Full wwPDB X-ray Structure Validation Report (i)

Aug 5, 2026 – 11:08 AM EDT

PDB ID	:	3OCD / pdb_00003ocd
Title	:	Diheme SoxAX - C236M mutant
Authors	:	Maher, M.J.
Deposited on	:	2010-08-09
Resolution	:	2.25 Å(reported)

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

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

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

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

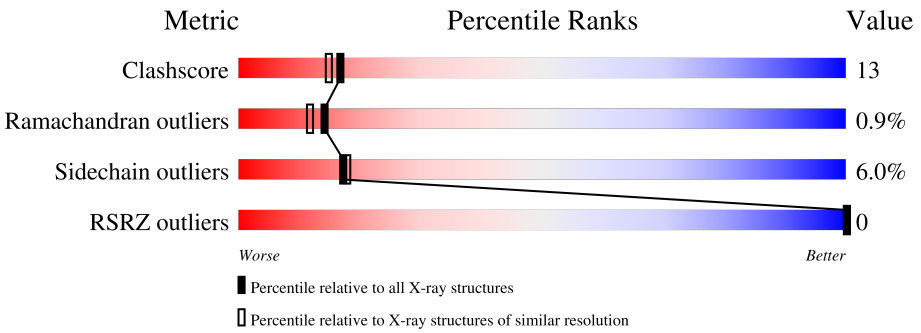
MolProbity	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Buster-report	:	wwPDB partial adaption of 1.1.7 (2018)
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.015 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.50

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))
Clashscore	190562	2005 (2.26-2.26)
Ramachandran outliers	187476	1965 (2.26-2.26)
Sidechain outliers	187428	1966 (2.26-2.26)
RSRZ outliers	180081	1898 (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	275	57% 25% • 17%
1	C	275	55% 27% • 17%
2	B	208	71% 13% • 13%
2	D	208	66% 20% 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
3	HEC	A	401	X	-	-	-
3	HEC	B	401	X	-	-	-
3	HEC	C	401	X	-	-	-
3	HEC	D	401	X	-	-	-

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 6543 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 SoxA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	229	1789	1126	316	333	14	0	1	0
1	C	229	1789	1126	316	333	14	0	1	0

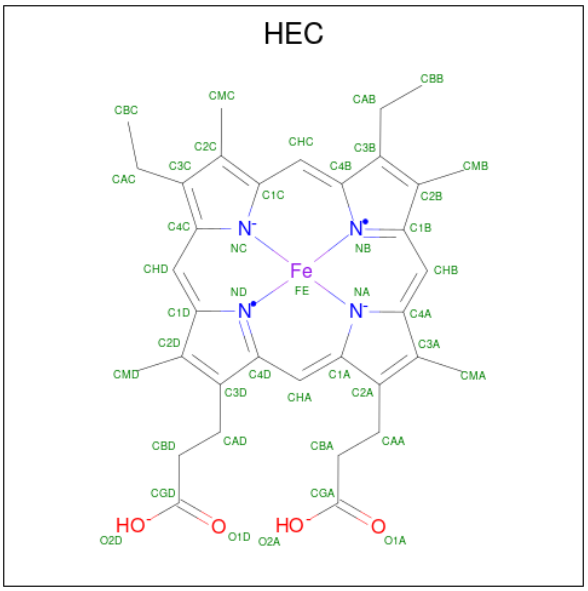
There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	236	MET	CYS	engineered mutation	UNP Q7BQR6
C	236	MET	CYS	engineered mutation	UNP Q7BQR6

- Molecule 2 is a protein called SoxX.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	180	1373	861	239	267	6	4	4	0
2	D	180	1377	865	239	267	6	11	5	0

- Molecule 3 is HEME C (CCD ID: HEC) (formula: $C_{34}H_{36}FeN_4O_4$).



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

- Molecule 4 is water.

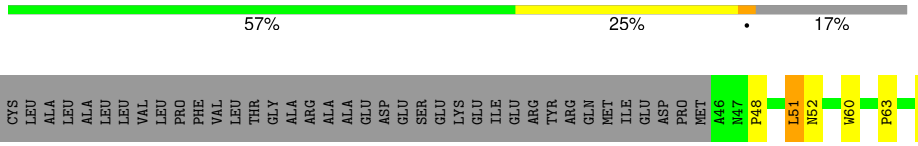
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	13	13	13	0	0
			Total	O		
4	B	9	9	9	0	0
			Total	O		
4	C	11	11	11	0	0
			Total	O		
4	D	10	10	10	0	0

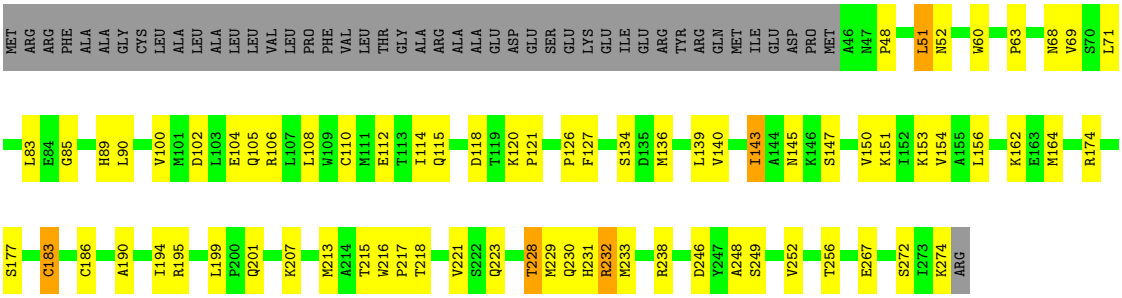
3 Residue-property plots (i)

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: SoxA

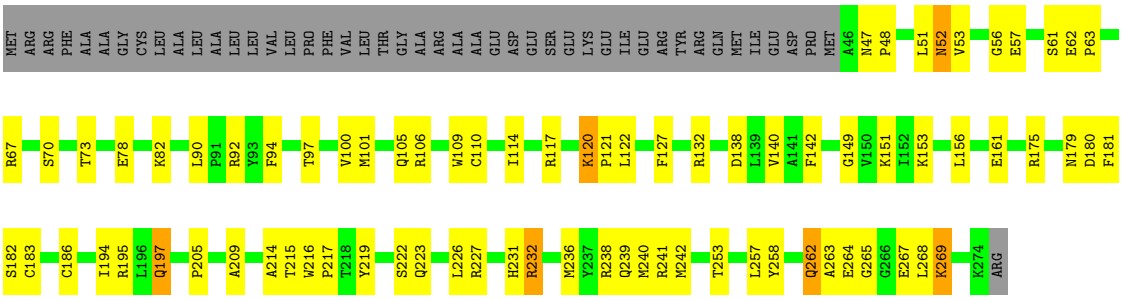
Chain A:





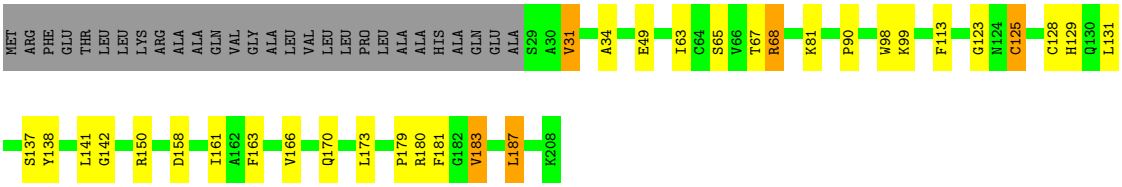
- Molecule 1: SoxA

Chain C:



- Molecule 2: SoxX

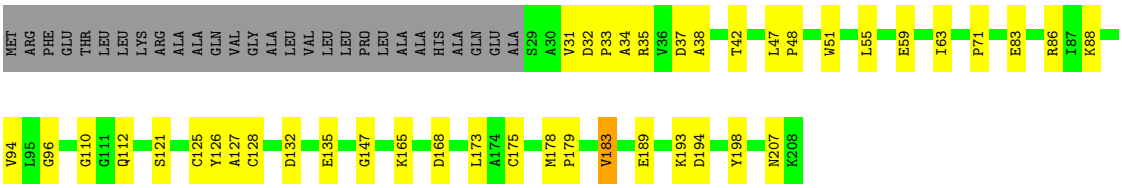
Chain B:



- Molecule 2: SoxX

Chain D:

66%
20%
13%


MET ARG PHE GLU THR LEU LEU LYS ARG ALA ALA GLN VAL GLY ALA LEU VAL LEU LEU PRO LEU ALA ALA HIS ALA GLN GLU ALA S29 A30 V31 D32 P33 A34 R35 Y36 D37 A38 T42 L47 P48 W51 L55 E59 I63 P71 E83 R86 I87 K88
V94 L95 G96 G110 G111 Q112 S121 C125 Y126 A127 C128 D132 E135 G147 K165 D168 L173 A174 C175 M178 P179 V183 E189 K193 D194 Y198 N207 K208

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	47.50Å 66.03Å 85.89Å 112.59° 94.33° 90.02°	Depositor
Resolution (Å)	41.18 – 2.25 41.18 – 2.25	Depositor EDS
% Data completeness (in resolution range)	90.2 (41.18-2.25) 90.3 (41.18-2.25)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.98 (at 2.24Å)	Xtriage
Refinement program	REFMAC 5.5.0088	Depositor
R, R_{free}	0.221 , 0.257 0.222 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å^2)	33.9	Xtriage
Anisotropy	0.651	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 14.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.479 for -h,k,-k-l	Xtriage
Reported twinning fraction	0.508 for H, K, L 0.492 for -H, K, -K-L	Depositor
Outliers	0 of 41131 reflections	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	6543	wwPDB-VP
Average B, all atoms (Å^2)	32.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 12.18% of the height of the origin peak. No significant pseudotranslation is detected.*

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: 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	0.53	0/1830	0.86	1/2473 (0.0%)
1	C	0.53	0/1830	0.83	0/2473
2	B	0.55	0/1414	0.90	2/1919 (0.1%)
2	D	0.53	0/1421	0.84	0/1929
All	All	0.53	0/6495	0.86	3/8794 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	142	GLY	CA-C-N	8.63	128.67	119.78
2	B	142	GLY	C-N-CA	8.63	128.67	119.78
1	A	223	GLN	N-CA-C	-5.56	107.50	114.56

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	1789	0	1789	53	0
1	C	1789	0	1789	62	0
2	B	1373	0	1355	24	0
2	D	1377	0	1364	30	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	43	0	32	8	0
3	B	43	0	32	9	0
3	C	43	0	32	12	0
3	D	43	0	32	13	0
4	A	13	0	0	4	0
4	B	9	0	0	1	0
4	C	11	0	0	1	0
4	D	10	0	0	0	0
All	All	6543	0	6425	165	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 13.

All (165) 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:183:CYS:SG	3:A:401:HEC:HAB	1.50	1.50
1:A:183:CYS:SG	3:A:401:HEC:CAB	2.26	1.22
2:B:125:CYS:SG	3:B:401:HEC:HAB	1.97	1.04
2:D:125:CYS:SG	3:D:401:HEC:CAB	2.48	1.01
1:C:186:CYS:SG	3:C:401:HEC:HAC	2.00	1.00
1:C:183:CYS:SG	3:C:401:HEC:HAB	2.03	0.98
2:D:125:CYS:SG	3:D:401:HEC:HAB	2.04	0.97
1:C:101:MET:HE1	1:C:106:ARG:HH11	1.29	0.95
1:C:183:CYS:SG	3:C:401:HEC:CAB	2.55	0.94
1:C:186:CYS:SG	3:C:401:HEC:CAC	2.55	0.94
2:B:128:CYS:SG	3:B:401:HEC:CAC	2.62	0.87
2:D:128:CYS:SG	3:D:401:HEC:HAC	2.14	0.87
2:B:125:CYS:HB2	2:B:131:LEU:HD11	1.57	0.86
2:B:125:CYS:SG	3:B:401:HEC:CAB	2.64	0.85
1:C:262:GLN:HE21	1:C:262:GLN:HA	1.41	0.85
1:A:194:ILE:HD12	1:A:199:LEU:HD21	1.60	0.83
1:A:156:LEU:HB3	1:A:162:LYS:HG2	1.62	0.81
2:B:34:ALA:HA	4:B:209:HOH:O	1.81	0.79
1:A:108:LEU:O	1:A:112:GLU:HG2	1.84	0.76
1:A:71:LEU:HD22	1:A:110:CYS:HB3	1.68	0.75
1:C:242:MET:HE1	2:D:55:LEU:HD23	1.71	0.72
2:D:128:CYS:SG	3:D:401:HEC:CAC	2.80	0.70
2:B:179:PRO:HD3	3:B:401:HEC:HBC2	1.75	0.68
1:A:233:MET:HG3	1:A:252:VAL:HG13	1.74	0.68
1:A:186:CYS:SG	3:A:401:HEC:HAC	2.35	0.67

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:246:ASP:HB2	1:A:249:SER:HB3	1.77	0.66
1:A:194:ILE:CD1	3:A:401:HEC:HHD	2.27	0.65
2:D:96:GLY:HA3	2:D:194:ASP:OD1	1.98	0.64
2:B:98:TRP:CZ2	2:B:99:LYS:HE3	2.34	0.63
2:B:181:PHE:HB3	2:B:187:LEU:HG	1.81	0.62
2:D:147:GLY:HA2	2:D:207:ASN:HD22	1.65	0.61
1:C:195:ARG:NH2	4:C:282:HOH:O	2.33	0.60
2:B:170:GLN:HA	2:B:173:LEU:O	2.01	0.60
1:C:92:ARG:HA	1:C:161:GLU:OE2	2.01	0.60
2:D:179:PRO:HD3	3:D:401:HEC:CBC	2.33	0.59
1:A:230:GLN:NE2	1:A:248:ALA:H	2.02	0.58
2:D:125:CYS:SG	3:D:401:HEC:C3B	2.92	0.58
1:A:85:GLY:O	1:A:89:HIS:ND1	2.36	0.57
2:D:59:GLU:O	2:D:63:ILE:HG13	2.05	0.57
1:A:102:ASP:H	1:A:105:GLN:HE21	1.51	0.57
3:D:401:HEC:HBB3	3:D:401:HEC:HMB3	1.86	0.57
1:C:262:GLN:HE21	1:C:262:GLN:CA	2.15	0.57
1:C:52:ASN:HD22	1:C:52:ASN:H	1.51	0.56
1:C:149:GLY:O	1:C:267:GLU:HG3	2.05	0.56
1:A:71:LEU:CD2	1:A:110:CYS:HB3	2.34	0.56
1:A:127:PHE:CG	1:A:231:HIS:HD2	2.24	0.56
1:C:101:MET:HE1	1:C:106:ARG:NH1	2.11	0.55
1:A:140:VAL:HG12	1:A:217:PRO:HG3	1.87	0.55
2:D:38:ALA:O	2:D:42:THR:HG23	2.07	0.55
2:D:110:GLY:HA3	2:D:121:SER:HA	1.89	0.55
1:A:186:CYS:SG	3:A:401:HEC:CAC	2.95	0.54
1:C:219:TYR:HA	1:C:226:LEU:HA	1.88	0.54
1:A:190:ALA:HB2	1:A:201:GLN:OE1	2.08	0.53
1:A:174:ARG:HA	2:B:65:SER:HB3	1.90	0.53
2:D:86:ARG:NH2	2:D:183:VAL:HG13	2.24	0.53
2:B:158:ASP:HA	2:B:161:ILE:HD12	1.91	0.52
1:C:238:ARG:O	2:D:112:GLN:HG2	2.09	0.52
1:A:232:ARG:HA	1:A:232:ARG:HE	1.73	0.52
1:C:110:CYS:O	1:C:114:ILE:HB	2.09	0.52
1:A:151:LYS:HG3	1:A:267:GLU:HG3	1.91	0.52
2:D:33:PRO:C	2:D:35:ARG:H	2.18	0.52
1:C:258:TYR:O	1:C:262:GLN:HG2	2.09	0.52
2:D:189:GLU:HG2	2:D:193:LYS:HD2	1.92	0.51
2:B:180:ARG:HB3	2:B:183:VAL:HG12	1.92	0.51
1:A:52:ASN:ND2	1:A:52:ASN:H	2.09	0.51
2:D:71:PRO:HG3	2:D:175:CYS:SG	2.50	0.51

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:48:PRO:O	2:D:51:TRP:HB2	2.11	0.51
1:C:78:GLU:HB2	1:C:82:LYS:HB3	1.92	0.50
1:C:239:GLN:CD	3:C:401:HEC:HBC2	2.35	0.50
1:A:194:ILE:HD12	3:A:401:HEC:HHD	1.94	0.50
1:A:69:VAL:O	1:A:115:GLN:HG2	2.11	0.50
1:C:53:VAL:HG13	1:C:142:PHE:HA	1.93	0.50
3:D:401:HEC:HBD1	3:D:401:HEC:HHA	1.93	0.50
1:C:194:ILE:HA	2:D:127:ALA:O	2.12	0.50
1:A:100:VAL:HB	1:A:164:MET:HE3	1.93	0.49
1:A:139:LEU:O	1:A:143:ILE:HD12	2.12	0.49
1:C:67:ARG:HG3	1:C:114:ILE:O	2.12	0.49
2:D:179:PRO:HD3	3:D:401:HEC:HBC3	1.94	0.49
1:C:127:PHE:O	1:C:132:ARG:NH1	2.44	0.49
1:C:262:GLN:HA	1:C:262:GLN:NE2	2.18	0.49
1:A:105:GLN:HG3	1:A:248:ALA:HB1	1.93	0.49
1:A:145:ASN:HB3	4:A:284:HOH:O	2.11	0.49
1:C:101:MET:HA	1:C:105:GLN:HE21	1.78	0.49
1:C:100:VAL:HG11	1:C:257:LEU:HD22	1.94	0.48
1:A:150:VAL:HG12	1:A:151:LYS:N	2.28	0.48
1:C:183:CYS:HA	3:C:401:HEC:CHC	2.43	0.48
2:B:128:CYS:SG	3:B:401:HEC:HAC	2.52	0.48
1:A:105:GLN:CG	1:A:248:ALA:HB1	2.44	0.48
1:A:228:THR:CG2	4:A:276:HOH:O	2.61	0.48
3:A:401:HEC:HHA	3:A:401:HEC:HBA1	1.96	0.48
1:C:183:CYS:SG	3:C:401:HEC:C3B	3.01	0.48
2:D:179:PRO:HD3	3:D:401:HEC:HBC2	1.94	0.48
1:C:101:MET:HB3	1:C:105:GLN:HB3	1.96	0.48
1:C:61:SER:HA	1:C:70:SER:HB2	1.94	0.48
1:A:215:THR:O	1:A:218:THR:HB	2.13	0.48
1:C:197:GLN:HE21	1:C:197:GLN:HA	1.80	0.47
1:A:89:HIS:HD2	1:A:153:LYS:O	1.98	0.47
1:A:118:ASP:O	1:A:121:PRO:HD2	2.14	0.47
1:A:238:ARG:HA	2:B:113:PHE:CE2	2.50	0.47
1:A:213:MET:HE1	1:A:229:MET:HE3	1.96	0.47
1:C:205:PRO:HG3	1:C:262:GLN:HE22	1.80	0.47
1:A:136:MET:O	1:A:140:VAL:HG23	2.15	0.47
2:D:83:GLU:HB3	2:D:168:ASP:CG	2.40	0.46
1:A:104:GLU:HG2	1:A:136:MET:HE3	1.97	0.46
1:A:194:ILE:HG13	1:A:199:LEU:HD11	1.98	0.46
1:C:214:ALA:HA	1:C:268:LEU:HD23	1.97	0.46
1:A:229:MET:HG2	1:A:256:THR:HG23	1.96	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:194:ILE:HD13	3:A:401:HEC:HHD	1.97	0.46
1:A:228:THR:HG21	4:A:276:HOH:O	2.16	0.46
2:B:67:THR:C	2:B:68:ARG:HG2	2.40	0.46
1:C:175:ARG:NE	1:C:180:ASP:O	2.44	0.46
1:C:236:MET:HB2	3:C:401:HEC:HBB2	1.98	0.46
1:C:179:ASN:HA	2:D:179:PRO:HB3	1.98	0.45
1:C:216:TRP:CZ2	1:C:232:ARG:HD2	2.51	0.45
2:B:123:GLY:HA3	3:B:401:HEC:HBB1	1.99	0.45
1:C:101:MET:HE3	1:C:106:ARG:HD3	1.99	0.45
2:B:166:VAL:HG21	3:B:401:HEC:HMB2	1.98	0.45
1:A:195:ARG:HD2	2:B:137:SER:OG	2.16	0.45
1:C:223:GLN:HE22	1:C:231:HIS:HE1	1.65	0.45
2:B:81:LYS:HE3	2:B:81:LYS:HB2	1.90	0.44
2:B:129:HIS:O	2:B:138:TYR:HB2	2.16	0.44
1:A:126:PRO:O	1:A:134:SER:OG	2.36	0.44
2:D:178:MET:HA	3:D:401:HEC:HBC2	1.98	0.44
1:C:151:LYS:HD2	1:C:265:GLY:O	2.17	0.44
1:C:94:PHE:CD2	1:C:101:MET:HE2	2.53	0.44
1:C:242:MET:HE1	2:D:55:LEU:CD2	2.45	0.44
2:D:126:TYR:OH	2:D:198:TYR:OH	2.18	0.44
1:C:62:GLU:HA	1:C:63:PRO:HD3	1.92	0.44
1:A:216:TRP:HA	1:A:217:PRO:C	2.42	0.44
1:C:47:ASN:HA	1:C:48:PRO:HD3	1.79	0.43
1:C:105:GLN:NE2	1:C:253:THR:OG1	2.51	0.43
1:C:232:ARG:HD3	3:C:401:HEC:C2A	2.48	0.43
1:A:52:ASN:H	1:A:52:ASN:HD22	1.64	0.43
2:D:132:ASP:HB3	2:D:135:GLU:HB2	2.00	0.43
1:A:60:TRP:CZ2	1:A:83:LEU:HD21	2.52	0.43
1:C:232:ARG:HH11	3:C:401:HEC:C1A	2.31	0.43
1:C:209:ALA:O	1:C:263:ALA:HB2	2.18	0.43
2:B:128:CYS:SG	3:B:401:HEC:C3C	3.07	0.43
2:D:88:LYS:HD3	2:D:88:LYS:HA	1.79	0.43
1:C:120:LYS:N	1:C:121:PRO:HD2	2.34	0.43
1:A:106:ARG:HA	1:A:106:ARG:HD2	1.88	0.43
1:C:140:VAL:HG12	1:C:217:PRO:HG3	2.01	0.42
1:C:175:ARG:HG2	1:C:182:SER:HB3	2.01	0.42
1:C:56:GLY:HA3	1:C:138:ASP:O	2.19	0.42
1:C:101:MET:HA	1:C:105:GLN:NE2	2.34	0.42
1:C:181:PHE:HB3	1:C:240:MET:HG2	2.00	0.42
1:A:218:THR:HG21	1:A:274:LYS:HD2	2.01	0.42
3:D:401:HEC:HHA	3:D:401:HEC:CBD	2.50	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:154:VAL:HG22	4:A:280:HOH:O	2.19	0.42
1:C:267:GLU:HG2	1:C:269:LYS:HE2	2.00	0.42
1:C:97:THR:HG22	1:C:109:TRP:CE2	2.55	0.41
2:D:168:ASP:OD2	2:D:183:VAL:HG21	2.20	0.41
2:B:90:PRO:HD3	2:B:163:PHE:CZ	2.54	0.41
1:C:227:ARG:HD2	1:C:231:HIS:CG	2.56	0.41
2:D:173:LEU:HD23	3:D:401:HEC:HBD2	2.03	0.41
1:A:63:PRO:HA	1:A:69:VAL:C	2.46	0.41
2:B:67:THR:O	2:B:68:ARG:HG2	2.21	0.41
1:C:239:GLN:C	1:C:241:ARG:H	2.29	0.41
1:A:48:PRO:O	1:A:51:LEU:HG	2.21	0.40
2:B:166:VAL:CG2	3:B:401:HEC:HMB2	2.51	0.40
1:C:117:ARG:NH2	1:C:122:LEU:HD11	2.37	0.40
1:C:183:CYS:HA	3:C:401:HEC:HHC	2.03	0.40
1:C:205:PRO:HG3	1:C:262:GLN:NE2	2.36	0.40
1:A:102:ASP:H	1:A:105:GLN:NE2	2.19	0.40
1:C:232:ARG:NH1	3:C:401:HEC:CHA	2.84	0.40

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	228/275 (83%)	212 (93%)	13 (6%)	3 (1%)	9 6
1	C	228/275 (83%)	207 (91%)	19 (8%)	2 (1%)	14 12
2	B	182/208 (88%)	169 (93%)	12 (7%)	1 (0%)	24 24
2	D	183/208 (88%)	170 (93%)	12 (7%)	1 (0%)	24 24
All	All	821/966 (85%)	758 (92%)	56 (7%)	7 (1%)	14 12

All (7) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	207	LYS
2	B	31	VAL
1	A	177	SER
2	D	34	ALA
1	C	264	GLU
1	C	222	SER
1	A	114	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	190/226 (84%)	179 (94%)	11 (6%)	18	19
1	C	190/226 (84%)	177 (93%)	13 (7%)	14	14
2	B	149/165 (90%)	140 (94%)	9 (6%)	17	18
2	D	150/165 (91%)	142 (95%)	8 (5%)	20	22
All	All	679/782 (87%)	638 (94%)	41 (6%)	17	18

All (41) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	51	LEU
1	A	68	ASN
1	A	90	LEU
1	A	120	LYS
1	A	143	ILE
1	A	147	SER
1	A	183	CYS
1	A	221	VAL
1	A	228	THR
1	A	232	ARG
1	A	272	SER
2	B	31	VAL
2	B	49	GLU
2	B	63	ILE
2	B	68	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	125	CYS
2	B	141	LEU
2	B	150	ARG
2	B	183	VAL
2	B	187	LEU
1	C	51	LEU
1	C	52	ASN
1	C	57	GLU
1	C	73	THR
1	C	90	LEU
1	C	120	LYS
1	C	153	LYS
1	C	156	LEU
1	C	197	GLN
1	C	215	THR
1	C	232	ARG
1	C	262	GLN
1	C	269	LYS
2	D	31	VAL
2	D	32	ASP
2	D	37	ASP
2	D	47	LEU
2	D	94[A]	VAL
2	D	94[B]	VAL
2	D	165	LYS
2	D	183	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (18) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	52	ASN
1	A	105	GLN
1	A	115	GLN
1	A	230	GLN
1	A	231	HIS
1	A	262	GLN
2	B	70	ASN
1	C	52	ASN
1	C	68	ASN
1	C	105	GLN
1	C	197	GLN
1	C	230	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	231	HIS
1	C	262	GLN
2	D	57	GLN
2	D	75	GLN
2	D	190	GLN
2	D	207	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 oligosaccharides in this entry.

5.6 Ligand geometry (i)

4 ligands are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
3	HEC	A	401	1	50,50,50	1.87	7 (14%)	68,82,82	1.32	7 (10%)
3	HEC	B	401	2	50,50,50	2.06	9 (18%)	68,82,82	1.45	8 (11%)
3	HEC	C	401	1	50,50,50	1.96	9 (18%)	68,82,82	1.32	5 (7%)
3	HEC	D	401	2	50,50,50	1.95	9 (18%)	68,82,82	1.34	5 (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
3	HEC	A	401	1	1/1/3/7	5/14/54/54	-
3	HEC	B	401	2	1/1/3/7	8/14/54/54	-
3	HEC	C	401	1	1/1/3/7	6/14/54/54	-
3	HEC	D	401	2	1/1/3/7	5/14/54/54	-

All (34) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	401	HEC	C3D-C2D	8.48	1.55	1.36
3	B	401	HEC	C3D-C2D	8.15	1.54	1.36
3	A	401	HEC	C3D-C2D	7.79	1.53	1.36
3	C	401	HEC	C3D-C2D	7.77	1.53	1.36
3	B	401	HEC	FE-NB	6.02	2.13	1.94
3	B	401	HEC	FE-ND	5.93	2.13	1.94
3	A	401	HEC	FE-ND	5.87	2.13	1.94
3	C	401	HEC	FE-NB	5.85	2.12	1.94
3	D	401	HEC	FE-NB	5.56	2.12	1.94
3	C	401	HEC	FE-ND	5.46	2.11	1.94
3	B	401	HEC	FE-NA	4.41	2.09	1.95
3	D	401	HEC	FE-ND	4.17	2.07	1.94
3	A	401	HEC	FE-NB	3.96	2.07	1.94
3	C	401	HEC	FE-NC	3.62	2.07	1.95
3	D	401	HEC	FE-NA	3.57	2.06	1.95
3	A	401	HEC	FE-NC	3.22	2.05	1.95
3	B	401	HEC	CMC-C2C	2.57	1.56	1.50
3	B	401	HEC	CMA-C3A	2.56	1.56	1.50
3	A	401	HEC	CMB-C2B	2.36	1.55	1.50
3	B	401	HEC	CMD-C2D	2.34	1.55	1.50
3	B	401	HEC	CMB-C2B	2.33	1.55	1.50
3	D	401	HEC	FE-NC	2.29	2.02	1.95
3	D	401	HEC	CMC-C2C	2.26	1.55	1.50
3	C	401	HEC	CMA-C3A	2.25	1.55	1.50
3	A	401	HEC	FE-NA	2.24	2.02	1.95
3	D	401	HEC	CMD-C2D	2.21	1.55	1.50
3	D	401	HEC	CMB-C2B	2.20	1.55	1.50
3	C	401	HEC	FE-NA	2.16	2.02	1.95
3	C	401	HEC	CMB-C2B	2.11	1.55	1.50
3	C	401	HEC	CMC-C2C	2.10	1.55	1.50
3	A	401	HEC	CMA-C3A	2.10	1.55	1.50
3	D	401	HEC	CAD-C3D	2.06	1.56	1.51

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	401	HEC	CAC-C3C	2.01	1.56	1.51
3	C	401	HEC	CMD-C2D	2.01	1.54	1.50

All (25) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	B	401	HEC	C1D-ND-C4D	6.07	112.40	105.21
3	D	401	HEC	C1D-ND-C4D	6.01	112.32	105.21
3	C	401	HEC	C1D-ND-C4D	5.27	111.44	105.21
3	A	401	HEC	C1D-ND-C4D	4.93	111.05	105.21
3	C	401	HEC	CAD-C3D-C4D	3.15	130.19	124.70
3	C	401	HEC	C2A-C1A-NA	-3.12	107.31	110.32
3	A	401	HEC	C3B-C4B-NB	-3.05	106.83	110.17
3	A	401	HEC	C4B-NB-C1B	2.98	108.73	105.21
3	B	401	HEC	C4B-NB-C1B	2.86	108.59	105.21
3	B	401	HEC	C3B-C4B-NB	-2.59	107.33	110.17
3	D	401	HEC	C3C-C4C-NC	-2.56	107.31	110.15
3	D	401	HEC	O2D-CGD-CBD	2.43	121.67	114.00
3	A	401	HEC	C2A-C1A-NA	-2.35	108.05	110.32
3	B	401	HEC	C3C-C4C-NC	-2.35	107.54	110.15
3	B	401	HEC	O2D-CGD-CBD	2.30	121.28	114.00
3	D	401	HEC	C4B-NB-C1B	2.30	107.93	105.21
3	B	401	HEC	C2A-C1A-NA	-2.29	108.12	110.32
3	A	401	HEC	CMD-C2D-C1D	2.24	128.54	125.03
3	D	401	HEC	O1D-CGD-CBD	-2.20	116.13	123.09
3	B	401	HEC	CAC-C3C-C4C	-2.18	121.91	124.91
3	B	401	HEC	CHC-C4B-C3B	2.17	129.23	125.23
3	C	401	HEC	C4B-NB-C1B	2.07	107.65	105.21
3	C	401	HEC	O2D-CGD-CBD	2.06	120.51	114.00
3	A	401	HEC	C3C-C4C-NC	-2.01	107.92	110.15
3	A	401	HEC	CAD-C3D-C4D	2.01	128.20	124.70

All (4) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	A	401	HEC	NA
3	B	401	HEC	NA
3	C	401	HEC	NA
3	D	401	HEC	NA

All (24) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	401	HEC	C3D-CAD-CBD-CGD
3	B	401	HEC	C2C-C3C-CAC-CBC
3	B	401	HEC	C4C-C3C-CAC-CBC
3	D	401	HEC	C2B-C3B-CAB-CBB
3	D	401	HEC	C3D-CAD-CBD-CGD
3	D	401	HEC	C4B-C3B-CAB-CBB
3	C	401	HEC	C3D-CAD-CBD-CGD
3	C	401	HEC	C2B-C3B-CAB-CBB
3	D	401	HEC	C4D-C3D-CAD-CBD
3	C	401	HEC	C4B-C3B-CAB-CBB
3	B	401	HEC	CAD-CBD-CGD-O1D
3	B	401	HEC	C2B-C3B-CAB-CBB
3	B	401	HEC	CAD-CBD-CGD-O2D
3	C	401	HEC	C4C-C3C-CAC-CBC
3	D	401	HEC	C2D-C3D-CAD-CBD
3	A	401	HEC	CAA-CBA-CGA-O2A
3	C	401	HEC	C4D-C3D-CAD-CBD
3	B	401	HEC	CAA-CBA-CGA-O2A
3	A	401	HEC	CAA-CBA-CGA-O1A
3	B	401	HEC	CAA-CBA-CGA-O1A
3	A	401	HEC	CAD-CBD-CGD-O1D
3	B	401	HEC	C4B-C3B-CAB-CBB
3	C	401	HEC	C2D-C3D-CAD-CBD
3	A	401	HEC	CAD-CBD-CGD-O2D

There are no ring outliers.

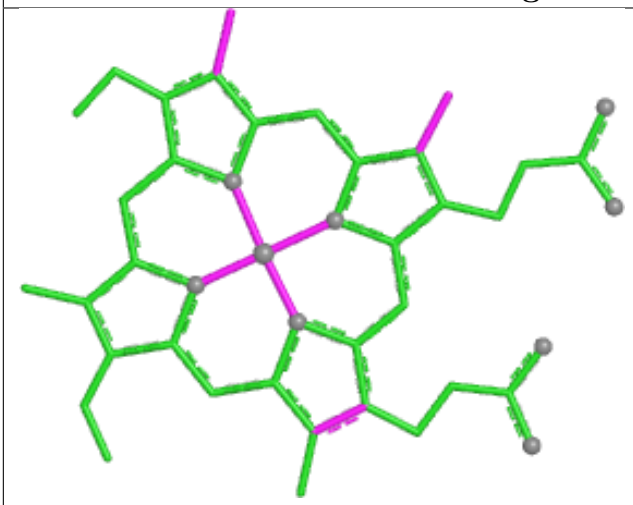
4 monomers are involved in 42 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	401	HEC	8	0
3	B	401	HEC	9	0
3	C	401	HEC	12	0
3	D	401	HEC	13	0

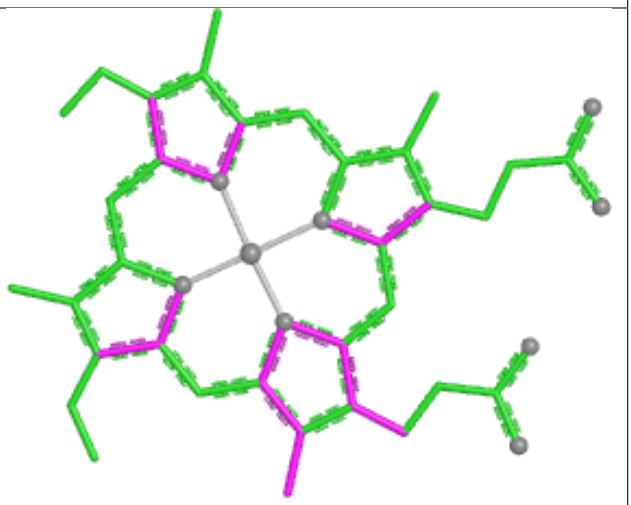
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included.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

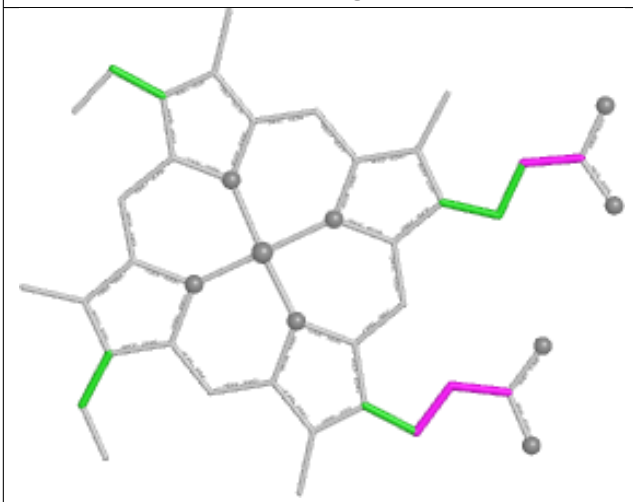
Ligand HEC A 401



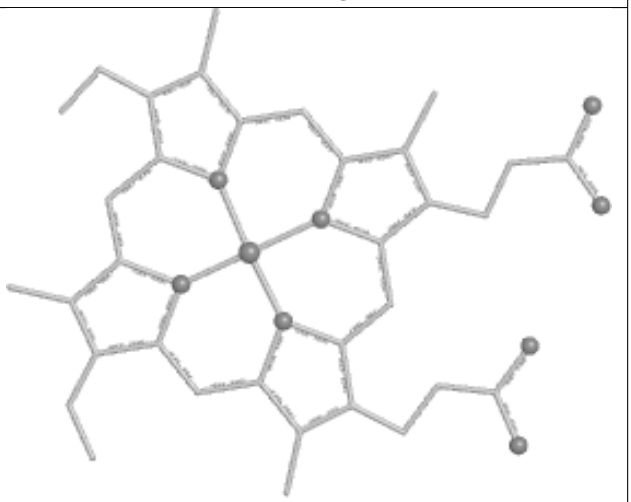
Bond lengths



Bond angles

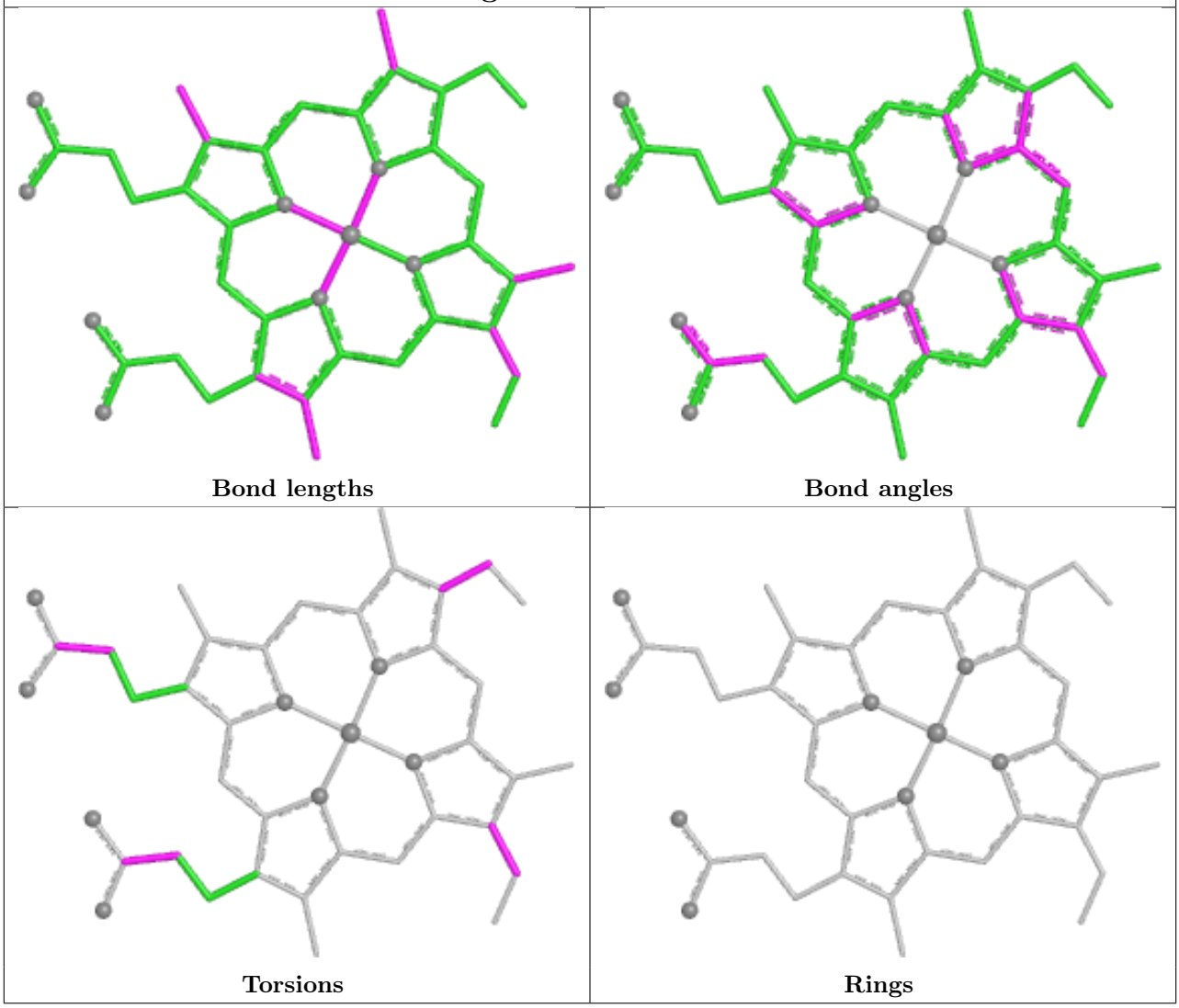


Torsions

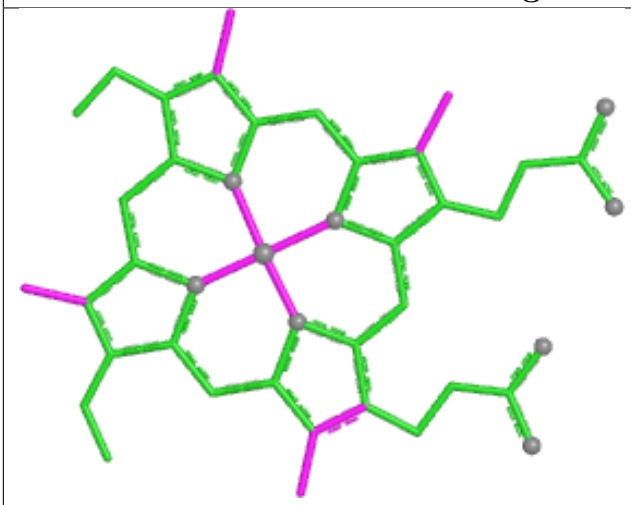


Rings

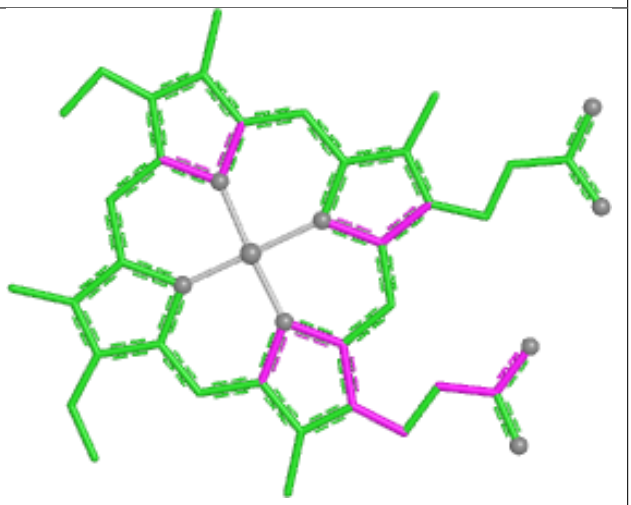
Ligand HEC B 401



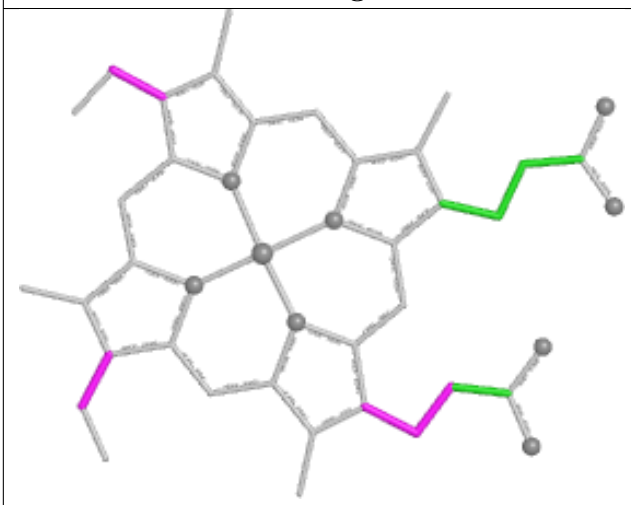
Ligand HEC C 401



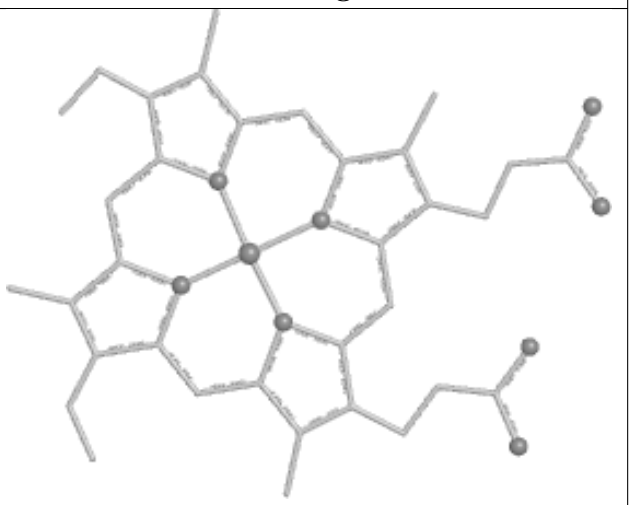
Bond lengths



Bond angles

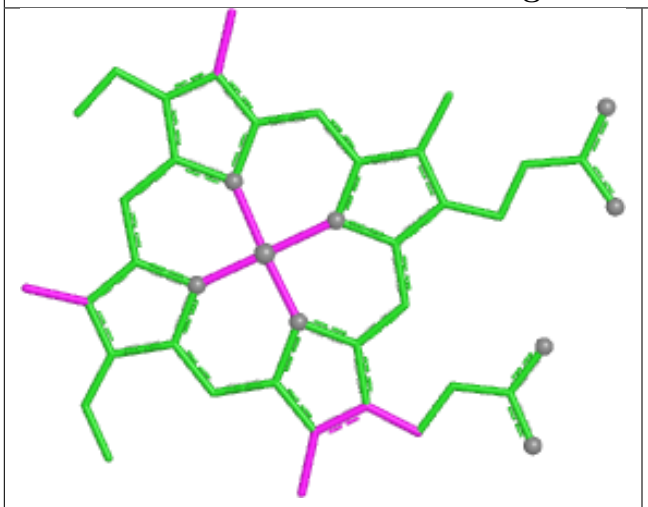


Torsions

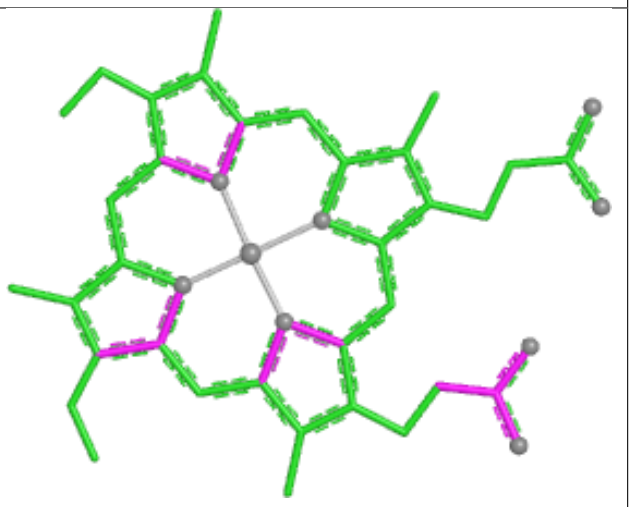


Rings

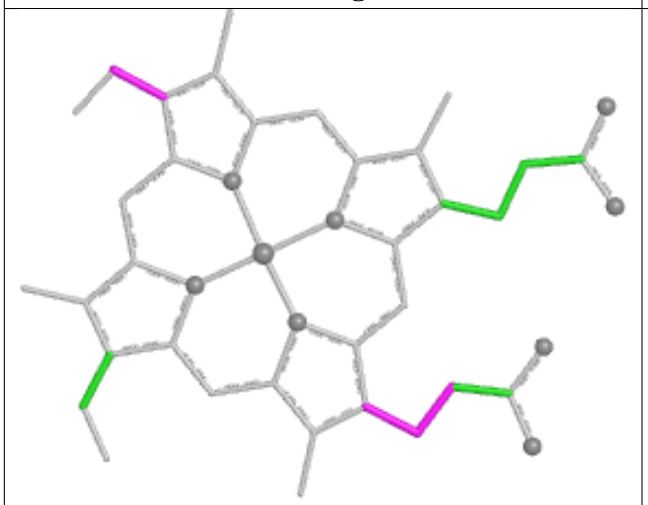
Ligand HEC D 401



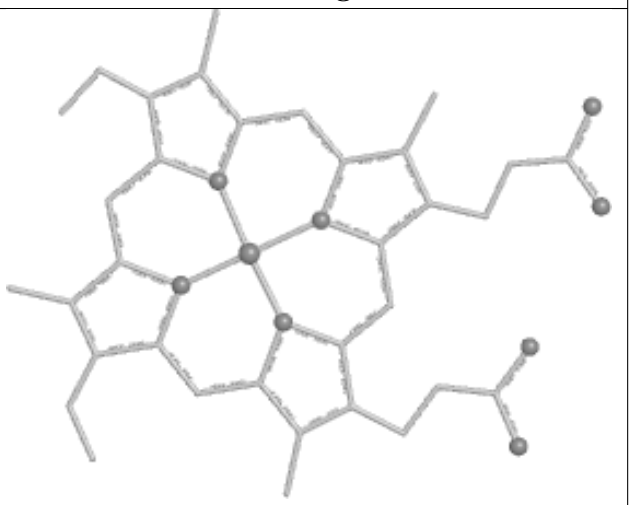
Bond lengths



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 (i)

6.1 Protein, DNA and RNA chains (i)

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, 95th 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	229/275 (83%)	-1.09	0	100	100	16, 32, 38, 42	1 (0%)
1	C	229/275 (83%)	-1.04	0	100	100	18, 33, 41, 50	1 (0%)
2	B	180/208 (86%)	-1.01	0	100	100	16, 33, 45, 53	4 (2%)
2	D	180/208 (86%)	-1.00	0	100	100	19, 33, 47, 51	5 (2%)
All	All	818/966 (84%)	-1.04	0	100	100	16, 33, 43, 53	11 (1%)

There are no RSRZ outliers to report.

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

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

6.3 Carbohydrates (i)

There are no oligosaccharides in this entry.

6.4 Ligands (i)

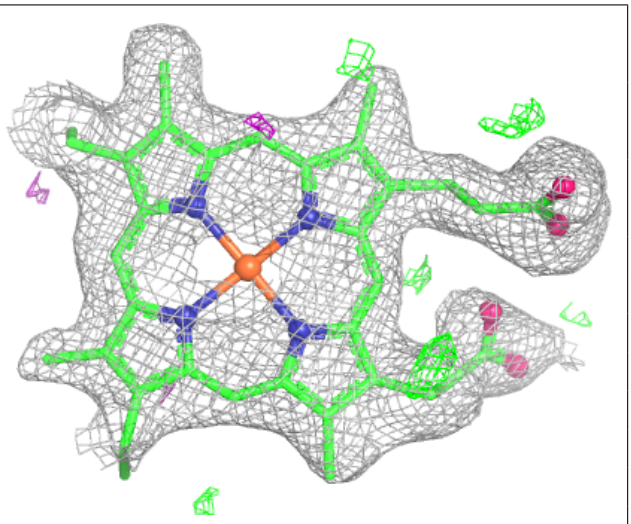
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

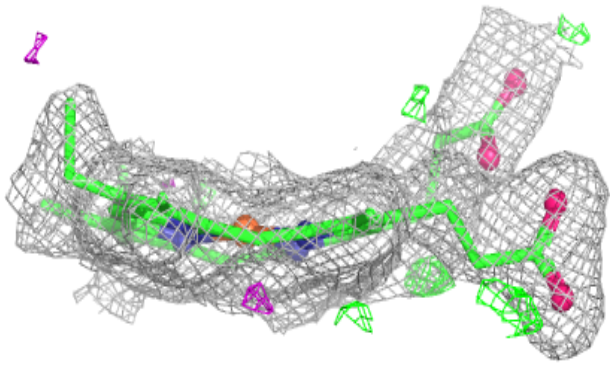
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	HEC	A	401	43/43	0.99	0.05	24,28,35,39	0
3	HEC	B	401	43/43	0.99	0.04	15,23,26,29	0
3	HEC	C	401	43/43	0.99	0.05	29,34,36,37	0
3	HEC	D	401	43/43	0.99	0.05	19,33,37,37	0

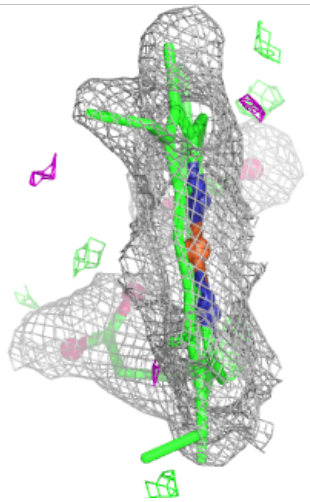
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around HEC A 401:

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

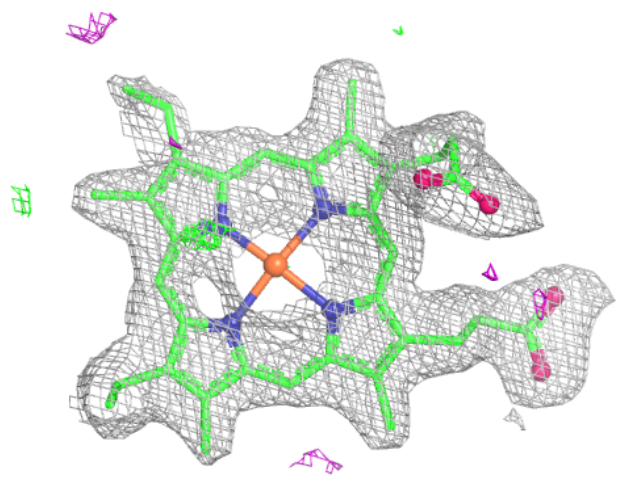


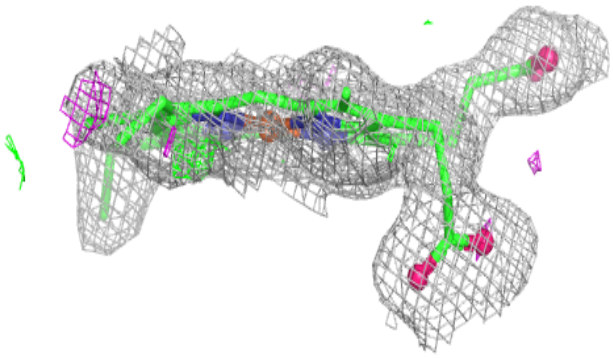


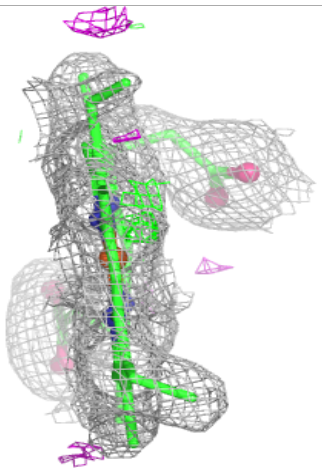


Electron density around HEC B 401:

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

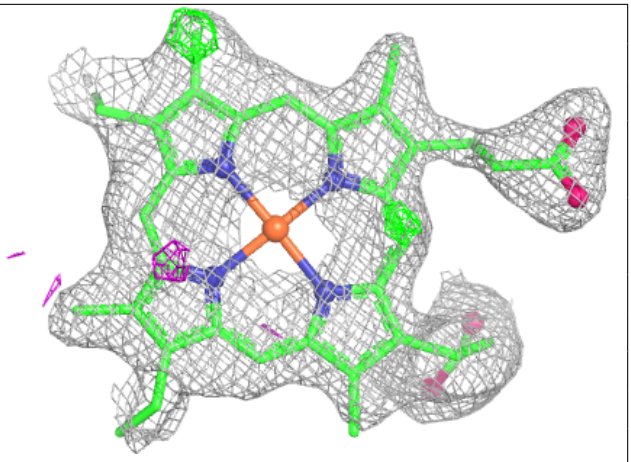


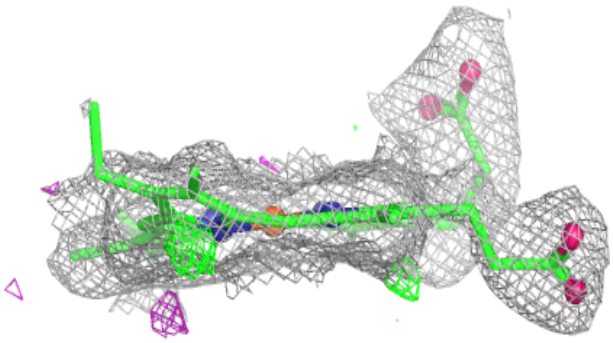


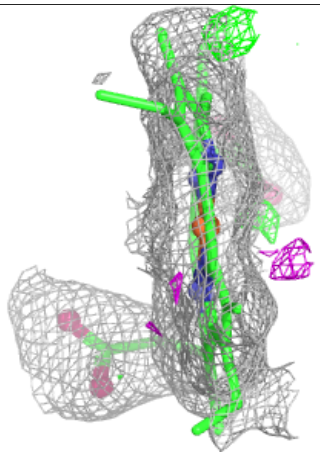


Electron density around HEC C 401:

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

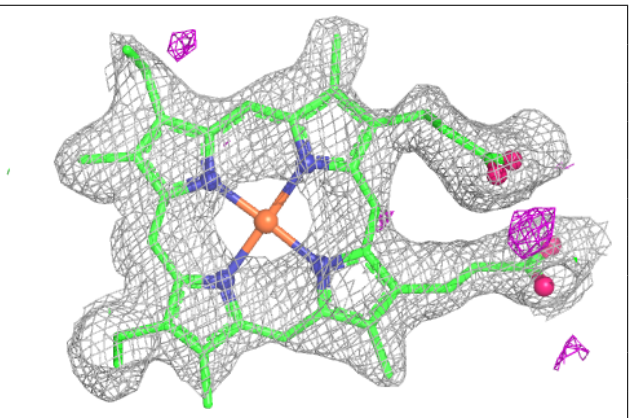


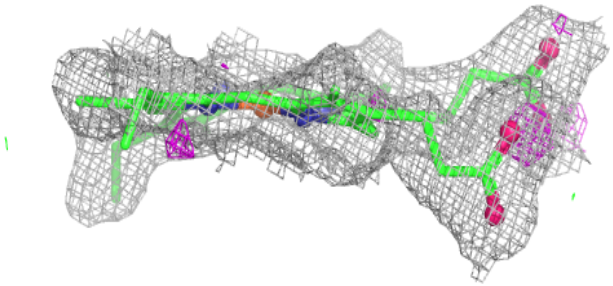


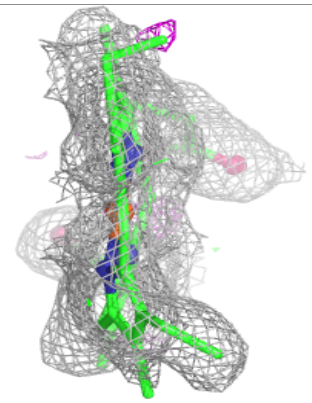


Electron density around HEC D 401:

$2mF_o$-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.