wwPDB X-ray Structure Validation Summary Report (i)

Aug 6, 2023 – 02:20 PM EDT

PDB ID	:	1KYO
Title	:	YEAST CYTOCHROME BC1 COMPLEX WITH BOUND SUBSTRATE CYTOCHROME C
Authors	:	Lange, C.; Hunte, C.
Deposited on	:	2002-02-05
Resolution	:	2.97 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 (1)) were used in the production of this report:

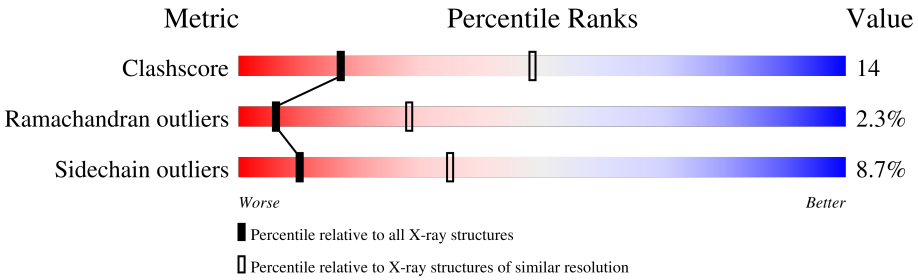
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 2.97 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	3103 (3.00-2.96)
Ramachandran outliers	138981	2993 (3.00-2.96)
Sidechain outliers	138945	2996 (3.00-2.96)

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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	430	70% 26% •
1	L	430	65% 30% 5%
2	B	352	58% 36% 6% •
2	M	352	58% 35% 7% •
3	C	385	70% 27% •
3	N	385	73% 23% •
4	D	248	79% 18% ••
4	O	248	75% 23% ••

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	E	185	69% 28%
5	P	185	75% 24%
6	F	74	70% 28%
6	Q	74	73% 24%
7	G	126	77% 21%
7	R	126	74% 23%
8	H	93	63% 30% 5%
8	S	93	66% 27% 6%
9	I	57	65% 26% 7%
9	T	57	74% 16% 7%
10	J	127	51% 38% 10%
10	U	127	50% 43% 7%
11	K	107	56% 36% 7%
11	V	107	56% 39% 5%
12	W	108	61% 31% 8%

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
14	SMA	C	505	X	-	-	-
14	SMA	N	525	X	-	-	-

2 Entry composition (i)

There are 15 unique types of molecules in this entry. The entry contains 35643 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 UBIQUINOL-CYTOCHROME C REDUCTASE COMPLEX CORE PROTEIN I.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	430	Total	C	N	O	S	0	0	0
			3338	2106	575	651	6			
1	L	430	Total	C	N	O	S	0	0	0
			3338	2106	575	651	6			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	SER	deletion	UNP P07256
A	152	ASP	GLU	conflict	UNP P07256
L	?	-	SER	deletion	UNP P07256
L	152	ASP	GLU	conflict	UNP P07256

- Molecule 2 is a protein called UBIQUINOL-CYTOCHROME C REDUCTASE COMPLEX CORE PROTEIN 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	352	Total	C	N	O	S	0	0	0
			2735	1747	453	534	1			
2	M	352	Total	C	N	O	S	0	0	0
			2735	1747	453	534	1			

- Molecule 3 is a protein called CYTOCHROME B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	385	Total	C	N	O	S	0	0	0
			3089	2080	484	504	21			
3	N	385	Total	C	N	O	S	0	0	0
			3089	2080	484	504	21			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	270	VAL	ASP	SEE REMARK 999	UNP P00163
N	270	VAL	ASP	SEE REMARK 999	UNP P00163

- Molecule 4 is a protein called CYTOCHROME C1, HEME PROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	245	Total	C	N	O	S	0	0	0
			1929	1229	332	359	9			
4	O	245	Total	C	N	O	S	0	0	0
			1929	1229	332	359	9			

- Molecule 5 is a protein called UBIQUINOL-CYTOCHROME C REDUCTASE IRON-SULFUR SUBUNIT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	185	Total	C	N	O	S	0	0	0
			1411	893	242	266	10			
5	P	185	Total	C	N	O	S	0	0	0
			1411	893	242	266	10			

- Molecule 6 is a protein called UBIQUINOL-CYTOCHROME C REDUCTASE COMPLEX 17 KD PROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	F	74	Total	C	N	O	S	0	0	0
			625	390	108	125	2			
6	Q	74	Total	C	N	O	S	0	0	0
			625	390	108	125	2			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	74	ASP	VAL	conflict	UNP P00127
Q	74	ASP	VAL	conflict	UNP P00127

- Molecule 7 is a protein called UBIQUINOL-CYTOCHROME C REDUCTASE COMPLEX 14 KD PROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	G	125	Total	C	N	O	S	0	0	0
			1012	648	172	190	2			
7	R	125	Total	C	N	O	S	0	0	0
			1012	648	172	190	2			

- Molecule 8 is a protein called UBIQUINOL-CYTOCHROME C REDUCTASE COMPLEX UBIQUINONE-BINDING PROTEIN QP-C.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	H	93	Total	C	N	O	S	0	0	0
			773	510	131	130	2			
8	S	93	Total	C	N	O	S	0	0	0
			773	510	131	130	2			

- Molecule 9 is a protein called UBIQUINOL-CYTOCHROME C REDUCTASE COMPLEX 7.3 KD PROTEIN.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
9	I	53	Total	C	N	O	0	0	0
			436	292	71	73			
9	T	53	Total	C	N	O	0	0	0
			436	292	71	73			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
I	55	LYS	ARG	conflict	UNP P22289
T	55	LYS	ARG	conflict	UNP P22289

- Molecule 10 is a protein called HEAVY CHAIN (VH) OF FV-FRAGMENT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	J	127	Total	C	N	O	S	0	0	0
			1015	644	167	201	3			
10	U	127	Total	C	N	O	S	0	0	0
			1015	644	167	201	3			

- Molecule 11 is a protein called LIGHT CHAIN (VL) OF FV-FRAGMENT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	K	107	Total	C	N	O	S	0	0	0
			842	536	141	163	2			
11	V	107	Total	C	N	O	S	0	0	0
			842	536	141	163	2			

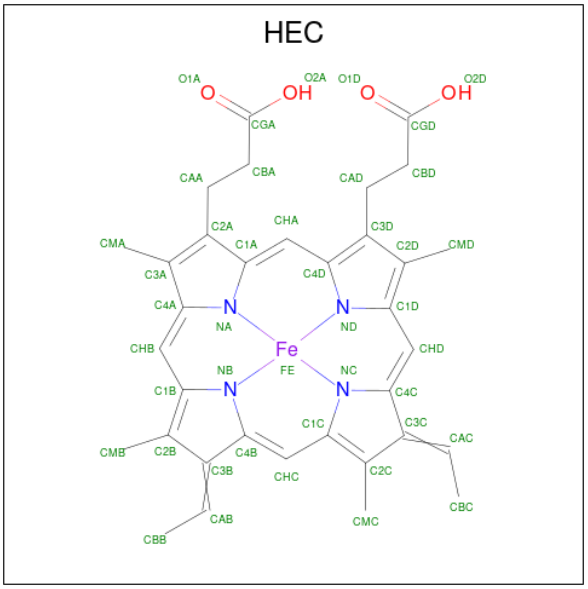
- Molecule 12 is a protein called CYTOCHROME C, ISO-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	W	108	Total	C	N	O	S	0	0	0
			850	537	151	157	5			

There is a discrepancy between the modelled and reference sequences:

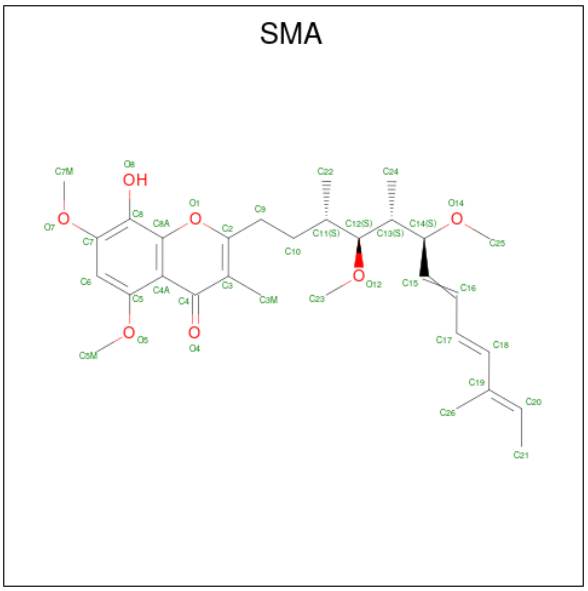
Chain	Residue	Modelled	Actual	Comment	Reference
W	77	M3L	LYS	modified residue	UNP P00044

- Molecule 13 is HEME C (three-letter code: HEC) (formula: $C_{34}H_{34}FeN_4O_4$).



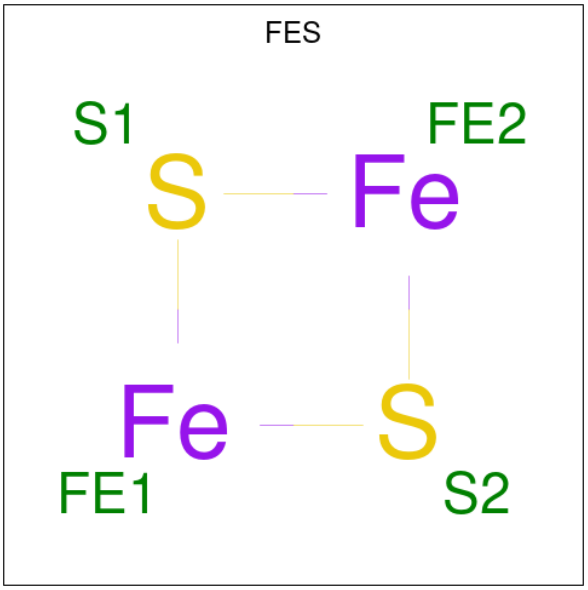
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
13	C	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
13	C	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
13	D	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
13	N	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
13	N	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
13	O	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
13	W	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		

- Molecule 14 is STIGMATELLIN A (three-letter code: SMA) (formula: $C_{30}H_{42}O_7$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
14	C	1	Total C O 37 30 7	0	0
14	N	1	Total C O 37 30 7	0	0

- Molecule 15 is FE2/S2 (INORGANIC) CLUSTER (three-letter code: FES) (formula: Fe_2S_2).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
15	E	1	Total Fe S 4 2 2	0	0
15	P	1	Total Fe S 4 2 2	0	0

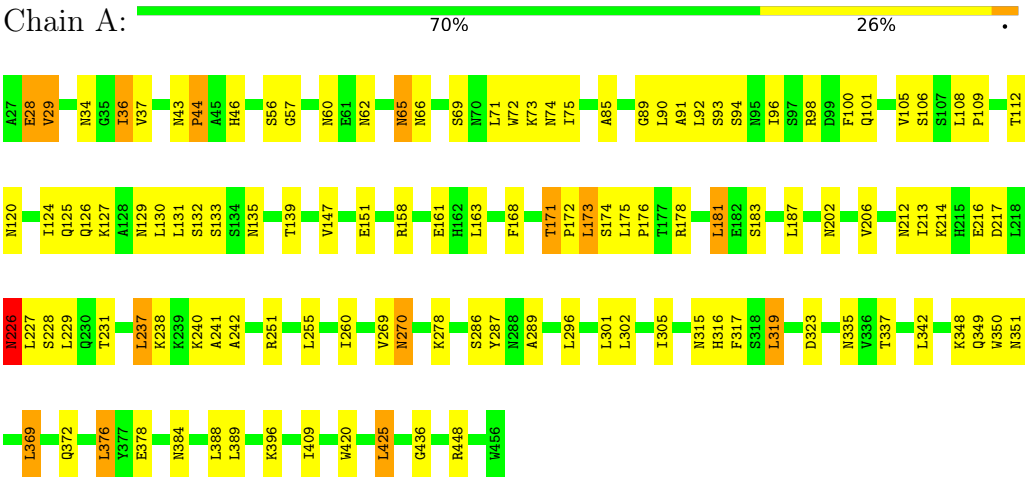
3 Residue-property plots (i)

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. 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.

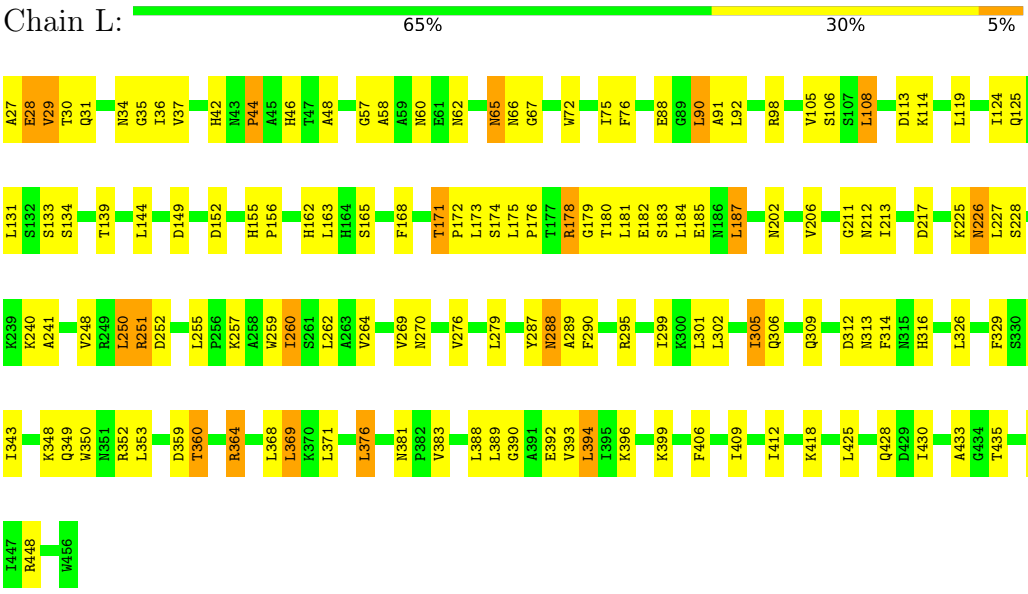
• Molecule 1: UBIQUINOL-CYTOCHROME C REDUCTASE COMPLEX CORE PROTEIN I

Chain A:



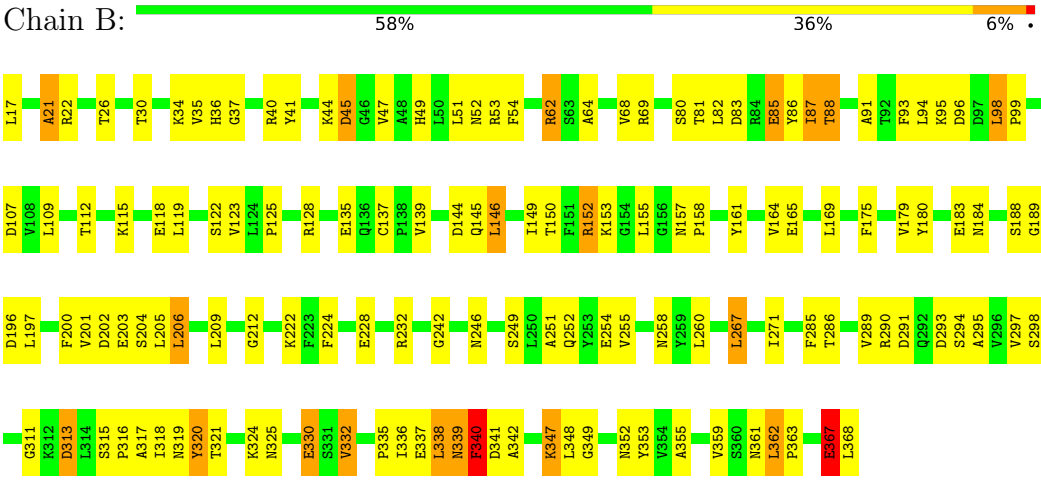
• Molecule 1: UBIQUINOL-CYTOCHROME C REDUCTASE COMPLEX CORE PROTEIN I

Chain L:



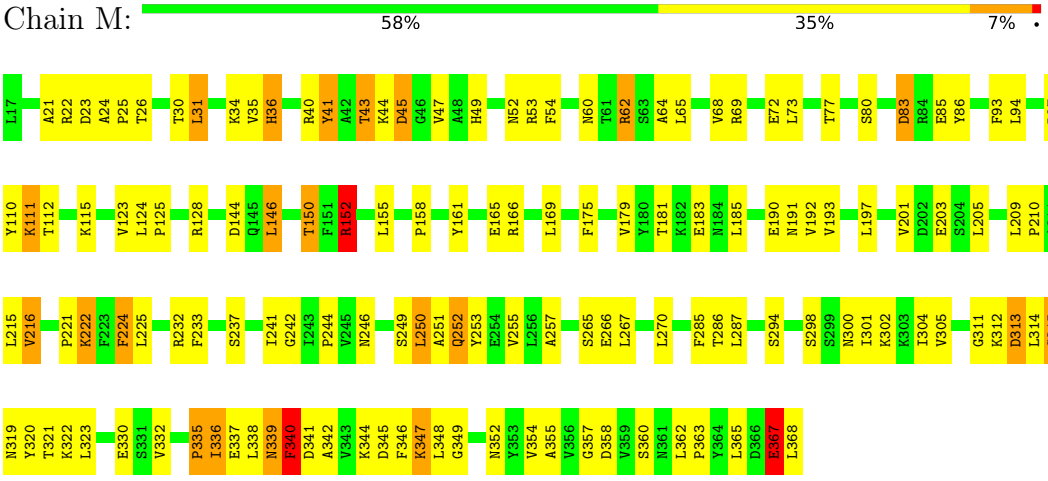
• Molecule 2: UBIQUINOL-CYTOCHROME C REDUCTASE COMPLEX CORE PROTEIN 2

Chain B:



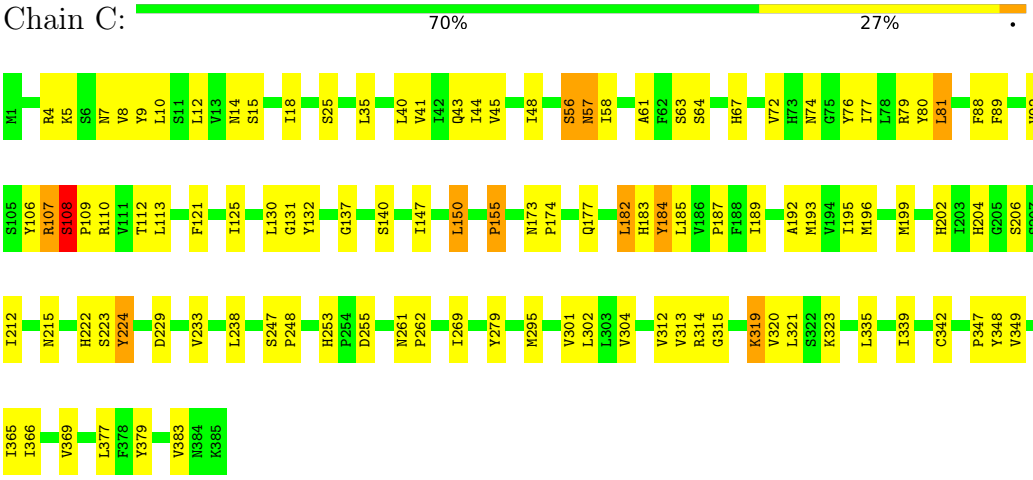
• Molecule 2: UBIQUINOL-CYTOCHROME C REDUCTASE COMPLEX CORE PROTEIN 2

Chain M:



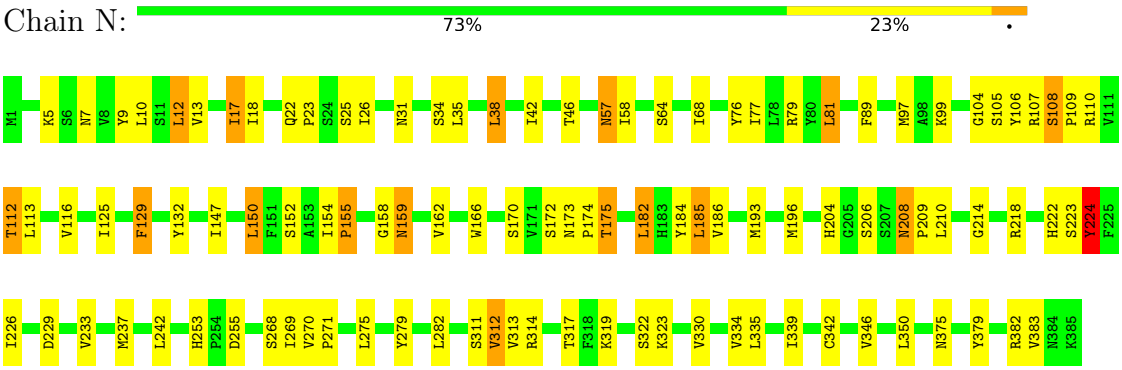
• Molecule 3: CYTOCHROME B

Chain C:



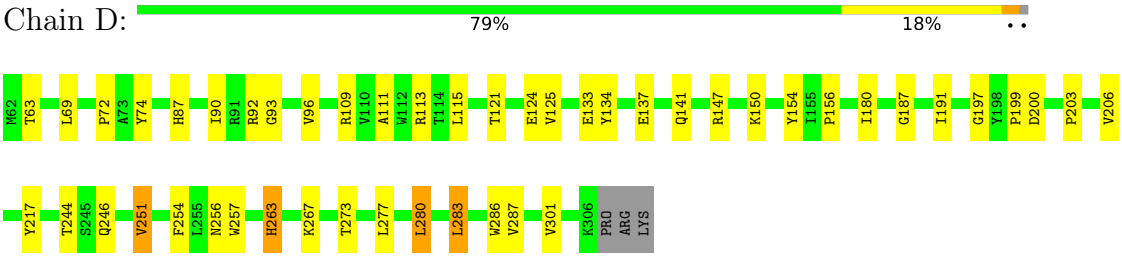
• Molecule 3: CYTOCHROME B

Chain N:



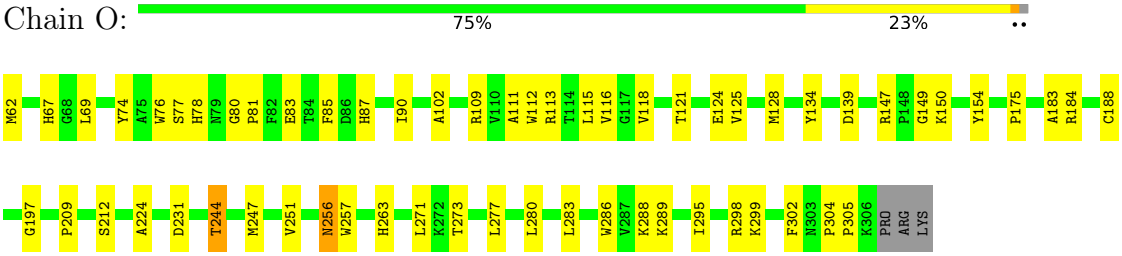
• Molecule 4: CYTOCHROME C1, HEME PROTEIN

Chain D:



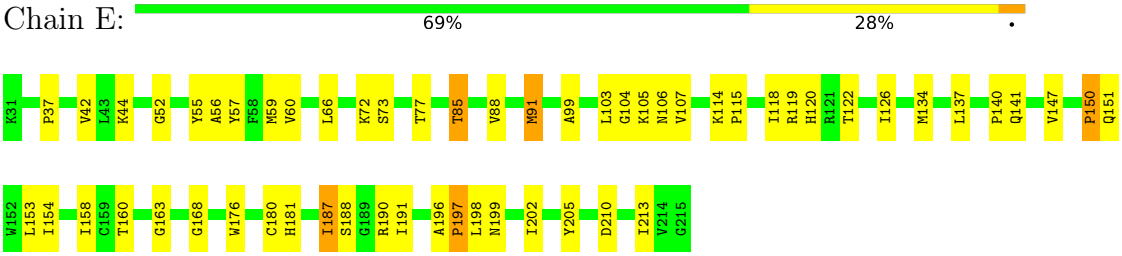
• Molecule 4: CYTOCHROME C1, HEME PROTEIN

Chain O:



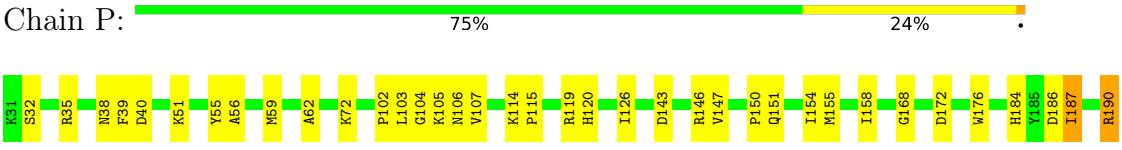
• Molecule 5: UBIQUINOL-CYTOCHROME C REDUCTASE IRON-SULFUR SUBUNIT

Chain E:



• Molecule 5: UBIQUINOL-CYTOCHROME C REDUCTASE IRON-SULFUR SUBUNIT

Chain P:

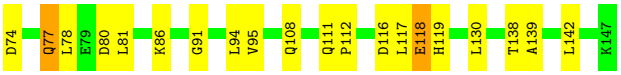


• Molecule 6: UBIQUINOL-CYTOCHROME C REDUCTASE COMPLEX 17 KD PROTEIN

Chain F:

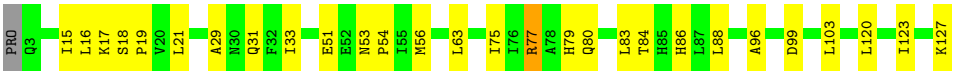
• Molecule 6: UBIQUINOL-CYTOCHROME C REDUCTASE COMPLEX 17 KD PROTEIN

Chain Q:



• Molecule 7: UBIQUINOL-CYTOCHROME C REDUCTASE COMPLEX 14 KD PROTEIN

Chain G:

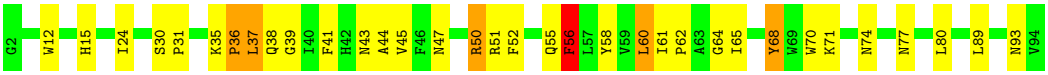


• Molecule 7: UBIQUINOL-CYTOCHROME C REDUCTASE COMPLEX 14 KD PROTEIN

Chain R:

• Molecule 8: UBIQUINOL-CYTOCHROME C REDUCTASE COMPLEX UBIQUINONE-BINDING PROTEIN QP-C

Chain H:

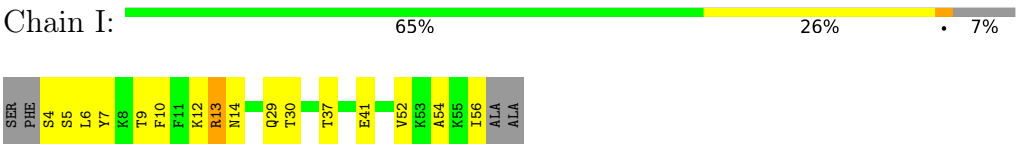


• Molecule 8: UBIQUINOL-CYTOCHROME C REDUCTASE COMPLEX UBIQUINONE-BINDING PROTEIN QP-C

Chain S:

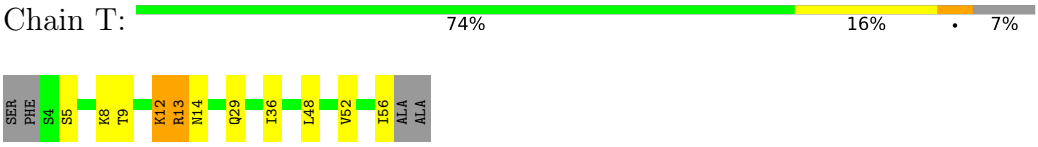
• Molecule 9: UBIQUINOL-CYTOCHROME C REDUCTASE COMPLEX 7.3 KD PROTEIN

Chain I:



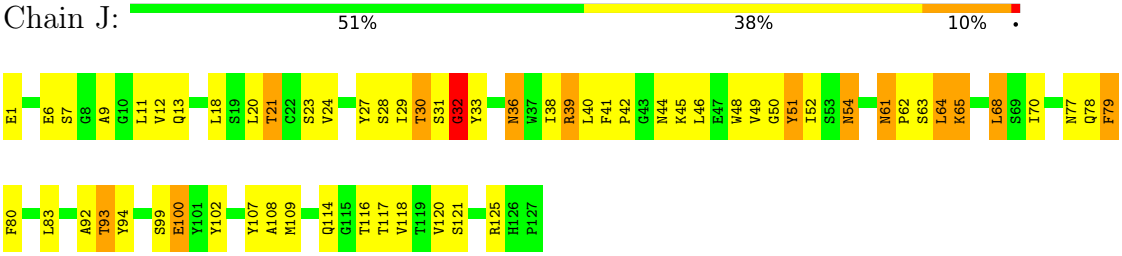
• Molecule 9: UBIQUINOL-CYTOCHROME C REDUCTASE COMPLEX 7.3 KD PROTEIN

Chain T:



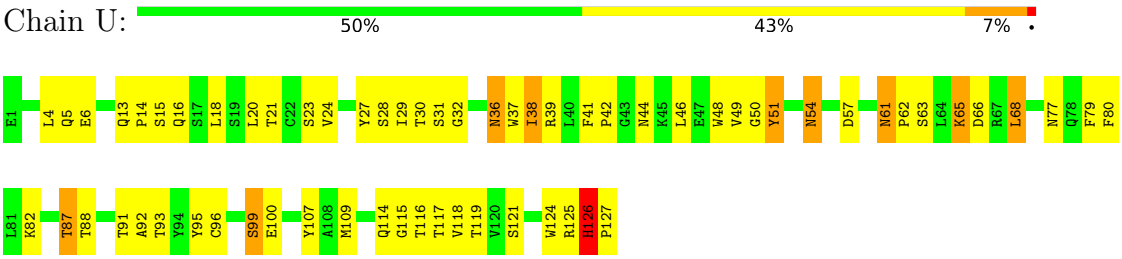
• Molecule 10: HEAVY CHAIN (VH) OF FV-FRAGMENT

Chain J:



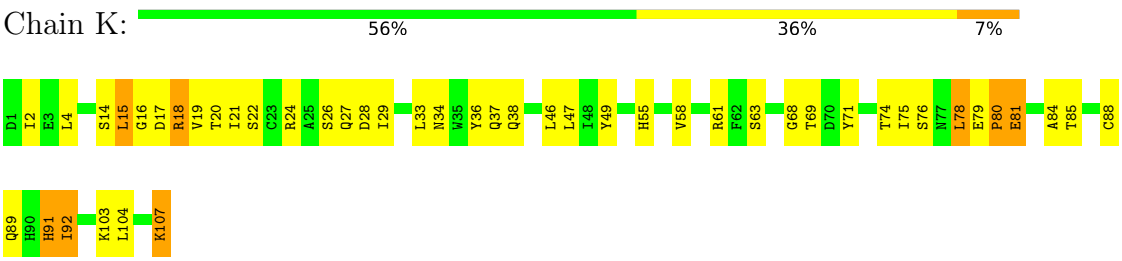
• Molecule 10: HEAVY CHAIN (VH) OF FV-FRAGMENT

Chain U:



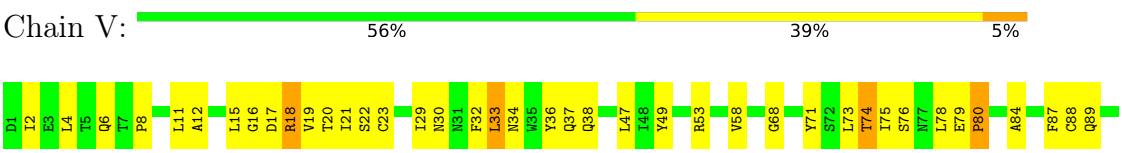
• Molecule 11: LIGHT CHAIN (VL) OF FV-FRAGMENT

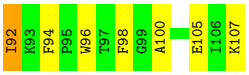
Chain K:



• Molecule 11: LIGHT CHAIN (VL) OF FV-FRAGMENT

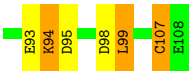
Chain V:





• Molecule 12: CYTOCHROME C, ISO-1

Chain W:



4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	147.22Å 165.53Å 195.89Å 90.00° 104.19° 90.00°	Depositor
Resolution (Å)	29.64 – 2.97	Depositor
% Data completeness (in resolution range)	93.7 (29.64-2.97)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	0.11	Depositor
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.229 , 0.268	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	35643	wwPDB-VP
Average B, all atoms (Å^2)	59.0	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: SMA, FES, M3L, HEC

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	0.40	0/3399	0.63	0/4606
1	L	0.40	0/3399	0.63	1/4606 (0.0%)
2	B	0.37	0/2781	0.68	2/3764 (0.1%)
2	M	0.38	0/2781	0.68	1/3764 (0.0%)
3	C	0.50	0/3191	0.68	0/4353
3	N	0.49	0/3191	0.70	0/4353
4	D	0.41	0/1989	0.64	0/2710
4	O	0.41	0/1989	0.63	0/2710
5	E	0.41	0/1444	0.67	1/1957 (0.1%)
5	P	0.39	0/1444	0.67	0/1957
6	F	0.39	0/639	0.59	0/859
6	Q	0.39	0/639	0.61	0/859
7	G	0.42	0/1032	0.69	0/1397
7	R	0.43	0/1032	0.68	0/1397
8	H	0.49	0/804	0.65	0/1088
8	S	0.52	0/804	0.61	0/1088
9	I	0.43	0/449	0.55	0/605
9	T	0.44	0/449	0.56	0/605
10	J	0.38	0/1043	0.68	1/1422 (0.1%)
10	U	0.38	0/1043	0.68	1/1422 (0.1%)
11	K	0.35	0/863	0.58	0/1172
11	V	0.36	0/863	0.62	0/1172
12	W	0.52	0/856	0.68	0/1145
All	All	0.42	0/36124	0.66	7/49011 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
3	N	0	1

There are no bond length outliers.

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
10	U	32	GLY	N-CA-C	8.27	133.77	113.10
10	J	32	GLY	N-CA-C	6.98	130.55	113.10
2	B	340	PHE	N-CA-C	-6.37	93.80	111.00
2	M	340	PHE	N-CA-C	-6.34	93.87	111.00
5	E	163	GLY	N-CA-C	5.42	126.66	113.10

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	N	224	TYR	Sidechain

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	3338	0	3316	87	0
1	L	3338	0	3316	101	0
2	B	2735	0	2774	98	0
2	M	2735	0	2774	107	0
3	C	3089	0	3125	88	0
3	N	3089	0	3125	85	0
4	D	1929	0	1844	26	0
4	O	1929	0	1844	43	0
5	E	1411	0	1386	49	0
5	P	1411	0	1386	38	0
6	F	625	0	576	11	0
6	Q	625	0	576	9	0
7	G	1012	0	1026	18	0
7	R	1012	0	1026	21	0
8	H	773	0	736	29	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
8	S	773	0	736	29	0
9	I	436	0	435	10	0
9	T	436	0	435	12	0
10	J	1015	0	959	48	0
10	U	1015	0	959	47	0
11	K	842	0	820	40	0
11	V	842	0	820	41	0
12	W	850	0	854	28	0
13	C	86	0	64	7	0
13	D	43	0	30	1	0
13	N	86	0	64	8	0
13	O	43	0	30	4	0
13	W	43	0	30	4	0
14	C	37	0	40	4	0
14	N	37	0	40	4	0
15	E	4	0	0	1	0
15	P	4	0	0	0	0
All	All	35643	0	35146	1000	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 14.

The worst 5 of 1000 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)
2:B:49:HIS:HD2	2:B:161:TYR:H	1.08	0.97
2:M:49:HIS:HD2	2:M:161:TYR:H	1.13	0.96
7:R:31:GLN:HE21	7:R:31:GLN:HA	1.31	0.96
1:L:62:ASN:H	1:L:65:ASN:HD21	1.13	0.95
1:A:62:ASN:HB2	1:A:65:ASN:HD21	1.31	0.95

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	428/430 (100%)	374 (87%)	47 (11%)	7 (2%)	9 38
1	L	428/430 (100%)	384 (90%)	37 (9%)	7 (2%)	9 38
2	B	350/352 (99%)	286 (82%)	47 (13%)	17 (5%)	2 11
2	M	350/352 (99%)	283 (81%)	48 (14%)	19 (5%)	2 9
3	C	383/385 (100%)	357 (93%)	21 (6%)	5 (1%)	12 43
3	N	383/385 (100%)	354 (92%)	24 (6%)	5 (1%)	12 43
4	D	243/248 (98%)	222 (91%)	20 (8%)	1 (0%)	34 70
4	O	243/248 (98%)	225 (93%)	14 (6%)	4 (2%)	9 38
5	E	183/185 (99%)	161 (88%)	19 (10%)	3 (2%)	9 38
5	P	183/185 (99%)	163 (89%)	20 (11%)	0	100 100
6	F	72/74 (97%)	65 (90%)	7 (10%)	0	100 100
6	Q	72/74 (97%)	62 (86%)	10 (14%)	0	100 100
7	G	123/126 (98%)	113 (92%)	8 (6%)	2 (2%)	9 38
7	R	123/126 (98%)	120 (98%)	2 (2%)	1 (1%)	19 55
8	H	91/93 (98%)	79 (87%)	8 (9%)	4 (4%)	2 13
8	S	91/93 (98%)	73 (80%)	13 (14%)	5 (6%)	2 9
9	I	51/57 (90%)	45 (88%)	6 (12%)	0	100 100
9	T	51/57 (90%)	45 (88%)	5 (10%)	1 (2%)	7 32
10	J	125/127 (98%)	110 (88%)	11 (9%)	4 (3%)	4 20
10	U	125/127 (98%)	107 (86%)	15 (12%)	3 (2%)	6 27
11	K	105/107 (98%)	85 (81%)	14 (13%)	6 (6%)	1 9
11	V	105/107 (98%)	83 (79%)	17 (16%)	5 (5%)	2 12
12	W	105/108 (97%)	89 (85%)	14 (13%)	2 (2%)	8 33
All	All	4413/4476 (99%)	3885 (88%)	427 (10%)	101 (2%)	6 28

5 of 101 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	226	ASN
2	B	95	LYS
2	B	152	ARG
2	B	335	PRO

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	336	ILE

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	369/369 (100%)	333 (90%)	36 (10%)	8 29
1	L	369/369 (100%)	328 (89%)	41 (11%)	6 23
2	B	301/301 (100%)	272 (90%)	29 (10%)	8 30
2	M	301/301 (100%)	269 (89%)	32 (11%)	6 25
3	C	338/338 (100%)	311 (92%)	27 (8%)	12 38
3	N	338/338 (100%)	311 (92%)	27 (8%)	12 38
4	D	202/206 (98%)	190 (94%)	12 (6%)	19 52
4	O	202/206 (98%)	193 (96%)	9 (4%)	27 62
5	E	151/151 (100%)	141 (93%)	10 (7%)	16 47
5	P	151/151 (100%)	145 (96%)	6 (4%)	31 66
6	F	67/67 (100%)	60 (90%)	7 (10%)	7 26
6	Q	67/67 (100%)	58 (87%)	9 (13%)	4 16
7	G	109/110 (99%)	104 (95%)	5 (5%)	27 61
7	R	109/110 (99%)	101 (93%)	8 (7%)	14 42
8	H	77/77 (100%)	70 (91%)	7 (9%)	9 32
8	S	77/77 (100%)	69 (90%)	8 (10%)	7 26
9	I	45/47 (96%)	41 (91%)	4 (9%)	9 33
9	T	45/47 (96%)	44 (98%)	1 (2%)	52 80
10	J	112/112 (100%)	99 (88%)	13 (12%)	5 21
10	U	112/112 (100%)	96 (86%)	16 (14%)	3 14
11	K	93/93 (100%)	87 (94%)	6 (6%)	17 48
11	V	93/93 (100%)	87 (94%)	6 (6%)	17 48
12	W	88/88 (100%)	76 (86%)	12 (14%)	3 15

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	3816/3830 (100%)	3485 (91%)	331 (9%)	10 35

5 of 331 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	N	17	ILE
7	R	83	LEU
3	N	89	PHE
4	O	113	ARG
10	U	39	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 139 such sidechains are listed below:

Mol	Chain	Res	Type
5	P	106	ASN
7	R	34	ASN
10	U	78	GLN
8	H	15	HIS
7	G	86	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)

1 non-standard protein/DNA/RNA residue is 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	$\#\|Z\| > 2$
12	M3L	W	77	12	10,11,12	0.61	0	9,14,16	1.23	1 (11%)

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
12	M3L	W	77	12	-	1/9/10/12	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
12	W	77	M3L	CM3-NZ-CM2	-2.86	101.61	108.97

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
12	W	77	M3L	CA-CB-CG-CD

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
12	W	77	M3L	1	0

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

11 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	$\#\|Z\|>2$
13	HEC	C	502	3	32,50,50	1.59	2 (6%)	24,82,82	1.49	5 (20%)
13	HEC	N	522	3	32,50,50	1.43	3 (9%)	24,82,82	1.84	5 (20%)
15	FES	P	524	5	0,4,4	-	-	-		
14	SMA	C	505	-	38,38,38	1.54	6 (15%)	48,52,52	2.18	11 (22%)
15	FES	E	504	5	0,4,4	-	-	-		
14	SMA	N	525	-	38,38,38	1.57	7 (18%)	48,52,52	2.25	11 (22%)
13	HEC	C	501	3	32,50,50	1.73	3 (9%)	24,82,82	1.34	3 (12%)
13	HEC	D	503	4	32,50,50	1.62	4 (12%)	24,82,82	1.19	2 (8%)
13	HEC	O	523	4	32,50,50	1.87	4 (12%)	24,82,82	1.09	2 (8%)
13	HEC	W	526	12	32,50,50	1.55	3 (9%)	24,82,82	1.51	6 (25%)
13	HEC	N	521	3	32,50,50	1.76	4 (12%)	24,82,82	1.68	4 (16%)

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
13	HEC	C	502	3	-	4/10/54/54	-
13	HEC	N	522	3	-	4/10/54/54	-
15	FES	P	524	5	-	-	0/1/1/1
14	SMA	C	505	-	2/2/5/10	22/34/34/34	0/2/2/2
15	FES	E	504	5	-	-	0/1/1/1
14	SMA	N	525	-	2/2/5/10	18/34/34/34	0/2/2/2
13	HEC	C	501	3	-	5/10/54/54	-
13	HEC	D	503	4	-	5/10/54/54	-
13	HEC	O	523	4	-	2/10/54/54	-
13	HEC	W	526	12	-	2/10/54/54	-
13	HEC	N	521	3	-	5/10/54/54	-

The worst 5 of 36 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
13	O	523	HEC	C2B-C3B	-6.05	1.34	1.40
13	C	501	HEC	C2B-C3B	-6.05	1.34	1.40
13	O	523	HEC	C3C-C2C	-5.92	1.34	1.40
13	N	521	HEC	C3C-C2C	-5.76	1.34	1.40
13	C	502	HEC	C3C-C2C	-5.50	1.35	1.40

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
14	N	525	SMA	C26-C19-C18	-6.63	107.62	118.08
14	C	505	SMA	C26-C19-C18	-6.39	108.02	118.08
14	N	525	SMA	O14-C14-C15	5.59	129.06	111.05
14	C	505	SMA	O14-C14-C13	5.49	119.46	107.98
14	C	505	SMA	O14-C14-C15	5.38	128.38	111.05

All (4) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
14	C	505	SMA	C12
14	C	505	SMA	C14
14	N	525	SMA	C12
14	N	525	SMA	C14

5 of 67 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
13	N	521	HEC	C2A-CAA-CBA-CGA
14	C	505	SMA	C11-C12-C13-C24
14	C	505	SMA	O12-C12-C13-C24
14	C	505	SMA	C13-C14-O14-C25
14	C	505	SMA	C14-C15-C16-C17

There are no ring outliers.

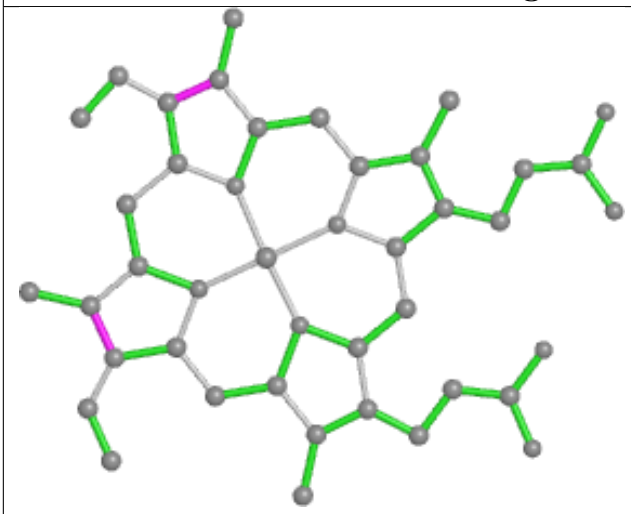
9 monomers are involved in 33 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
13	N	522	HEC	4	0
14	C	505	SMA	4	0
15	E	504	FES	1	0
14	N	525	SMA	4	0
13	C	501	HEC	7	0
13	D	503	HEC	1	0
13	O	523	HEC	4	0
13	W	526	HEC	4	0
13	N	521	HEC	4	0

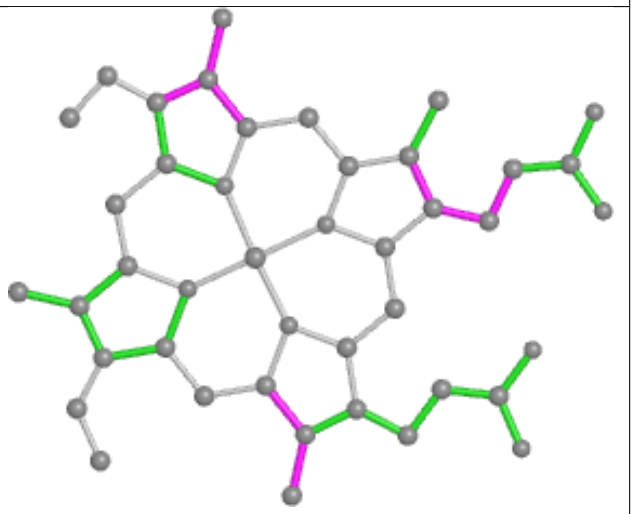
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. 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.

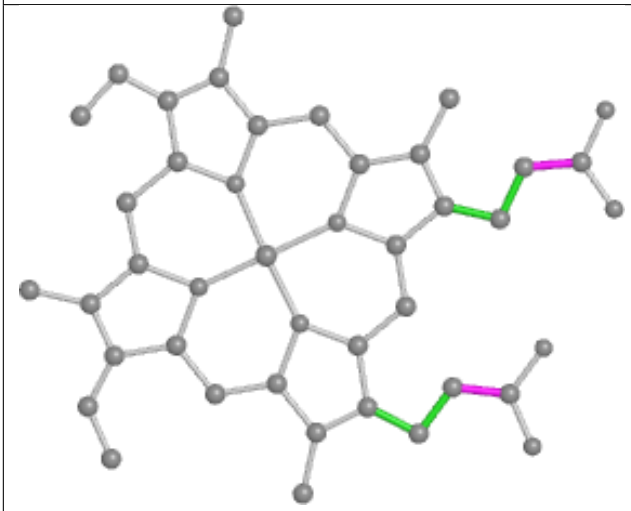
Ligand HEC C 502



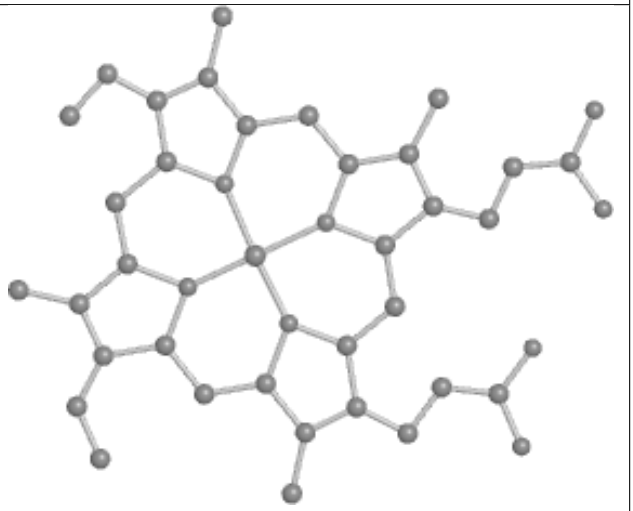
Bond lengths



Bond angles

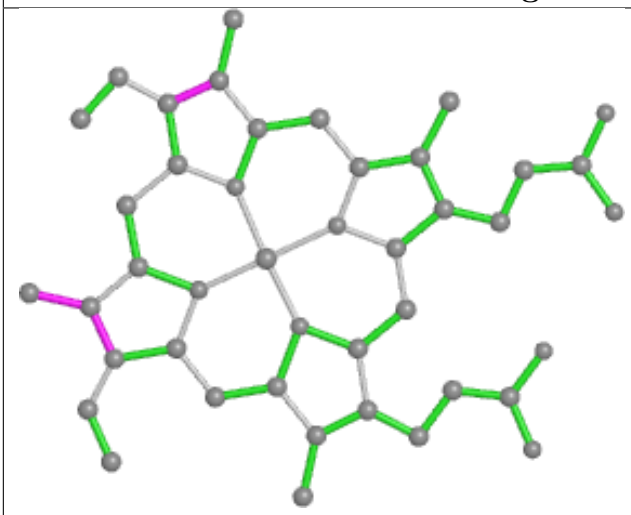


Torsions

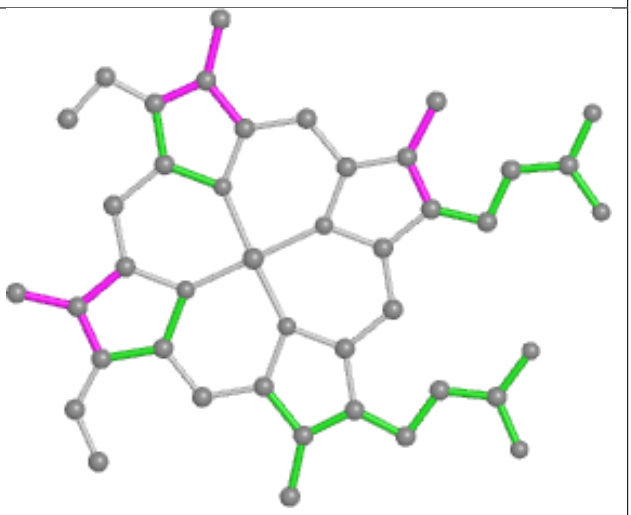


Rings

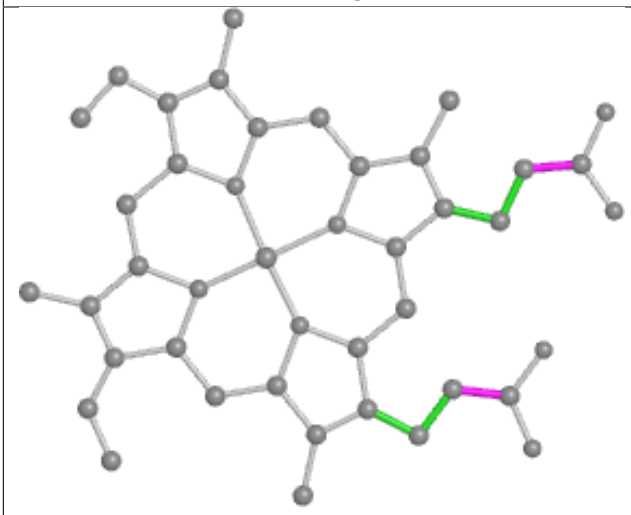
Ligand HEC N 522



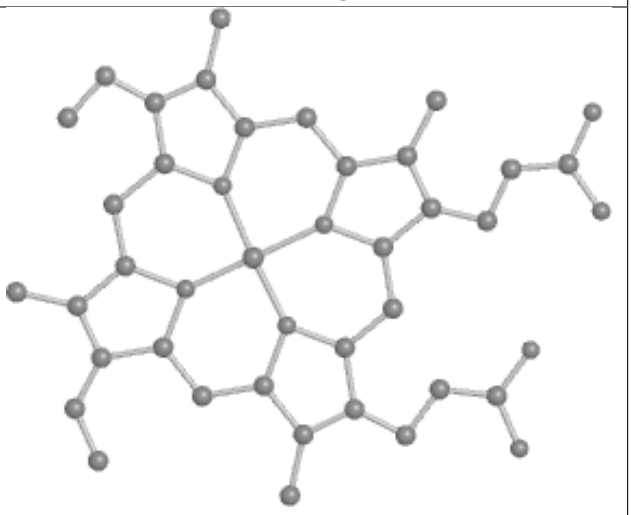
Bond lengths



Bond angles

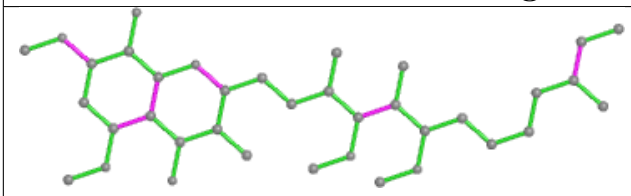


Torsions

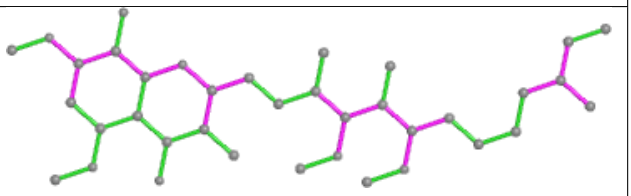


Rings

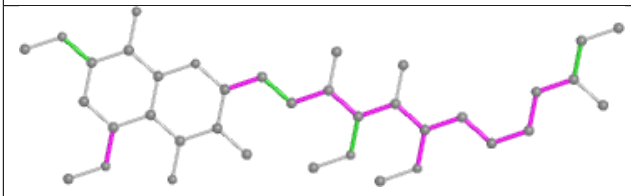
Ligand SMA C 505



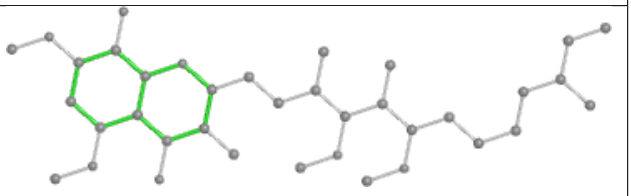
Bond lengths



Bond angles

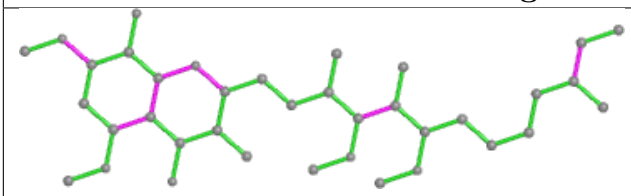


Torsions

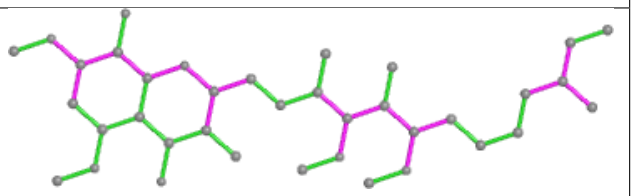


Rings

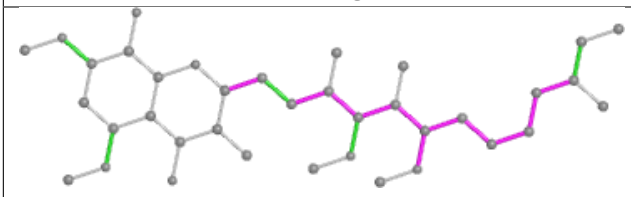
Ligand SMA N 525



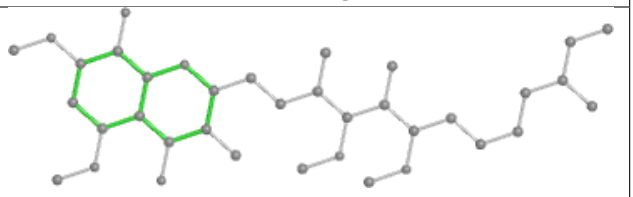
Bond lengths



Bond angles

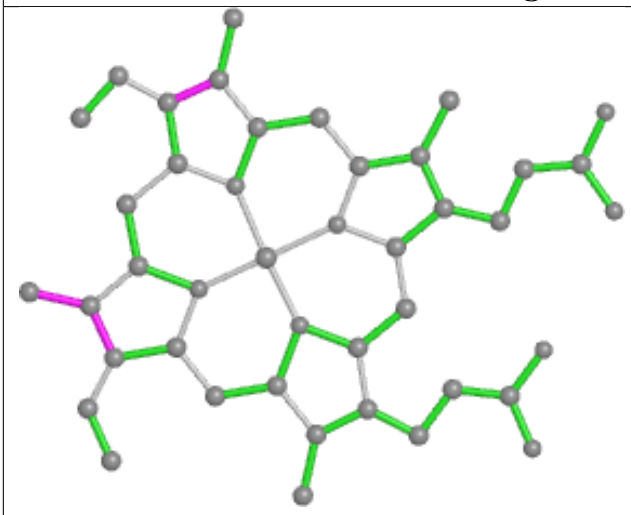


Torsions

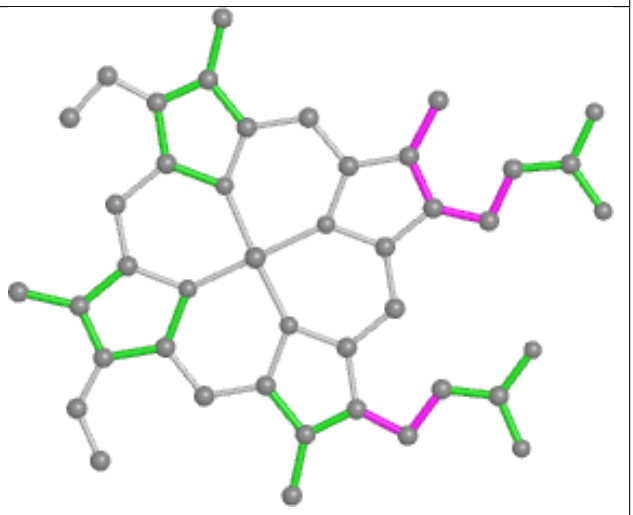


Rings

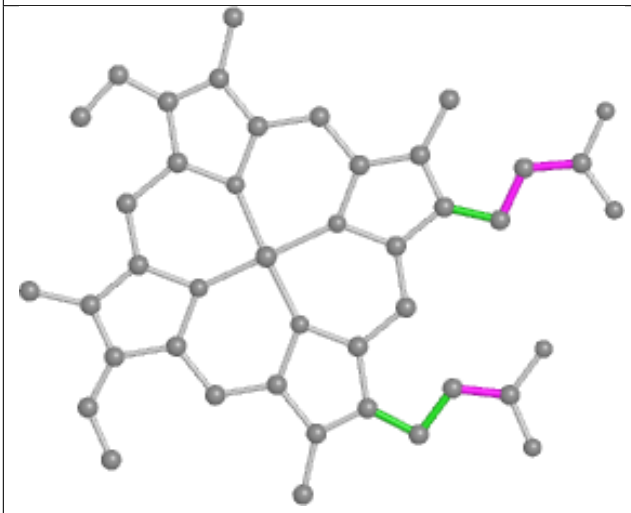
Ligand HEC C 501



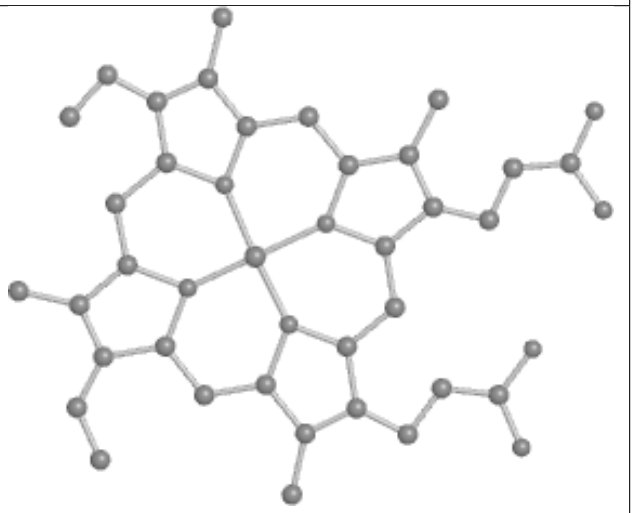
Bond lengths



Bond angles

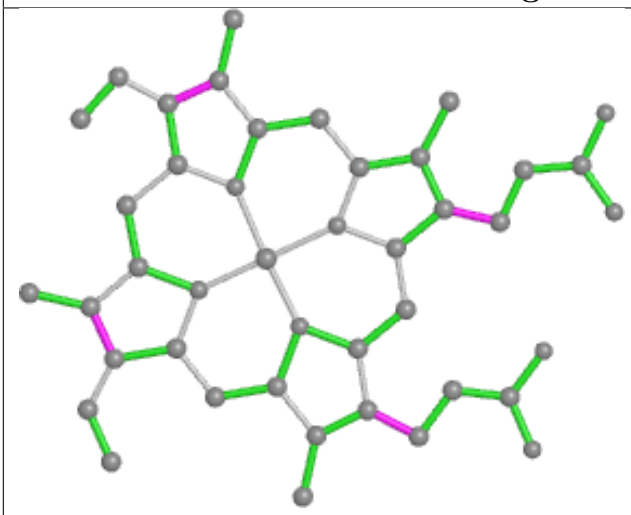


Torsions

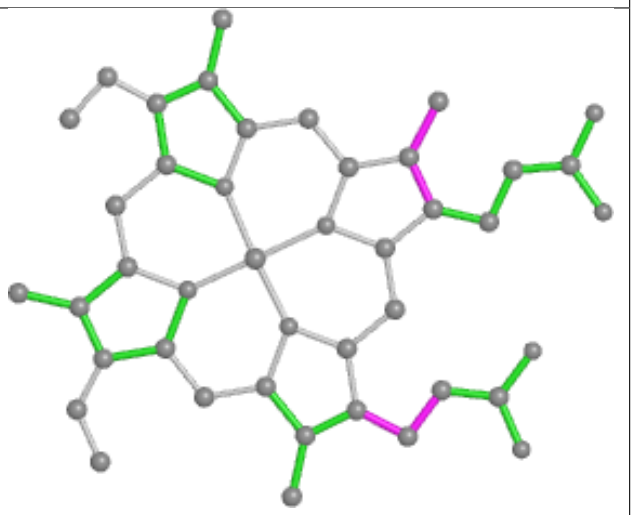


Rings

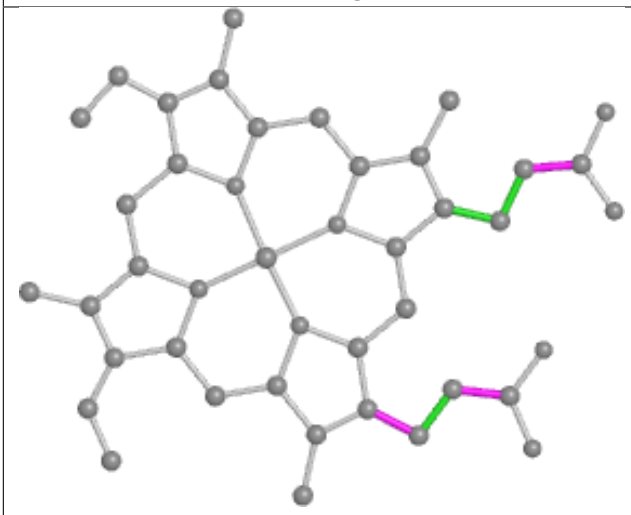
Ligand HEC D 503



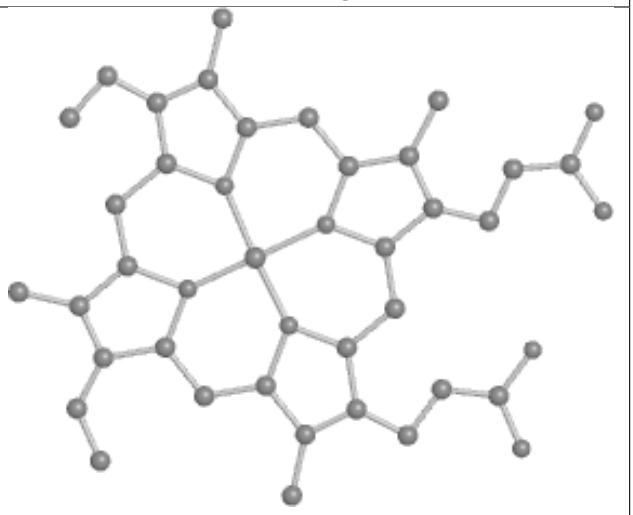
Bond lengths



Bond angles

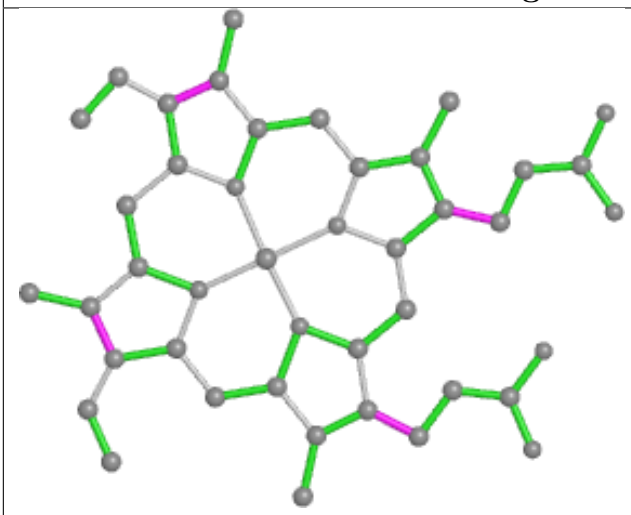


Torsions

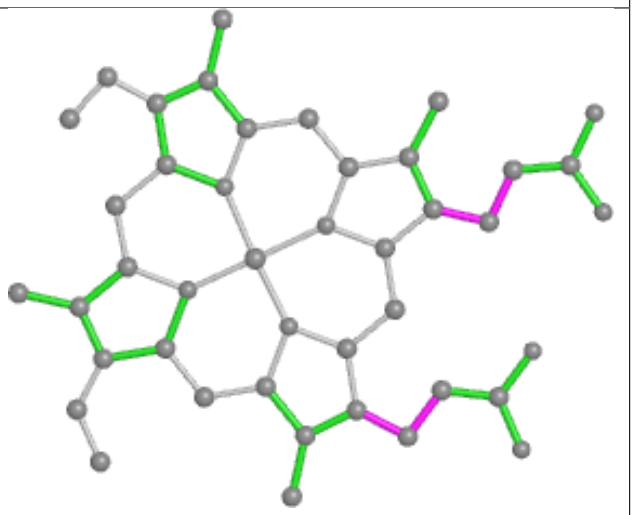


Rings

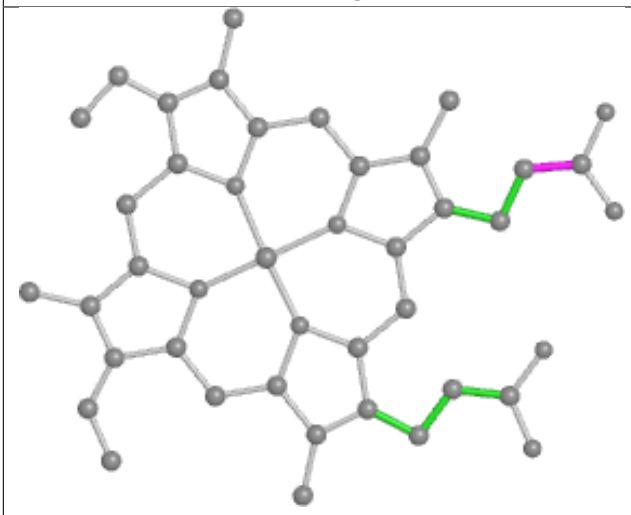
Ligand HEC O 523



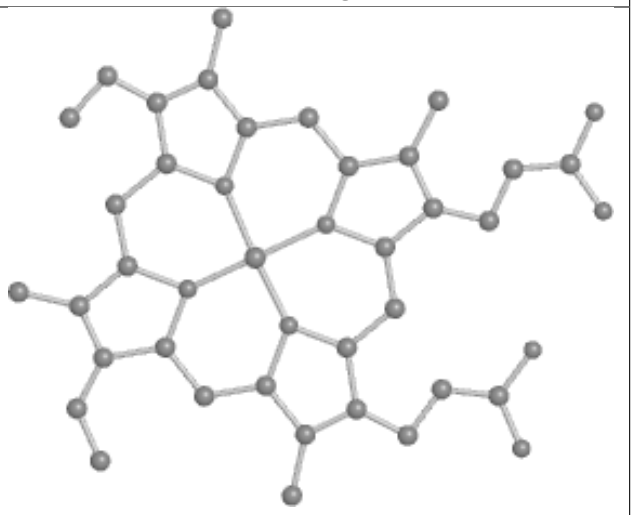
Bond lengths



Bond angles

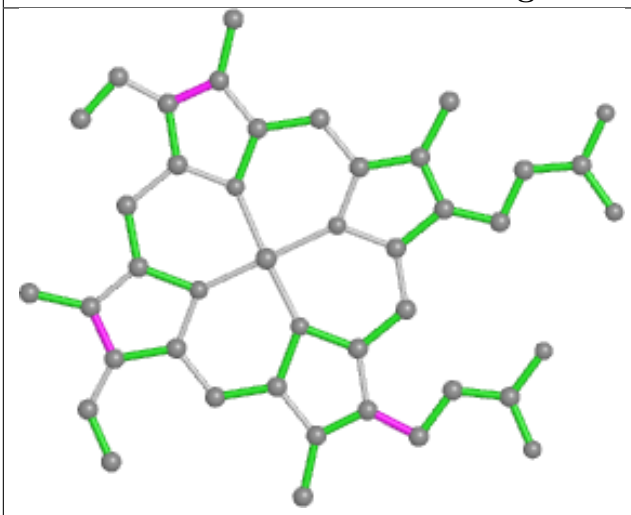


Torsions

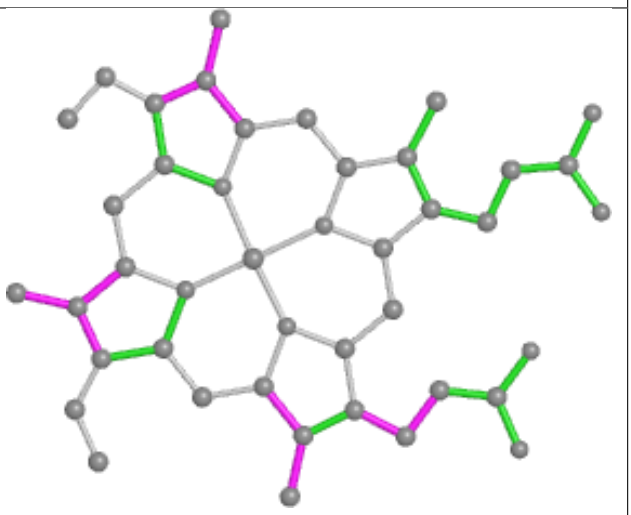


Rings

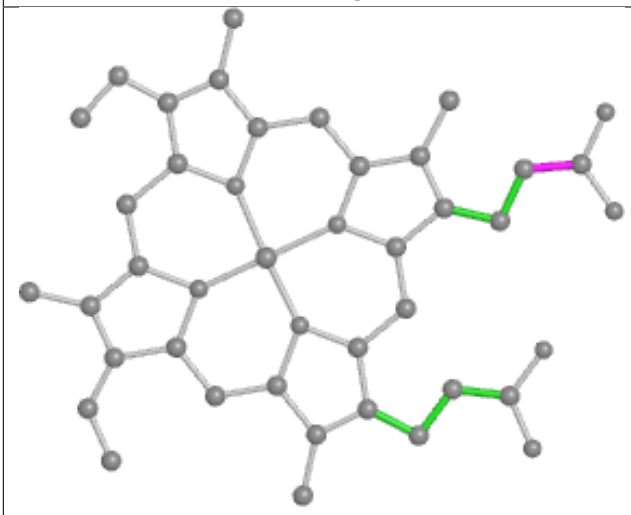
Ligand HEC W 526



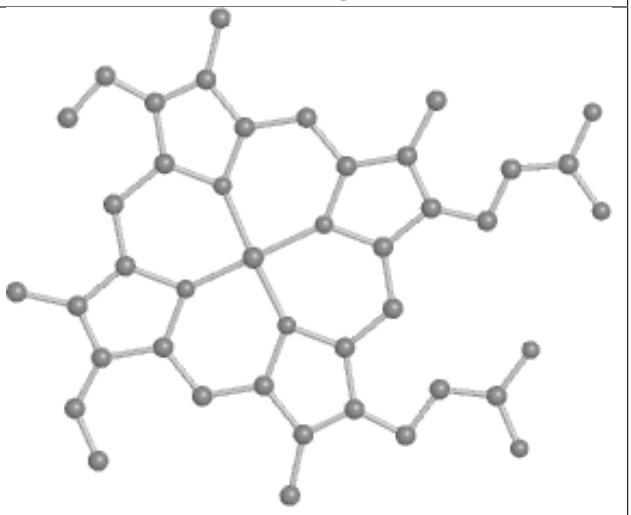
Bond lengths



Bond angles

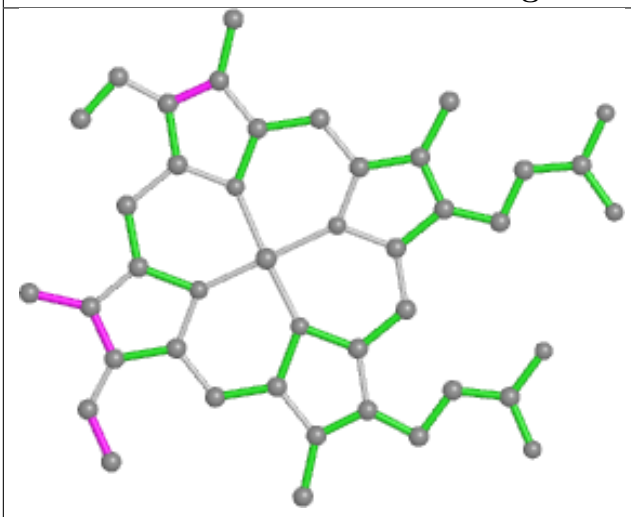


Torsions

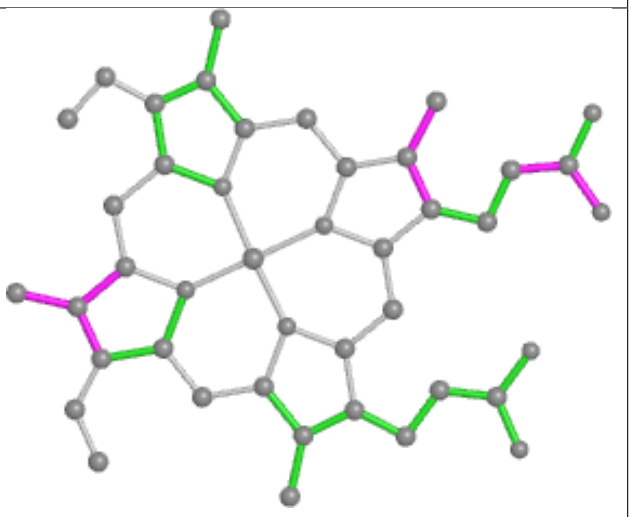


Rings

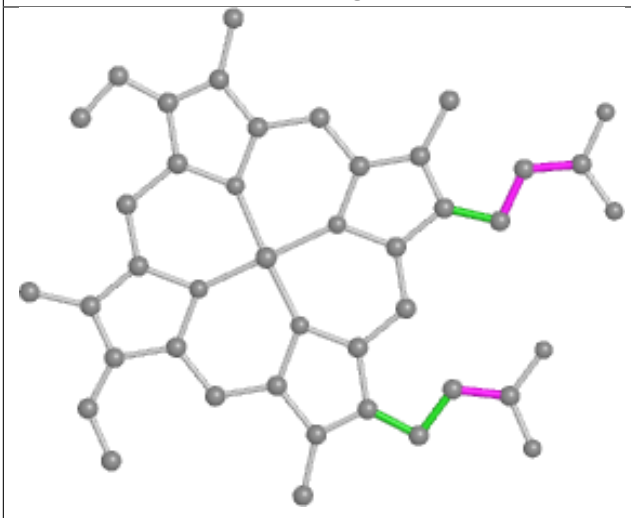
Ligand HEC N 521



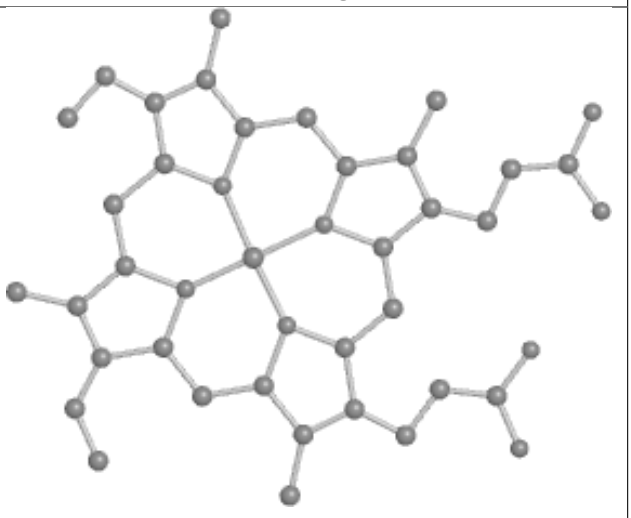
Bond lengths



Bond angles



Torsions



Rings

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.