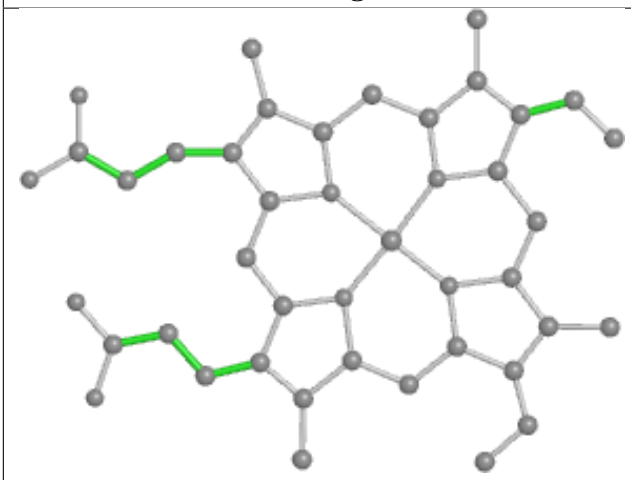
Ligand HEM I 200



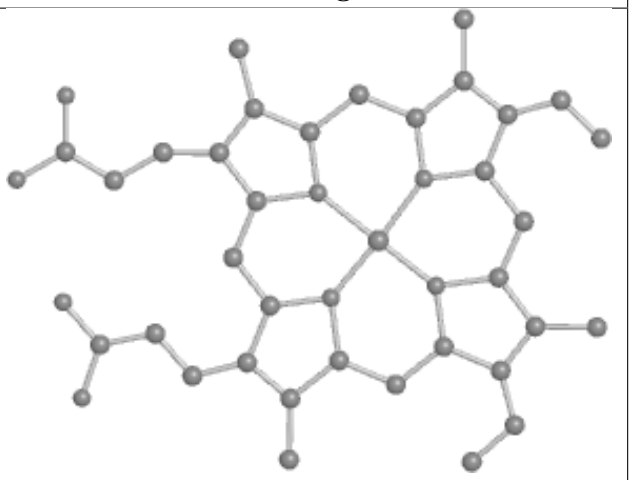
Bond lengths



Bond angles

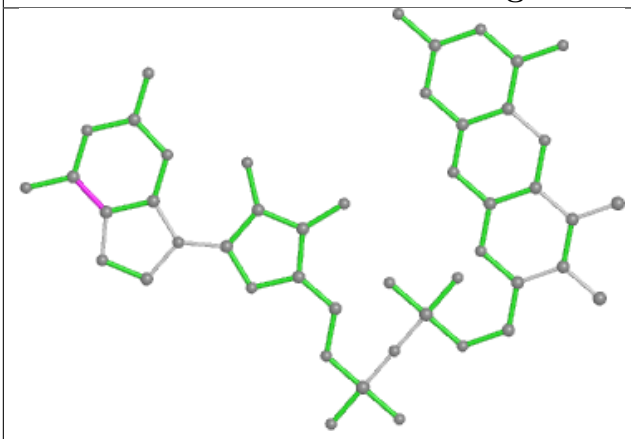


Torsions

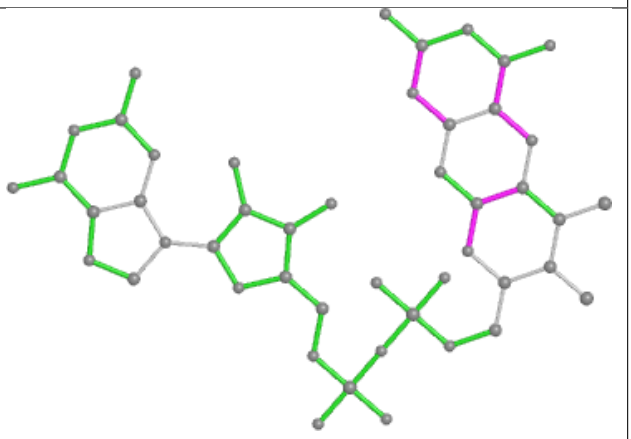


Rings

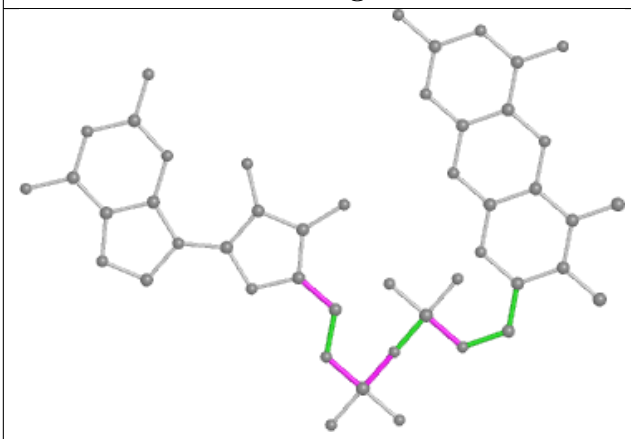
Ligand MGD G 2001



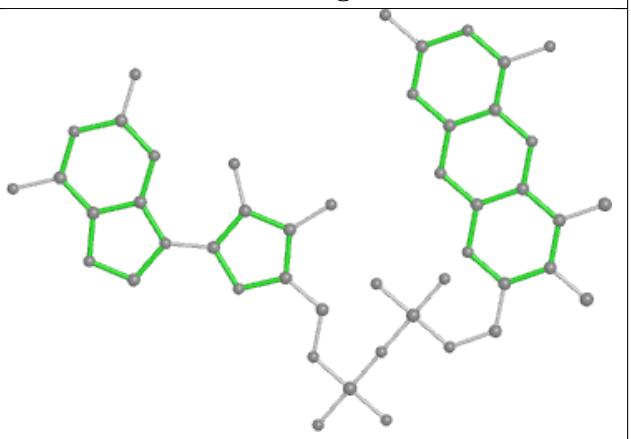
Bond lengths



Bond angles

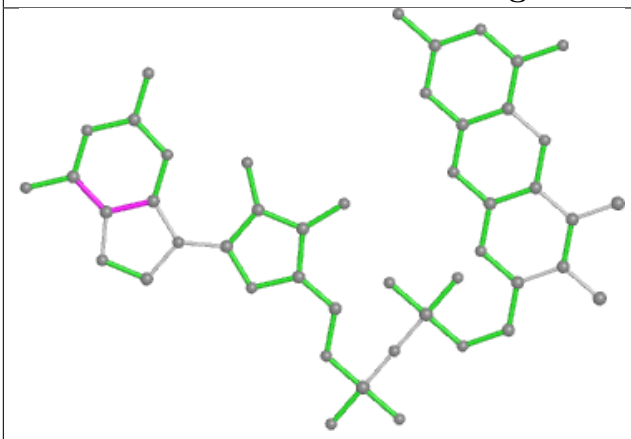


Torsions

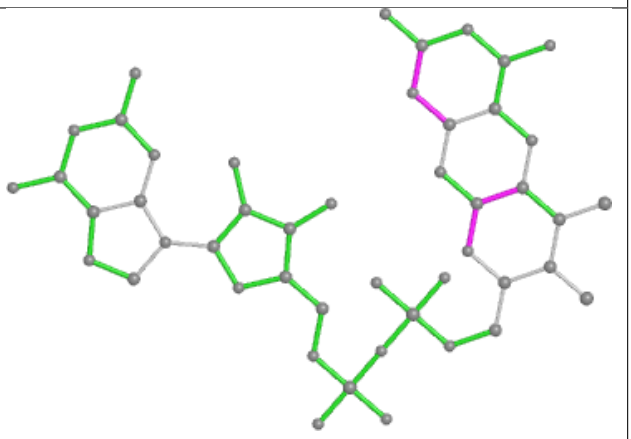


Rings

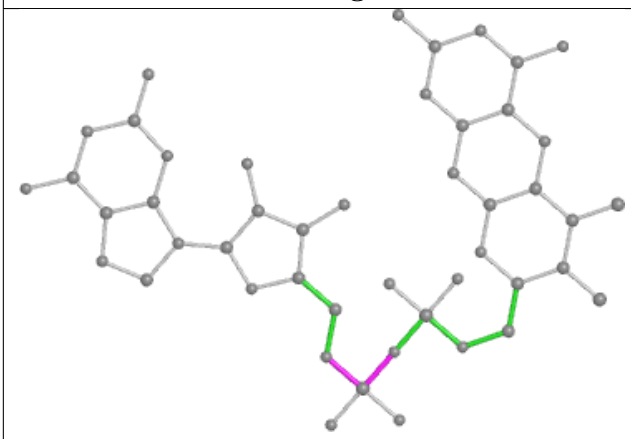
Ligand MGD E 2002



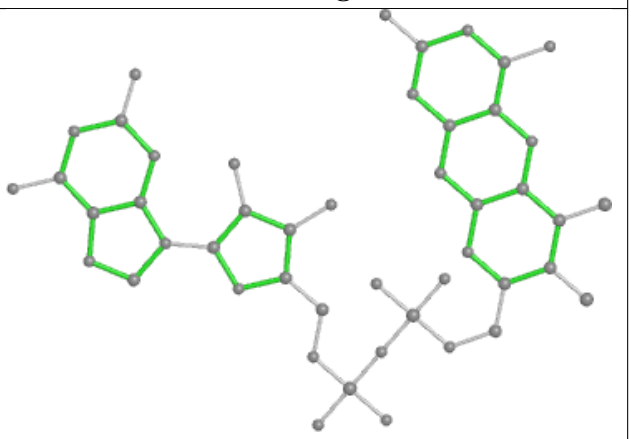
Bond lengths



Bond angles

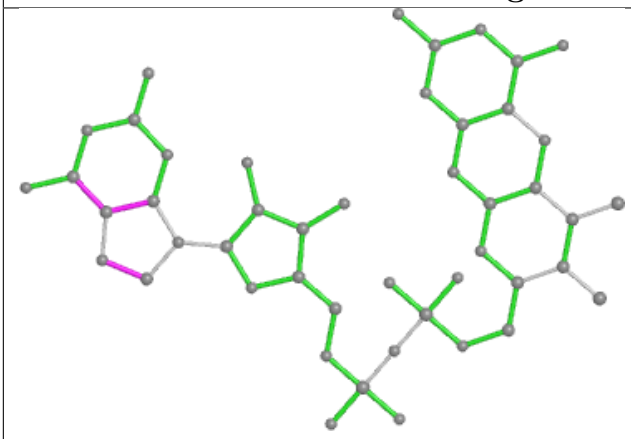


Torsions

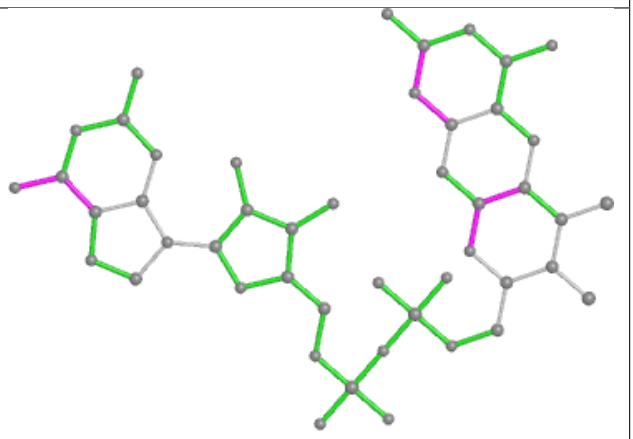


Rings

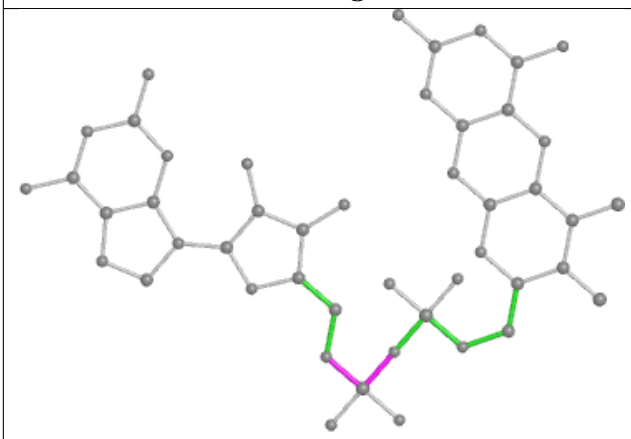
Ligand MGD A 2002



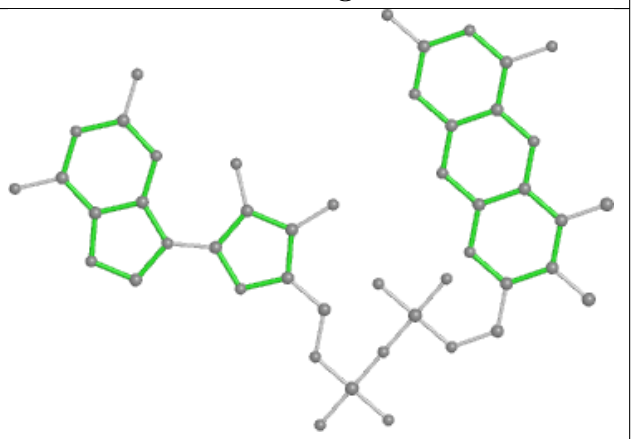
Bond lengths



Bond angles

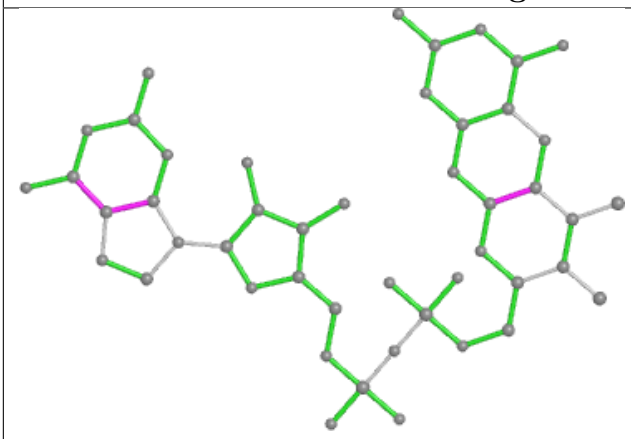


Torsions

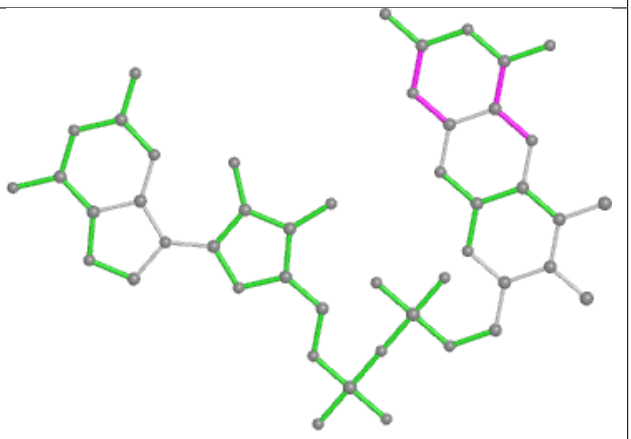


Rings

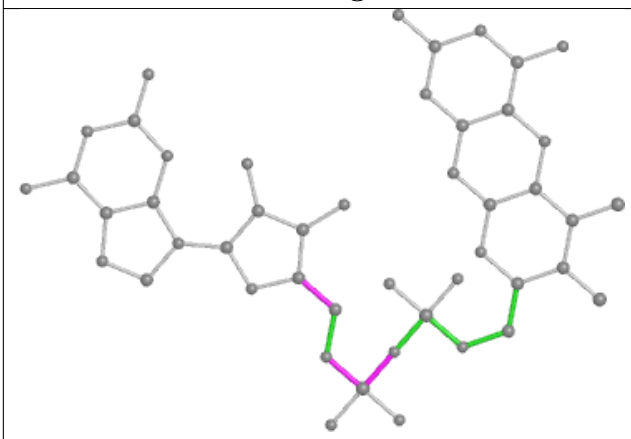
Ligand MGD E 2001



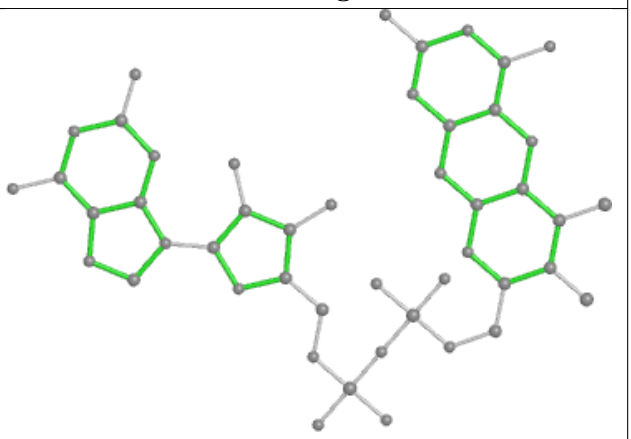
Bond lengths



Bond angles

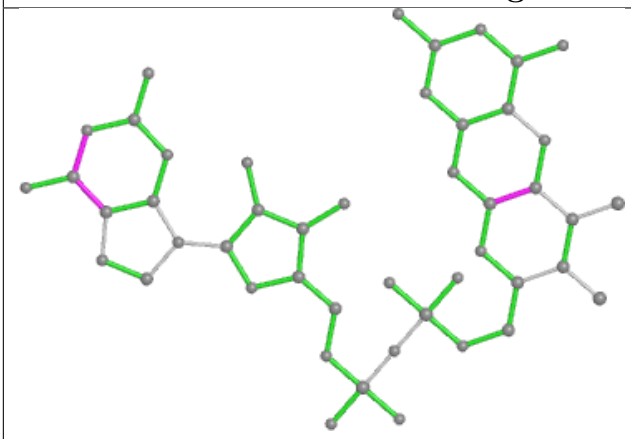


Torsions

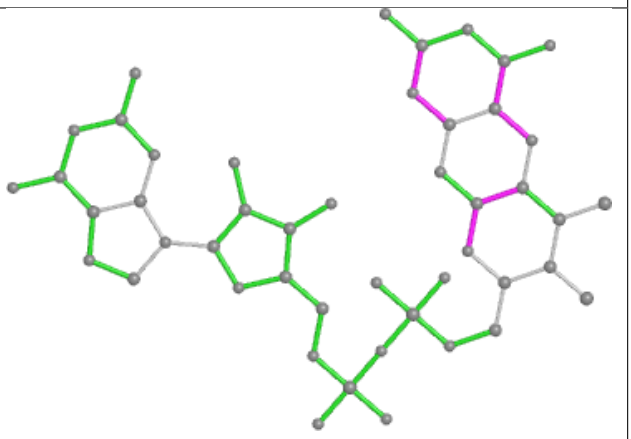


Rings

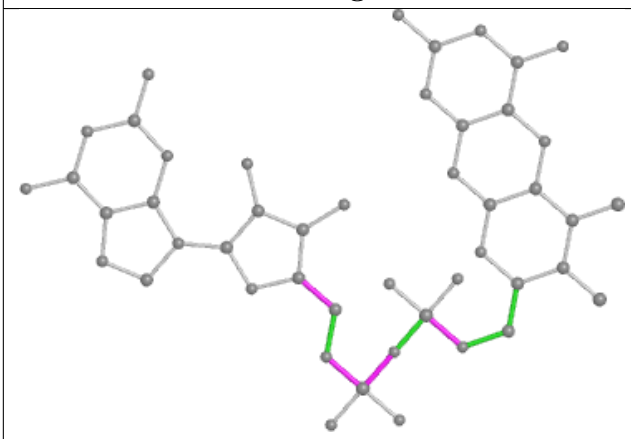
Ligand MGD A 2001



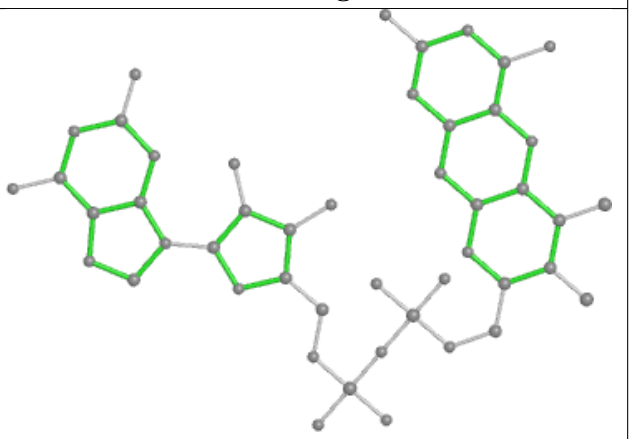
Bond lengths



Bond angles

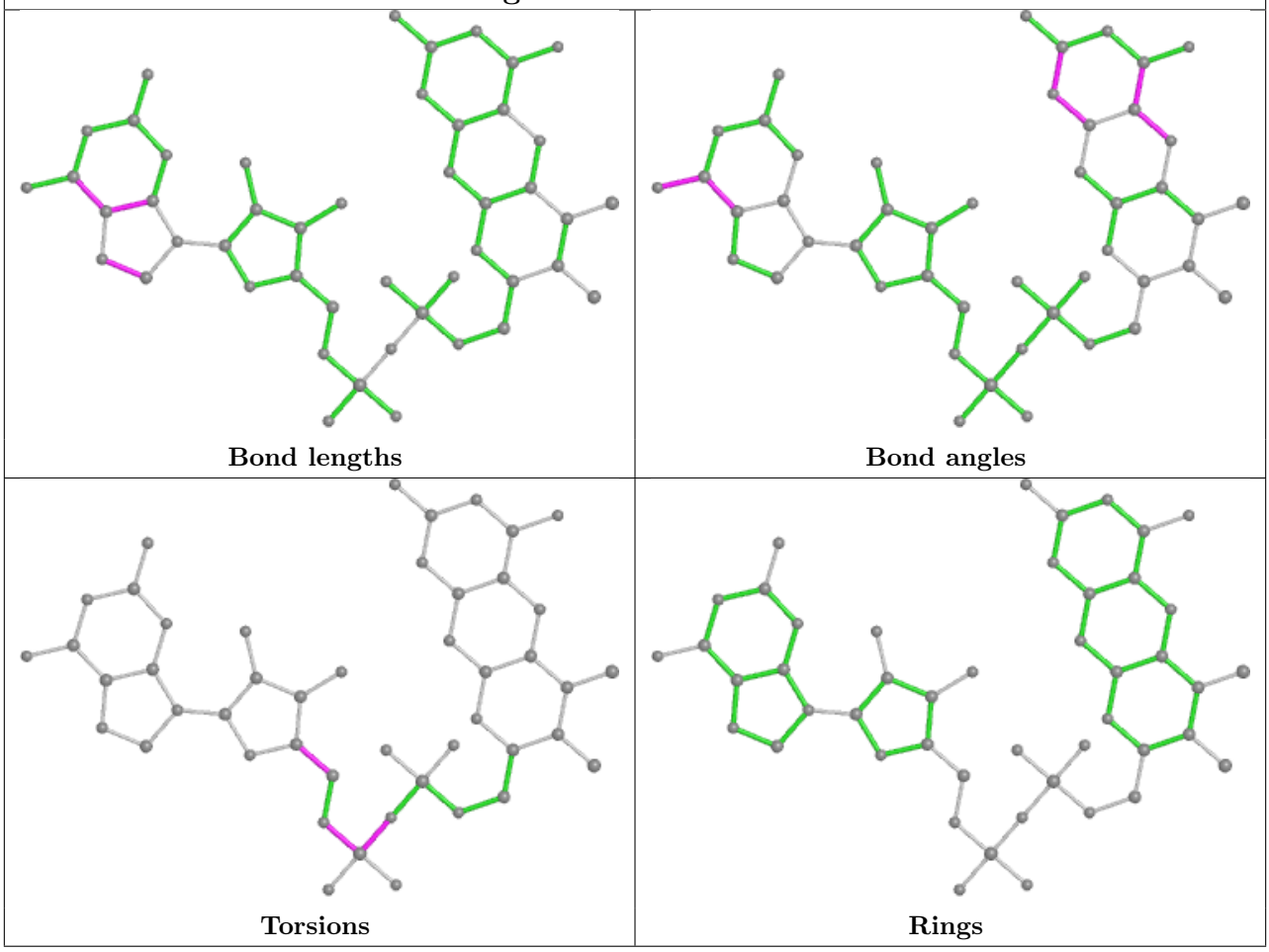


Torsions



Rings

Ligand MGD C 2001



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.

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	843/844 (99%)	-0.21	4 (0%) 91 91	0, 8, 27, 50	0
1	C	843/844 (99%)	-0.21	5 (0%) 89 89	0, 10, 27, 46	0
1	E	843/844 (99%)	-0.41	1 (0%) 95 96	0, 4, 18, 35	0
1	G	843/844 (99%)	-0.39	2 (0%) 95 96	0, 1, 16, 38	0
2	B	132/162 (81%)	0.38	5 (3%) 40 43	5, 20, 39, 47	0
2	D	132/162 (81%)	0.77	15 (11%) 5 4	11, 24, 43, 51	0
2	F	132/162 (81%)	-0.33	1 (0%) 86 87	1, 9, 17, 28	0
2	H	131/162 (80%)	-0.35	0 100 100	1, 10, 18, 25	0
3	I	105/117 (89%)	1.37	24 (22%) 0 0	15, 37, 57, 67	0
3	J	102/117 (87%)	2.48	60 (58%) 0 0	23, 57, 85, 103	0
3	K	103/117 (88%)	1.57	36 (34%) 0 0	17, 40, 63, 81	0
3	L	102/117 (87%)	1.77	38 (37%) 0 0	20, 41, 60, 62	0
All	All	4311/4492 (95%)	-0.05	191 (4%) 34 37	0, 8, 43, 103	0

The worst 5 of 191 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	J	96	ALA	8.3
3	J	80	TRP	7.4
3	K	124	PHE	6.2
3	L	40	VAL	5.9
3	J	74	ALA	5.7

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 (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

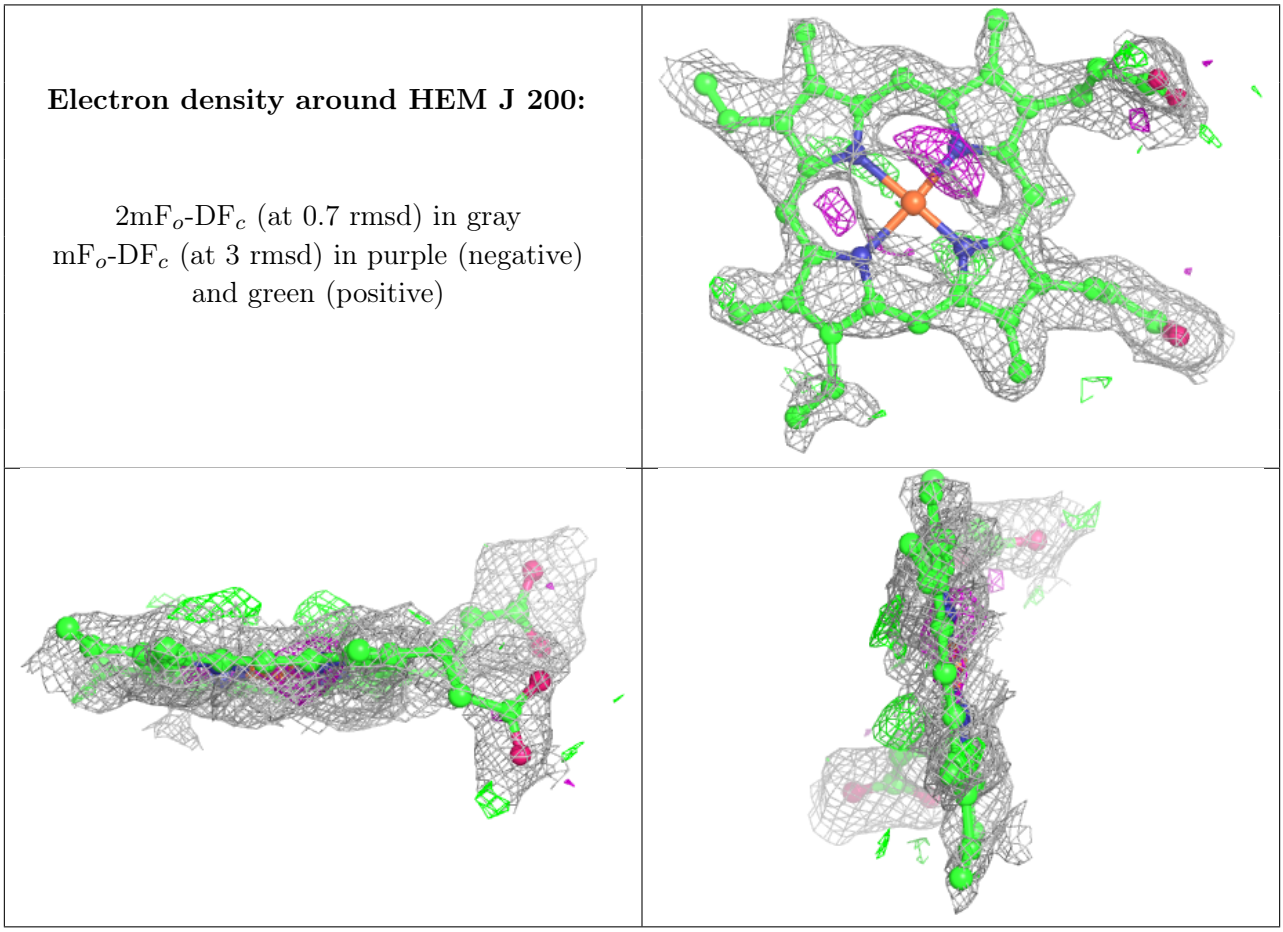
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' 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
11	PG4	E	2007	13/13	0.80	0.24	12,21,31,34	0
9	HEM	J	200	43/43	0.84	0.23	37,41,44,45	0
11	PG4	G	2006	13/13	0.84	0.34	0,0,0,0	13
10	GOL	E	2006	6/6	0.88	0.39	0,0,0,0	6
9	HEM	I	200	43/43	0.90	0.20	24,27,29,29	0
9	HEM	L	200	43/43	0.92	0.21	22,29,34,36	0
12	HEC	K	200	43/43	0.93	0.18	18,20,22,23	0
8	FES	D	201	4/4	0.97	0.07	11,13,14,14	0
4	MGD	C	2002	47/47	0.97	0.10	3,4,4,4	0
6	F3S	C	2004	7/7	0.98	0.06	2,2,3,3	0
7	O	A	2005	1/1	0.98	0.24	30,30,30,30	0
8	FES	B	201	4/4	0.98	0.05	9,10,10,10	0
4	MGD	A	2002	47/47	0.98	0.10	0,0,0,0	0
4	MGD	C	2001	47/47	0.98	0.11	0,0,0,0	0
4	MGD	A	2001	47/47	0.98	0.11	0,0,0,0	0
4	MGD	E	2001	47/47	0.98	0.11	0,0,0,0	0
4	MGD	E	2002	47/47	0.98	0.09	0,0,0,0	0
4	MGD	G	2001	47/47	0.98	0.13	0,0,0,0	0
4	MGD	G	2002	47/47	0.98	0.12	0,0,0,0	0
6	F3S	A	2004	7/7	0.98	0.08	2,2,2,2	0
8	FES	F	201	4/4	0.99	0.06	7,7,7,7	0
8	FES	H	201	4/4	0.99	0.05	5,5,5,6	0
6	F3S	E	2004	7/7	0.99	0.07	0,0,0,0	0
6	F3S	G	2004	7/7	0.99	0.08	0,0,0,0	0
5	4MO	G	2003	1/1	0.99	0.08	1,1,1,1	0
7	O	E	2005	1/1	0.99	0.14	30,30,30,30	0
7	O	G	2005	1/1	0.99	0.20	30,30,30,30	0
5	4MO	C	2003	1/1	0.99	0.07	2,2,2,2	0
5	4MO	E	2003	1/1	0.99	0.07	0,0,0,0	0
7	O	C	2005	1/1	1.00	0.16	30,30,30,30	0
5	4MO	A	2003	1/1	1.00	0.06	2,2,2,2	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.

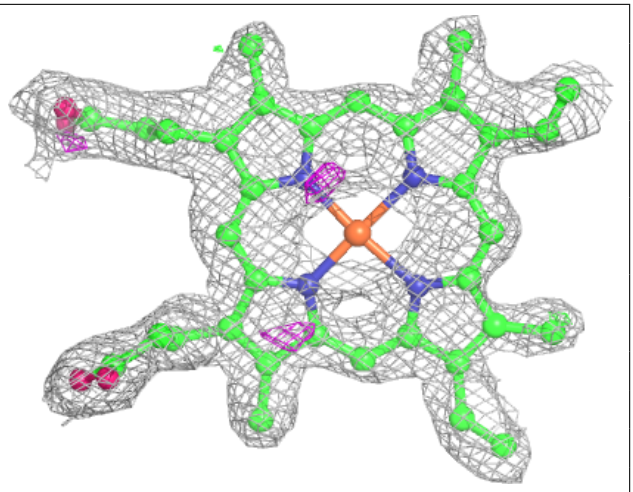
Electron density around HEM J 200:

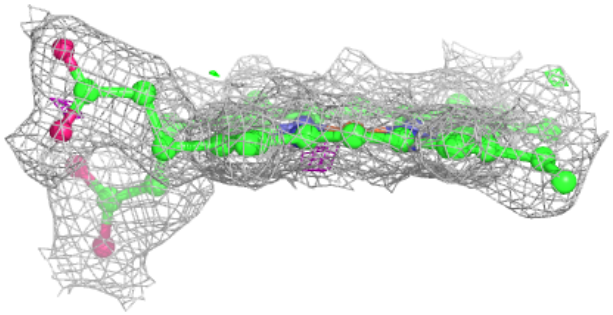
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

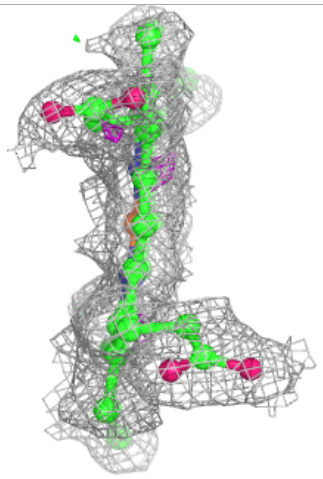


Electron density around HEM I 200:

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

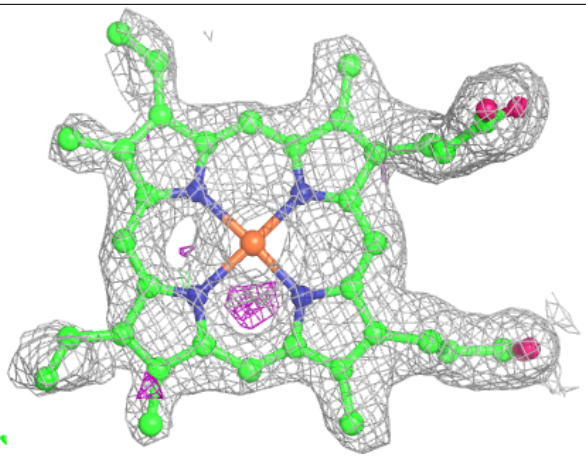


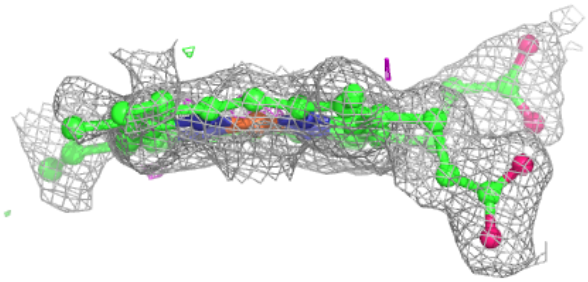


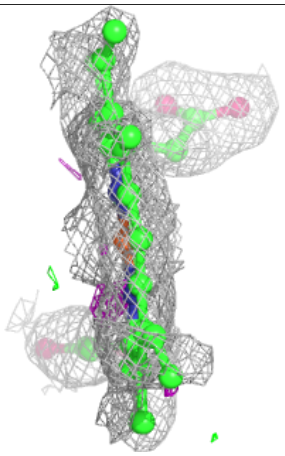


Electron density around HEM L 200:

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

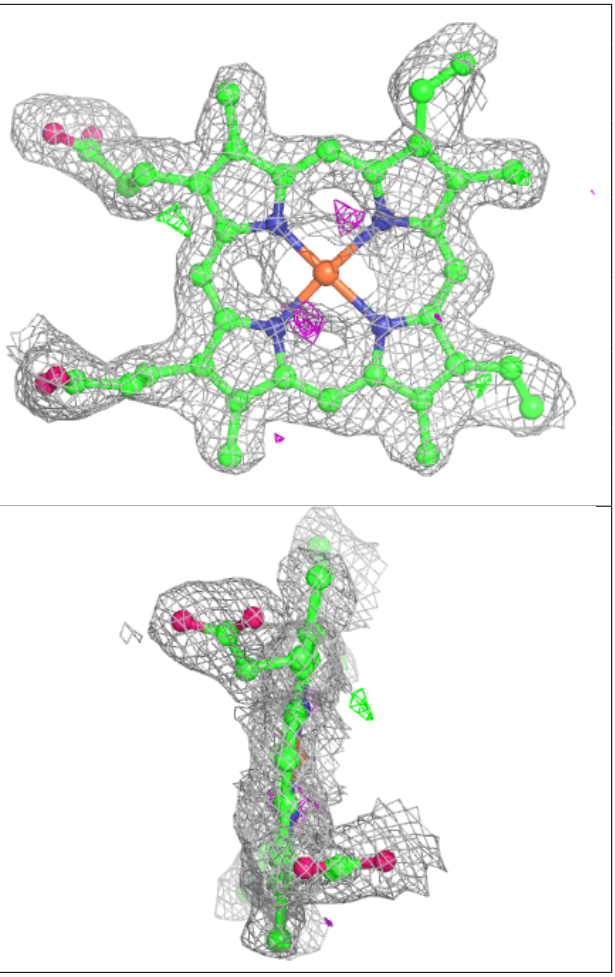






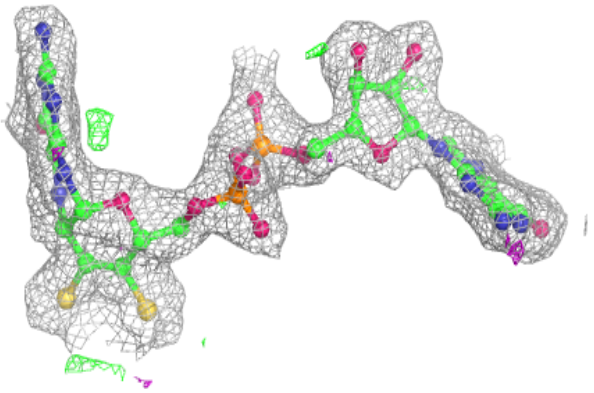
Electron density around HEC K 200:

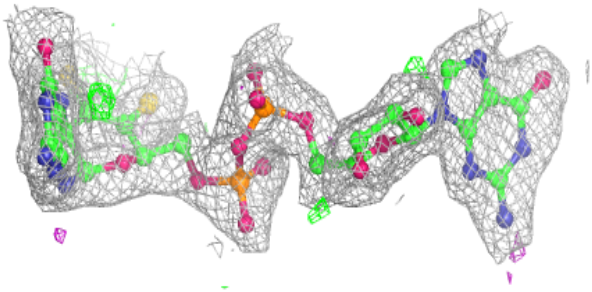
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

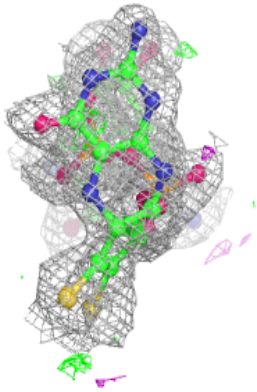


Electron density around MGD C 2002:

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

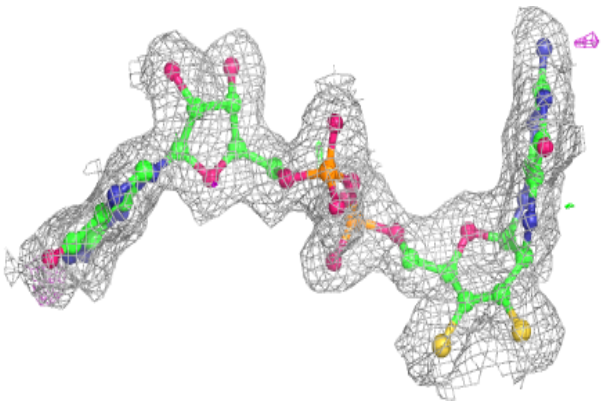


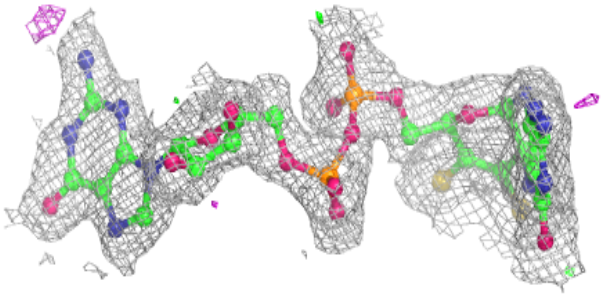


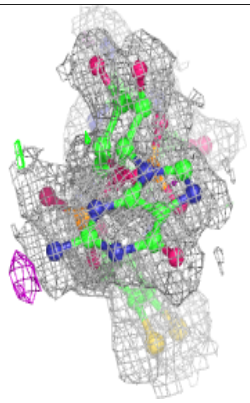


Electron density around MGD A 2002:

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

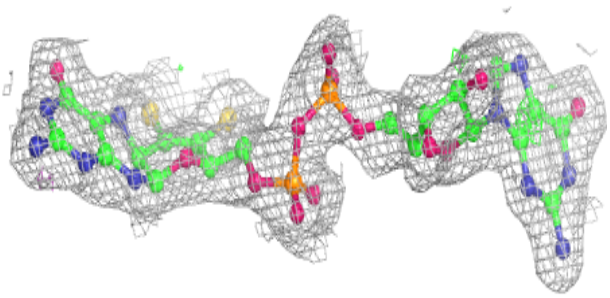


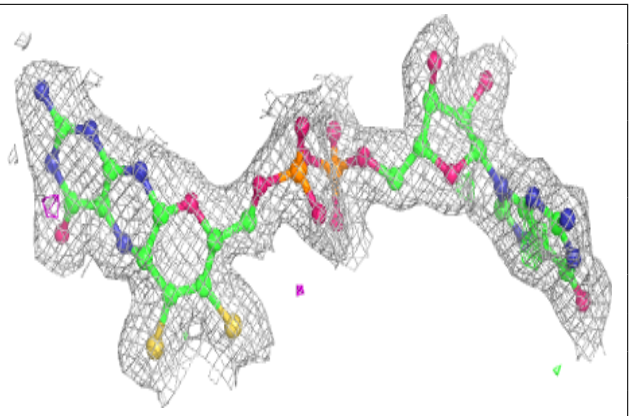


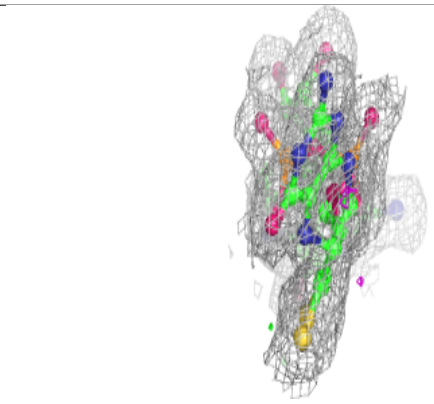


Electron density around MGD C 2001:

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

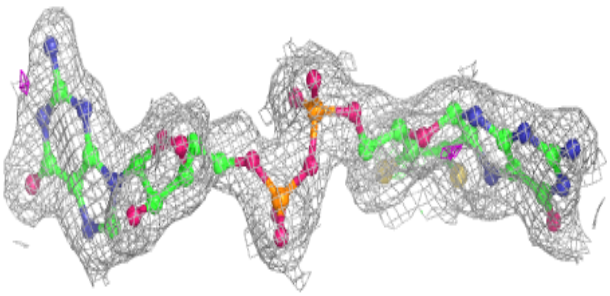


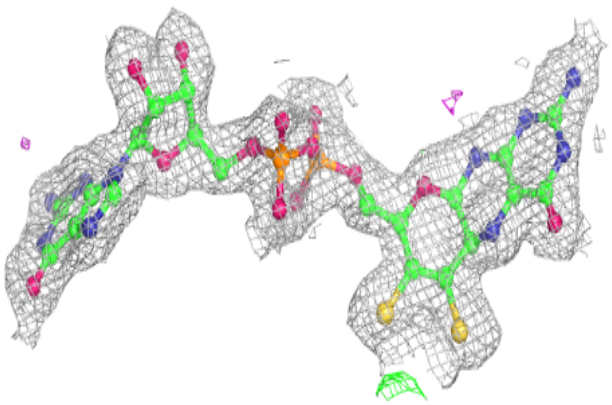


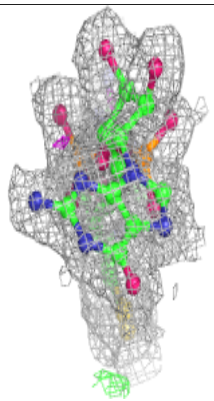


Electron density around MGD A 2001:

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

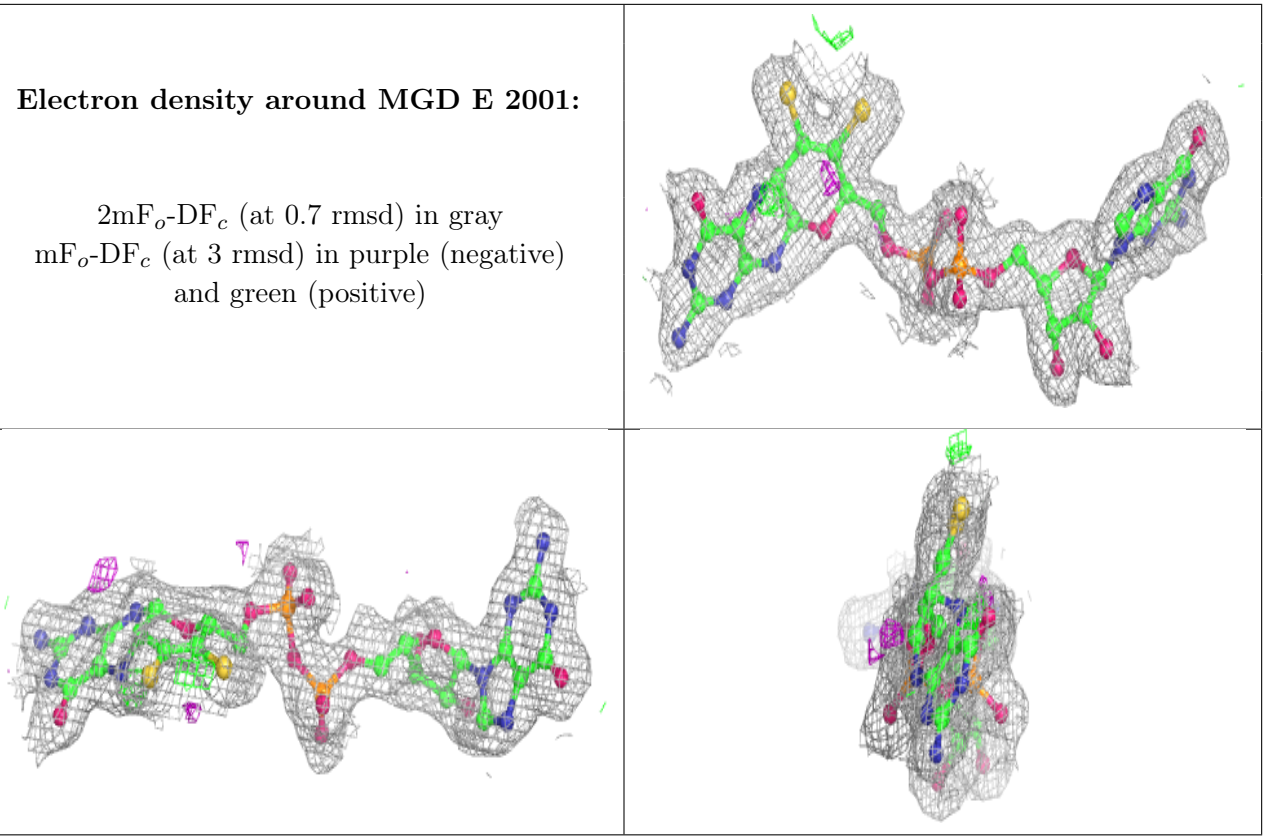






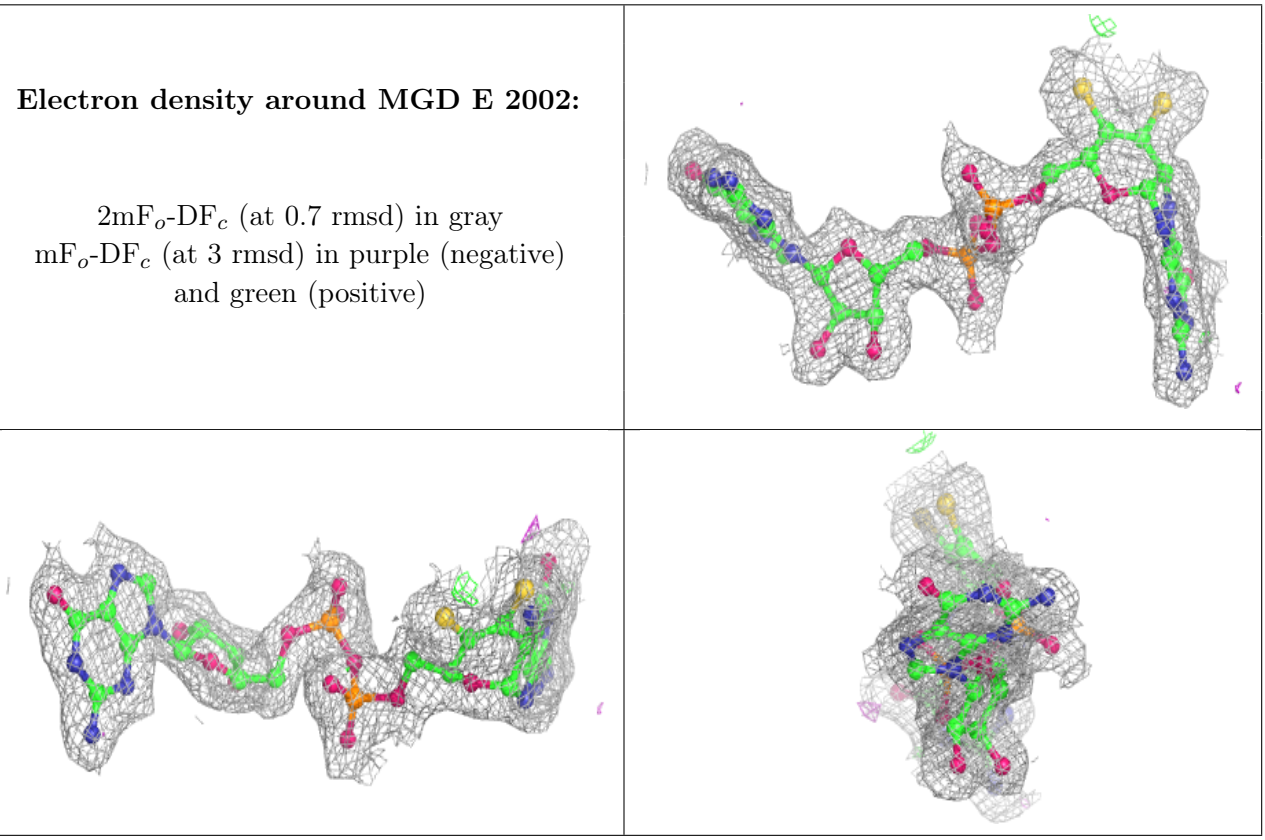
Electron density around MGD E 2001:

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



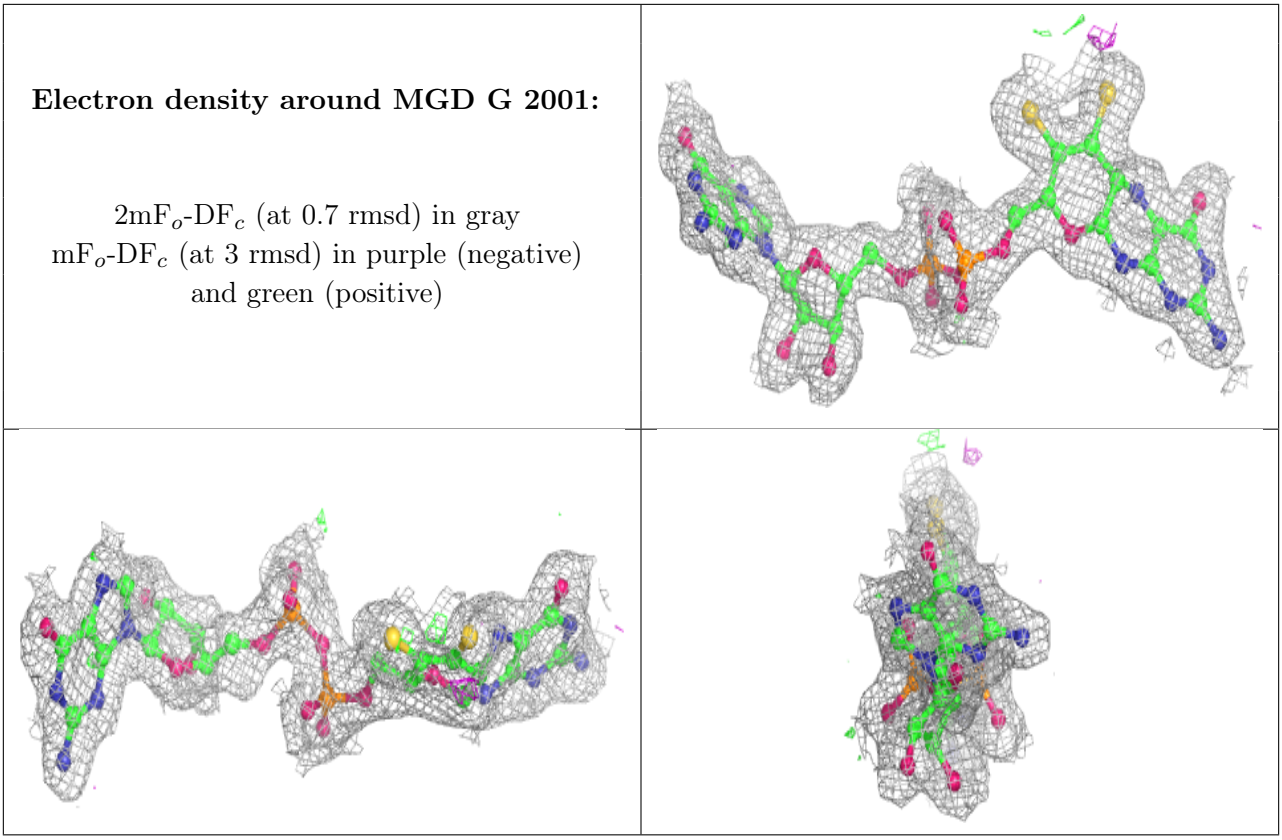
Electron density around MGD E 2002:

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



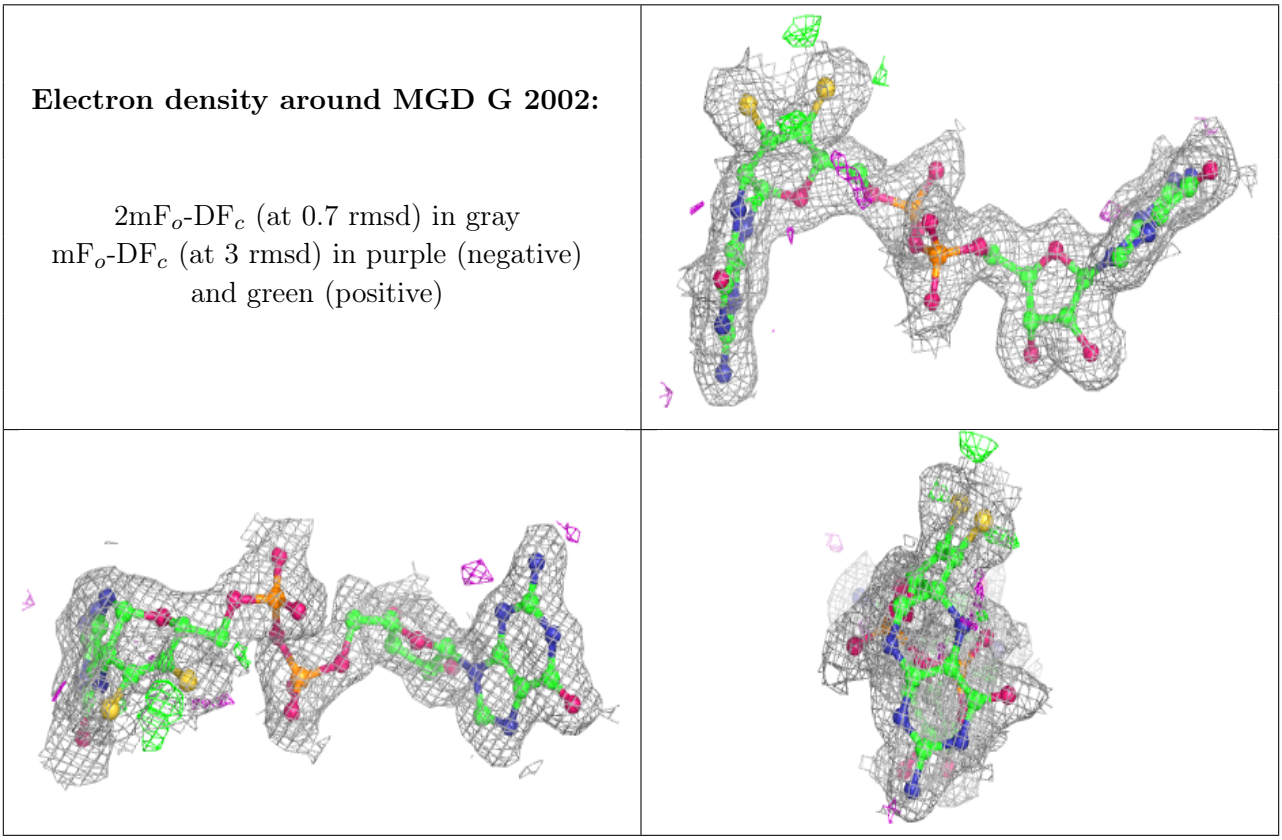
Electron density around MGD G 2001:

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



Electron density around MGD G 2002:

$2mF_o$-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.