wwPDB X-ray Structure Validation Summary Report (i)

May 26, 2020 – 01:02 pm BST

PDB ID	:	2D2N
Title	:	Structure of an extracellular giant hemoglobin of the gutless beard worm Oligobrachia mashikoi
Authors	:	Numoto, N.; Nakagawa, T.; Kita, A.; Sasayama, Y.; Fukumori, Y.; Miki, K.
Deposited on	:	2005-09-12
Resolution	:	3.20 Å(reported)

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

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

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

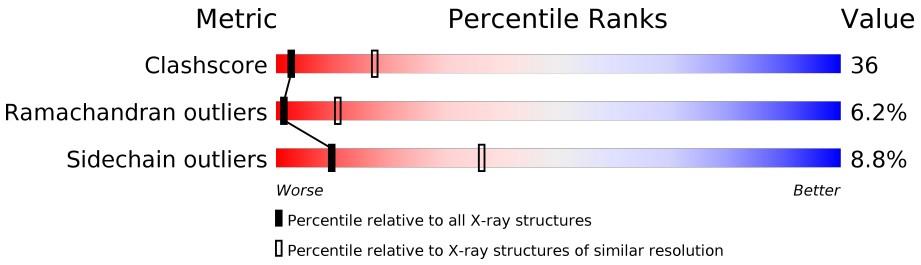
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 3.20 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))
Clashscore	141614	1253 (3.20-3.20)
Ramachandran outliers	138981	1234 (3.20-3.20)
Sidechain outliers	138945	1233 (3.20-3.20)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	140	39% 53% 7%
2	B	142	44% 48% 8%
3	C	147	44% 46% 10%
4	D	145	45% 46% 9%

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 4452 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 Giant hemoglobin, A1(b) globin chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	140	Total	C	N	O	S	0	0	0
			1065	656	196	206	7			

- Molecule 2 is a protein called Giant hemoglobin, A2(a5) globin chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	142	Total	C	N	O	S	0	0	0
			1075	662	198	209	6			

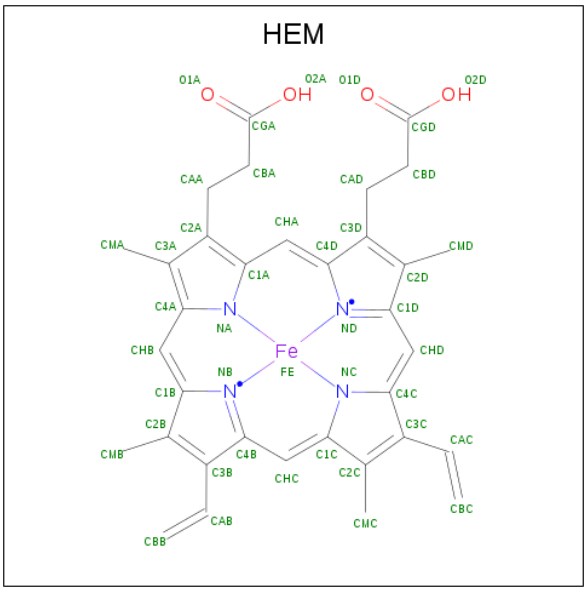
- Molecule 3 is a protein called Giant hemoglobin, B2(c) globin chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	146	Total	C	N	O	S	0	0	0
			1088	668	201	210	9			

- Molecule 4 is a protein called Giant hemoglobin, B1(d) globin chain.

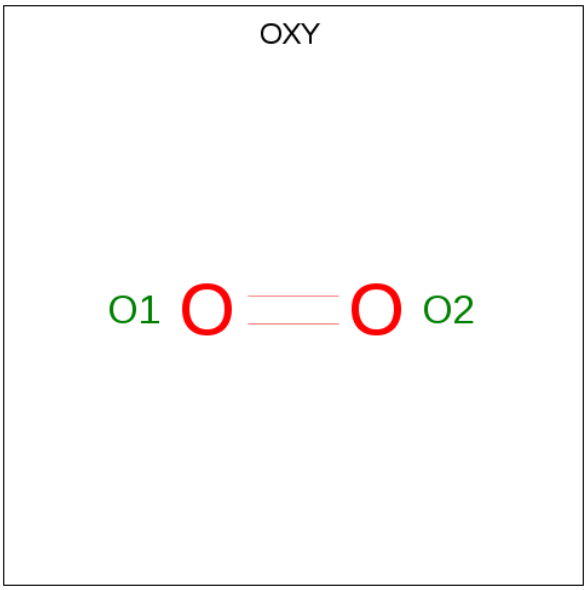
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	145	Total	C	N	O	S	0	0	0
			1036	642	178	209	7			

- Molecule 5 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (formula: $C_{34}H_{32}FeN_4O_4$).



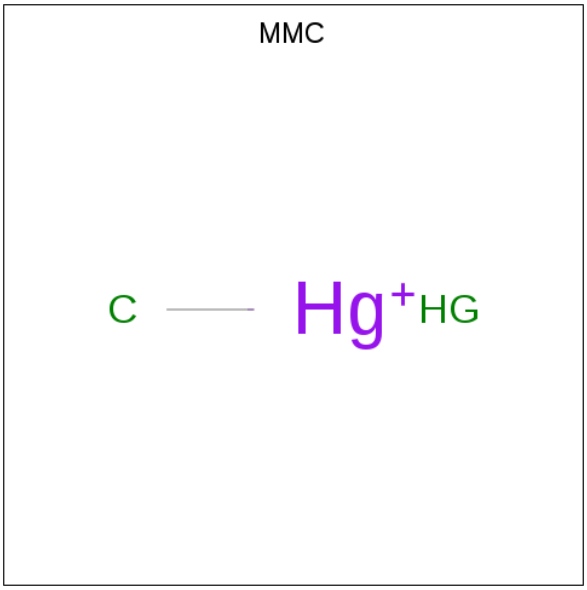
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	A	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
5	B	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
5	C	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
5	D	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		

- Molecule 6 is OXYGEN MOLECULE (three-letter code: OXY) (formula: O_2).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total O 2 2	0	0
6	B	1	Total O 2 2	0	0
6	C	1	Total O 2 2	0	0
6	D	1	Total O 2 2	0	0

- Molecule 7 is METHYL MERCURY ION (three-letter code: MMC) (formula: CH_3Hg).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	1	Total C Hg 2 1 1	0	0
7	B	1	Total C Hg 2 1 1	0	0
7	C	1	Total C Hg 2 1 1	0	0
7	D	1	Total C Hg 2 1 1	0	0

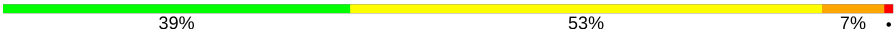
3 Residue-property plots (i)

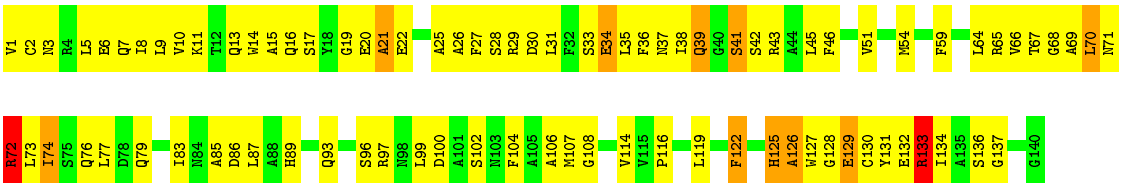
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according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

Note EDS was not executed.

- Molecule 1: Giant hemoglobin, A1(b) globin chain

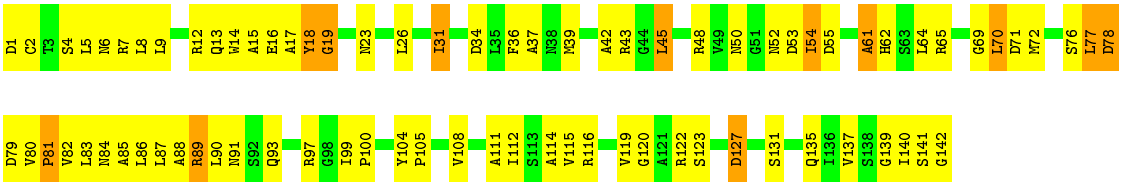
Chain A:





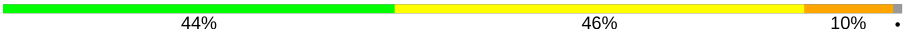
- Molecule 2: Giant hemoglobin, A2(a5) globin chain

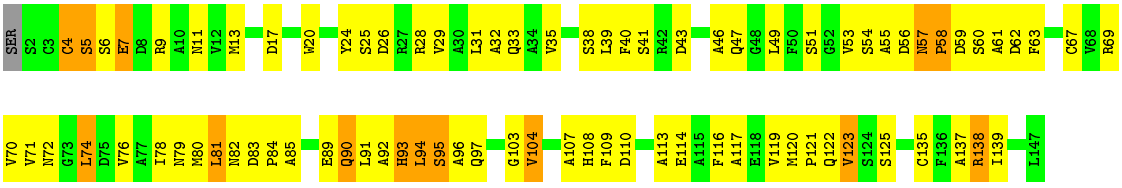
Chain B:



- Molecule 3: Giant hemoglobin, B2(c) globin chain

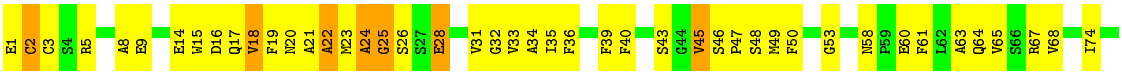
Chain C:





- Molecule 4: Giant hemoglobin, B1(d) globin chain

Chain D:




T80
N81
T84
L88
L89
S90
H91
L92
N93
R97
V102
T103
V107
T108
H109
L110
S111
Q112
A113
I114
S115
S116
V117
V118
A119
Q120
S124
A125
H126
I127
D128
A129
W130
M134
A138
A139
G140
A143
G144
L145

4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	H 3 2	Depositor
Cell constants a, b, c, α, β, γ	111.75Å 111.75Å 276.71Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	50.00 – 3.20	Depositor
% Data completeness (in resolution range)	97.6 (50.00-3.20)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	0.09	Depositor
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.242 , 0.297	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	4452	wwPDB-VP
Average B, all atoms (Å^2)	70.0	wwPDB-VP

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: HEM, OXY, MMC

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	0.46	0/1083	0.71	0/1464
2	B	0.48	0/1093	0.67	1/1481 (0.1%)
3	C	0.46	0/1110	0.66	0/1507
4	D	0.45	0/1055	0.68	0/1436
All	All	0.46	0/4341	0.68	1/5888 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	45	LEU	CA-CB-CG	5.24	127.35	115.30

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	1065	0	1014	98	0
2	B	1075	0	1038	65	0
3	C	1088	0	1012	84	0
4	D	1036	0	986	76	0
5	A	43	0	30	3	0
5	B	43	0	30	3	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	C	43	0	30	1	0
5	D	43	0	30	3	0
6	A	2	0	0	0	0
6	B	2	0	0	0	0
6	C	2	0	0	0	0
6	D	2	0	0	0	0
7	A	2	0	0	0	0
7	B	2	0	0	0	0
7	C	2	0	0	0	0
7	D	2	0	0	0	0
All	All	4452	0	4170	311	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 36.

The worst 5 of 311 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:46:PHE:HB2	1:A:51:VAL:HG11	1.24	1.08
2:B:9:LEU:HD21	2:B:13:GLN:HE21	1.42	0.84
3:C:25:SER:HA	3:C:28:ARG:CD	2.10	0.81
1:A:14:TRP:HE1	1:A:71:ASN:ND2	1.79	0.80
3:C:25:SER:HA	3:C:28:ARG:HD3	1.64	0.79

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	A	138/140 (99%)	102 (74%)	27 (20%)	9 (6%)	1 10
2	B	140/142 (99%)	103 (74%)	32 (23%)	5 (4%)	3 23

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
3	C	144/147 (98%)	107 (74%)	25 (17%)	12 (8%)	1 5
4	D	143/145 (99%)	110 (77%)	24 (17%)	9 (6%)	1 10
All	All	565/574 (98%)	422 (75%)	108 (19%)	35 (6%)	1 11

5 of 35 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	21	ALA
1	A	22	GLU
3	C	54	SER
3	C	59	ASP
3	C	94	LEU

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	110/110 (100%)	101 (92%)	9 (8%)	11 41
2	B	113/113 (100%)	103 (91%)	10 (9%)	10 36
3	C	110/111 (99%)	98 (89%)	12 (11%)	6 26
4	D	108/108 (100%)	100 (93%)	8 (7%)	13 46
All	All	441/442 (100%)	402 (91%)	39 (9%)	10 36

5 of 39 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	127	ASP
3	C	7	GLU
4	D	118	VAL
2	B	135	GLN
3	C	4	CYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 22 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	84	ASN
3	C	47	GLN
3	C	141	ASN
2	B	134	ASN
3	C	11	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 carbohydrates in this entry.

5.6 Ligand geometry (i)

12 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	$\#\|Z\| > 2$
6	OXY	A	201	5	1,1,1	0.04	0	-		
7	MMC	A	500	1	0,1,1	0.00	-	-		
6	OXY	D	3201	5	1,1,1	0.04	0	-		
5	HEM	C	200	3,6	27,50,50	0.86	1 (3%)	17,82,82	1.70	5 (29%)
5	HEM	D	200	4,6	27,50,50	0.91	1 (3%)	17,82,82	1.44	3 (17%)
5	HEM	A	200	1,6	27,50,50	0.91	1 (3%)	17,82,82	1.55	5 (29%)
5	HEM	B	200	2,6	27,50,50	0.92	1 (3%)	17,82,82	1.55	4 (23%)
6	OXY	C	2201	5	1,1,1	0.01	0	-		
6	OXY	B	1201	5	1,1,1	0.02	0	-		

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
7	MMC	D	3500	4	0,1,1	0.00	-	-		
7	MMC	B	1500	2	0,1,1	0.00	-	-		
7	MMC	C	2500	3	0,1,1	0.00	-	-		

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
5	HEM	C	200	3,6	-	1/6/54/54	-
5	HEM	D	200	4,6	-	1/6/54/54	-
5	HEM	A	200	1,6	-	1/6/54/54	-
5	HEM	B	200	2,6	-	1/6/54/54	-

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	200	HEM	C3B-C2B	-2.48	1.36	1.40
5	D	200	HEM	C3B-C2B	-2.43	1.37	1.40
5	C	200	HEM	C3B-C2B	-2.29	1.37	1.40
5	B	200	HEM	C3B-C2B	-2.22	1.37	1.40

The worst 5 of 17 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	C	200	HEM	CAD-CBD-CGD	-3.65	106.55	112.67
5	B	200	HEM	CAD-CBD-CGD	-3.14	107.41	112.67
5	D	200	HEM	CMB-C2B-C3B	2.92	130.14	124.68
5	B	200	HEM	CMB-C2B-C3B	2.79	129.90	124.68
5	A	200	HEM	CMC-C2C-C3C	2.67	129.67	124.68

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	C	200	HEM	C2A-CAA-CBA-CGA
5	D	200	HEM	C3D-CAD-CBD-CGD
5	B	200	HEM	C3D-CAD-CBD-CGD
5	A	200	HEM	C3D-CAD-CBD-CGD

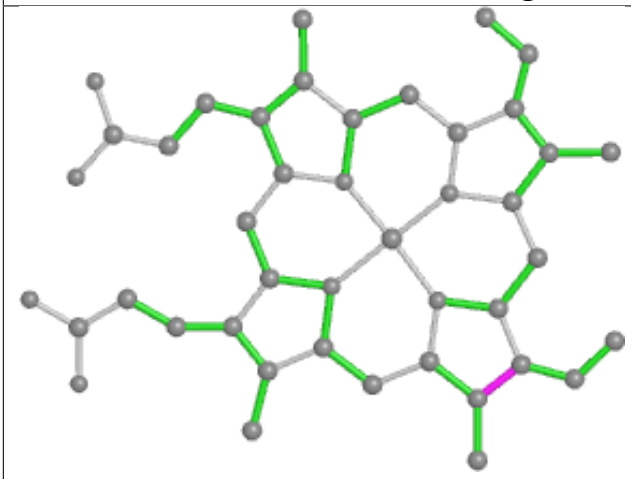
There are no ring outliers.

4 monomers are involved in 10 short contacts:

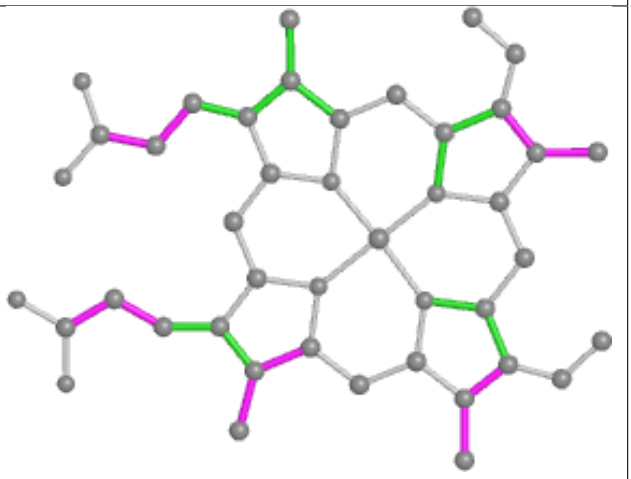
Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	C	200	HEM	1	0
5	D	200	HEM	3	0
5	A	200	HEM	3	0
5	B	200	HEM	3	0

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

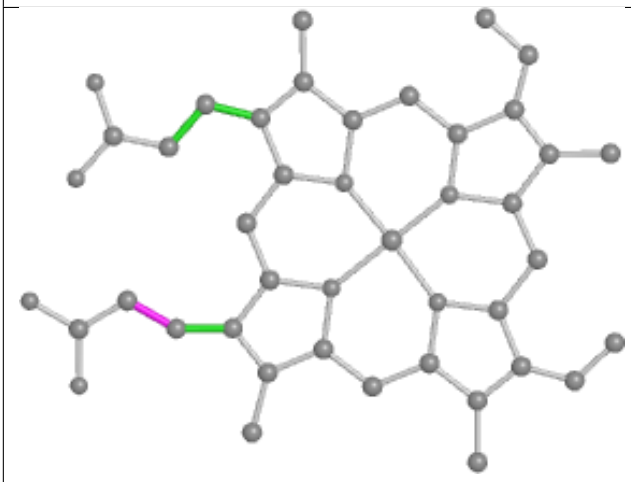
Ligand HEM C 200



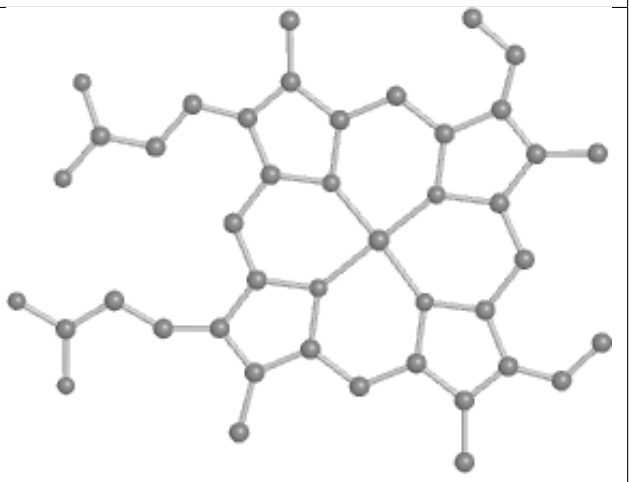
Bond lengths



Bond angles

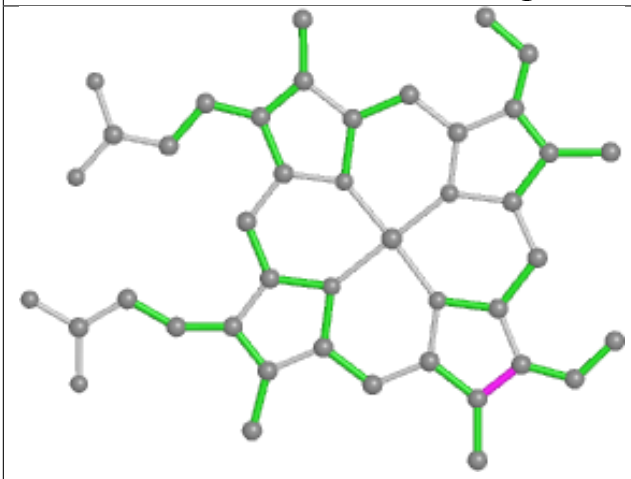


Torsions

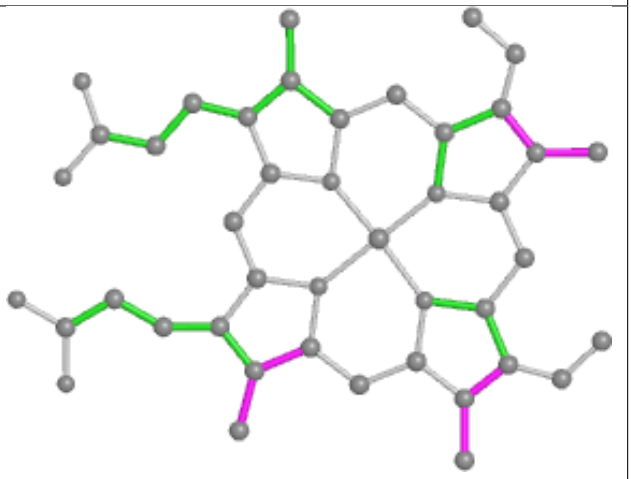


Rings

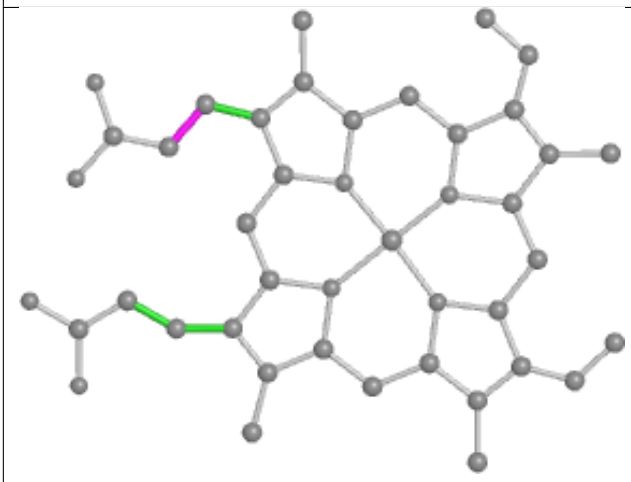
Ligand HEM D 200



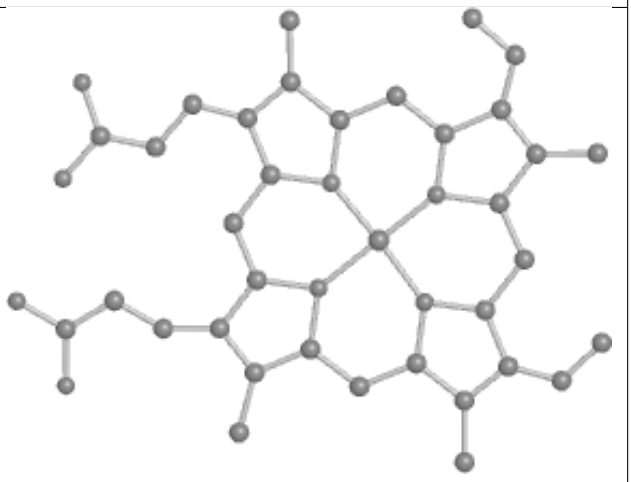
Bond lengths



Bond angles

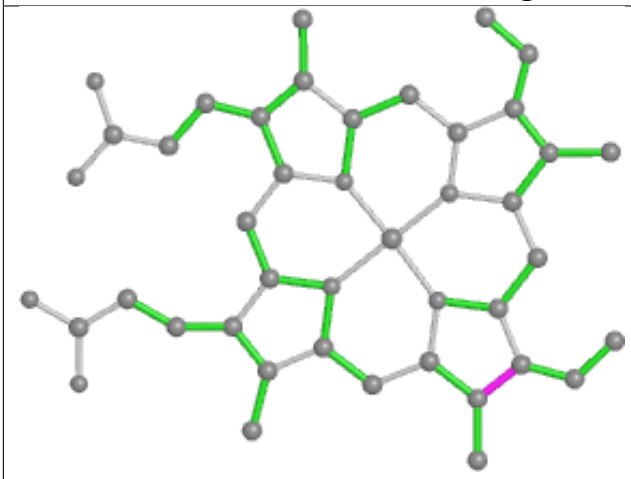


Torsions

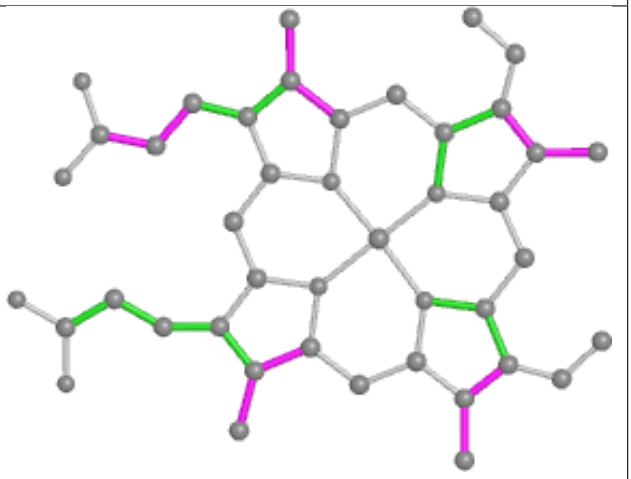


Rings

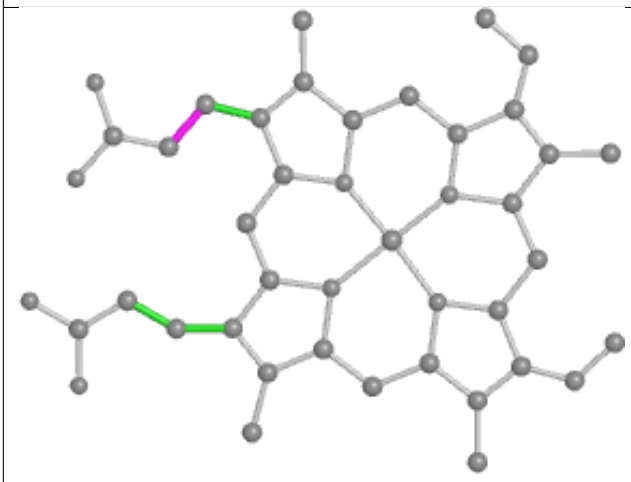
Ligand HEM A 200



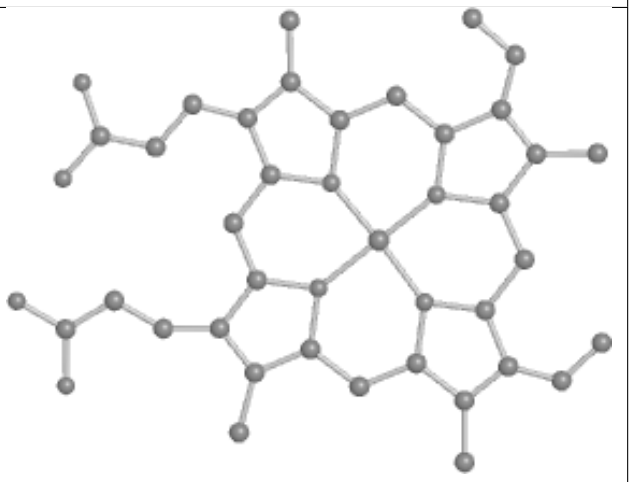
Bond lengths



Bond angles

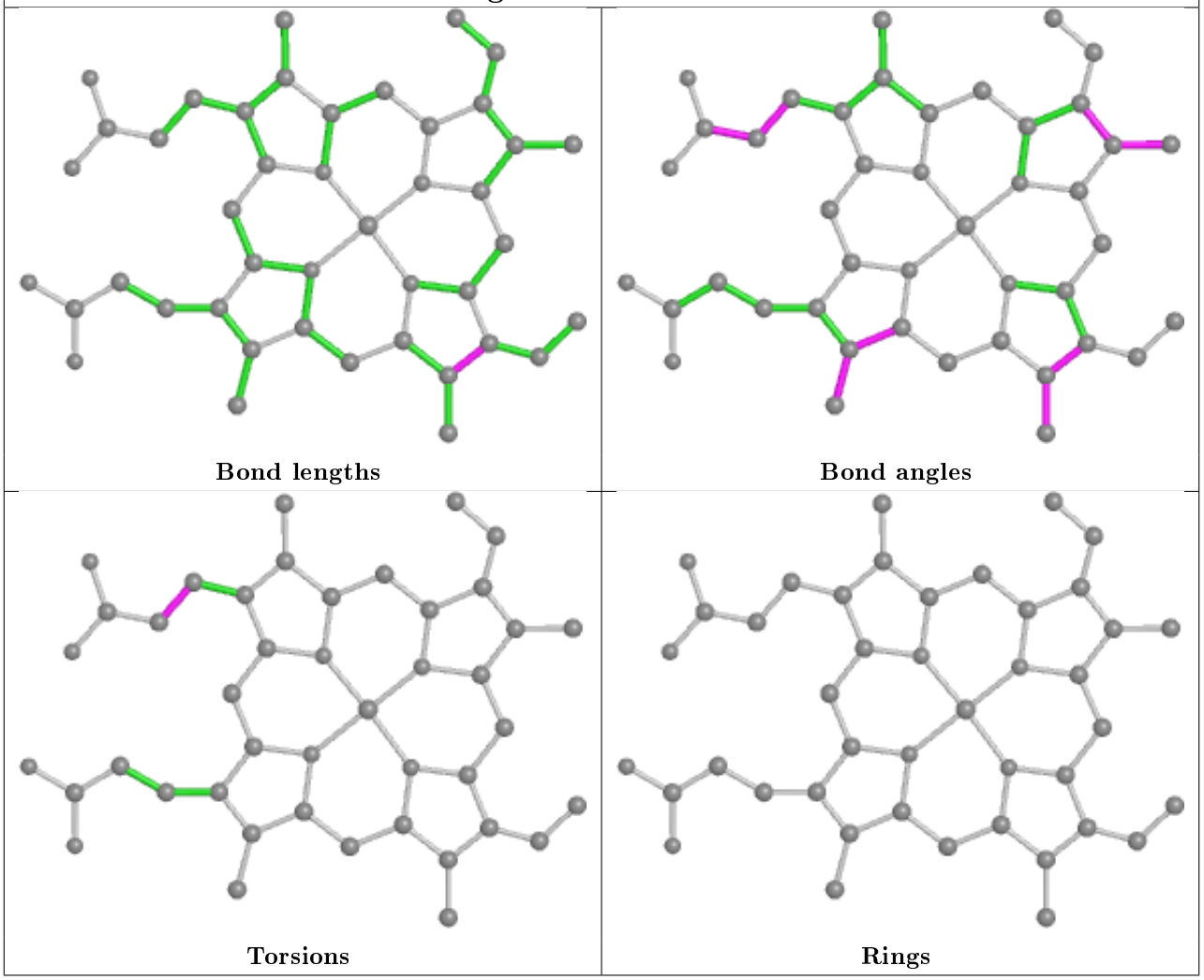


Torsions



Rings

Ligand HEM B 200



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

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

6.4 Ligands

EDS was not executed - this section is therefore empty.

6.5 Other polymers

EDS was not executed - this section is therefore empty.