Full wwPDB X-ray Structure Validation Report (i)

Sep 17, 2023 – 02:07 PM EDT

PDB ID	:	4Y8W
Title	:	Crystal Structure of Human Cytochrome P450 21A2 Progesterone Complex
Authors	:	Pallan, P.S.; Lei, L.; Egli, M.
Deposited on	:	2015-02-16
Resolution	:	2.64 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:

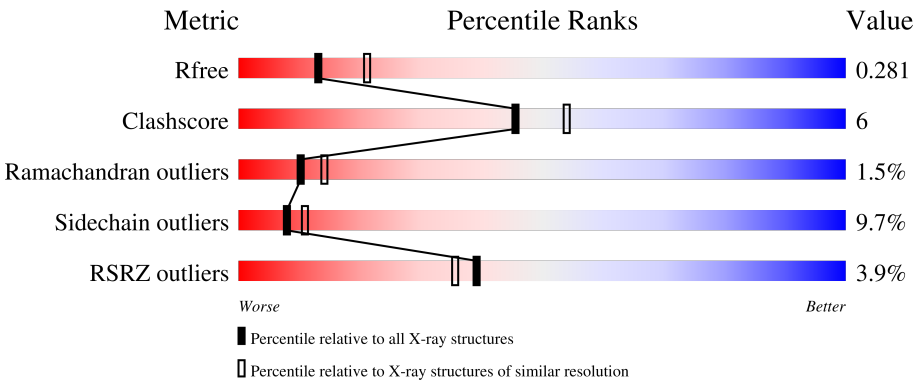
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 2.64 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1426 (2.66-2.62)
Clashscore	141614	1472 (2.66-2.62)
Ramachandran outliers	138981	1446 (2.66-2.62)
Sidechain outliers	138945	1446 (2.66-2.62)
RSRZ outliers	127900	1408 (2.66-2.62)

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	482	4% 69% 18% • 9%
1	B	482	4% 70% 18% • 9%
1	C	482	2% 75% 15% • 8%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 10832 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 Cytochrome P450 21-hydroxylase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	440	3493	2258	606	612	17	0	0	0
1	B	440	3513	2270	611	615	17	0	0	0
1	C	442	3531	2279	617	618	17	0	0	0

There are 48 discrepancies between the modelled and reference sequences:

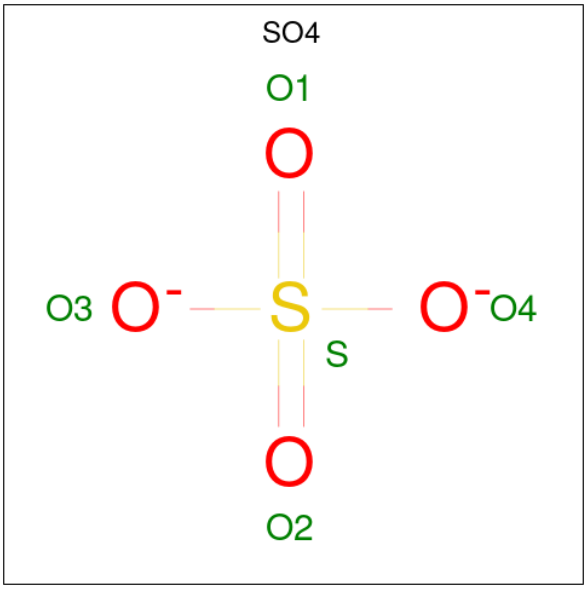
Chain	Residue	Modelled	Actual	Comment	Reference
A	20	MET	-	initiating methionine	UNP Q16874
A	21	ALA	-	expression tag	UNP Q16874
A	22	LYS	-	expression tag	UNP Q16874
A	23	LYS	-	expression tag	UNP Q16874
A	24	THR	-	expression tag	UNP Q16874
A	25	SER	-	expression tag	UNP Q16874
A	26	SER	-	expression tag	UNP Q16874
A	27	LYS	-	expression tag	UNP Q16874
A	28	GLY	-	expression tag	UNP Q16874
A	29	LYS	-	expression tag	UNP Q16874
A	496	HIS	-	expression tag	UNP Q16874
A	497	HIS	-	expression tag	UNP Q16874
A	498	HIS	-	expression tag	UNP Q16874
A	499	HIS	-	expression tag	UNP Q16874
A	500	HIS	-	expression tag	UNP Q16874
A	501	HIS	-	expression tag	UNP Q16874
B	20	MET	-	initiating methionine	UNP Q16874
B	21	ALA	-	expression tag	UNP Q16874
B	22	LYS	-	expression tag	UNP Q16874
B	23	LYS	-	expression tag	UNP Q16874
B	24	THR	-	expression tag	UNP Q16874
B	25	SER	-	expression tag	UNP Q16874
B	26	SER	-	expression tag	UNP Q16874

Continued on next page...

Continued from previous page...

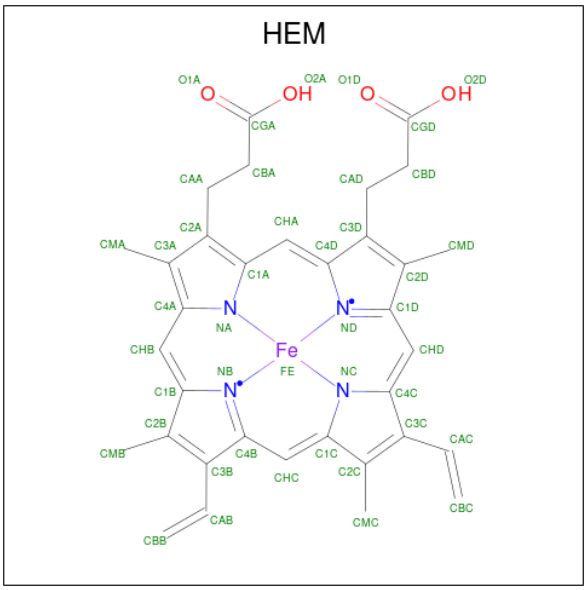
Chain	Residue	Modelled	Actual	Comment	Reference
B	27	LYS	-	expression tag	UNP Q16874
B	28	GLY	-	expression tag	UNP Q16874
B	29	LYS	-	expression tag	UNP Q16874
B	496	HIS	-	expression tag	UNP Q16874
B	497	HIS	-	expression tag	UNP Q16874
B	498	HIS	-	expression tag	UNP Q16874
B	499	HIS	-	expression tag	UNP Q16874
B	500	HIS	-	expression tag	UNP Q16874
B	501	HIS	-	expression tag	UNP Q16874
C	20	MET	-	initiating methionine	UNP Q16874
C	21	ALA	-	expression tag	UNP Q16874
C	22	LYS	-	expression tag	UNP Q16874
C	23	LYS	-	expression tag	UNP Q16874
C	24	THR	-	expression tag	UNP Q16874
C	25	SER	-	expression tag	UNP Q16874
C	26	SER	-	expression tag	UNP Q16874
C	27	LYS	-	expression tag	UNP Q16874
C	28	GLY	-	expression tag	UNP Q16874
C	29	LYS	-	expression tag	UNP Q16874
C	496	HIS	-	expression tag	UNP Q16874
C	497	HIS	-	expression tag	UNP Q16874
C	498	HIS	-	expression tag	UNP Q16874
C	499	HIS	-	expression tag	UNP Q16874
C	500	HIS	-	expression tag	UNP Q16874
C	501	HIS	-	expression tag	UNP Q16874

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



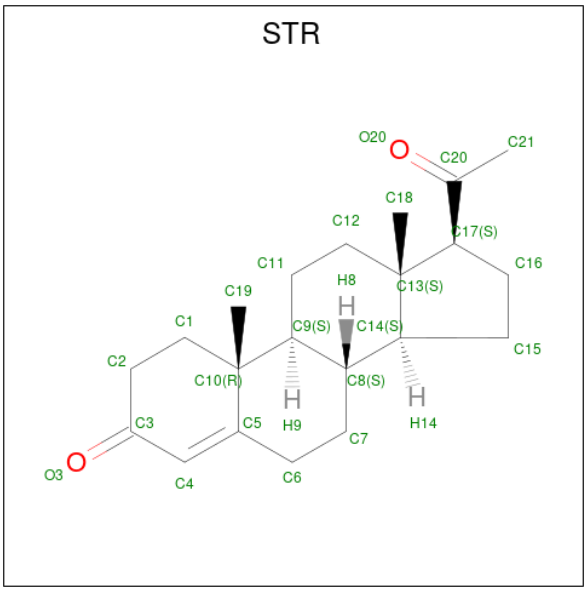
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O	S		
2	A	1	5	4	1	0	0
			Total	O	S		
2	A	1	5	4	1	0	0
			Total	O	S		
2	B	1	5	4	1	0	0
			Total	O	S		
2	C	1	5	4	1	0	0

- Molecule 3 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (formula: $C_{34}H_{32}FeN_4O_4$).



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

- Molecule 4 is PROGESTERONE (three-letter code: STR) (formula: $C_{21}H_{30}O_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C O 23 21 2	0	0
4	B	1	Total C O 23 21 2	0	0
4	C	1	Total C O 23 21 2	0	0

- Molecule 5 is water.

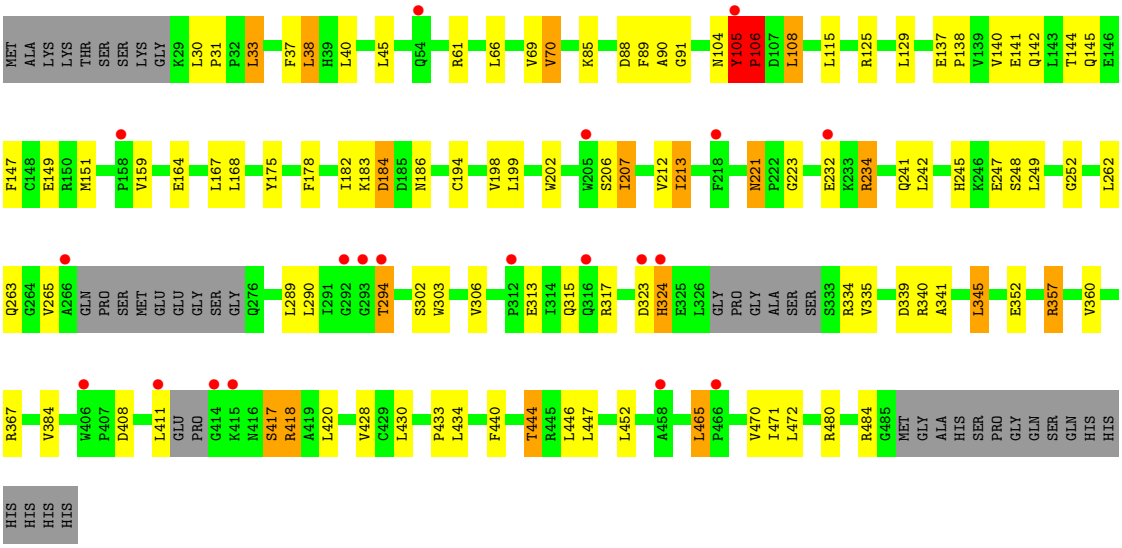
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	22	Total O 22 22	0	0
5	B	25	Total O 25 25	0	0
5	C	30	Total O 30 30	0	0

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.

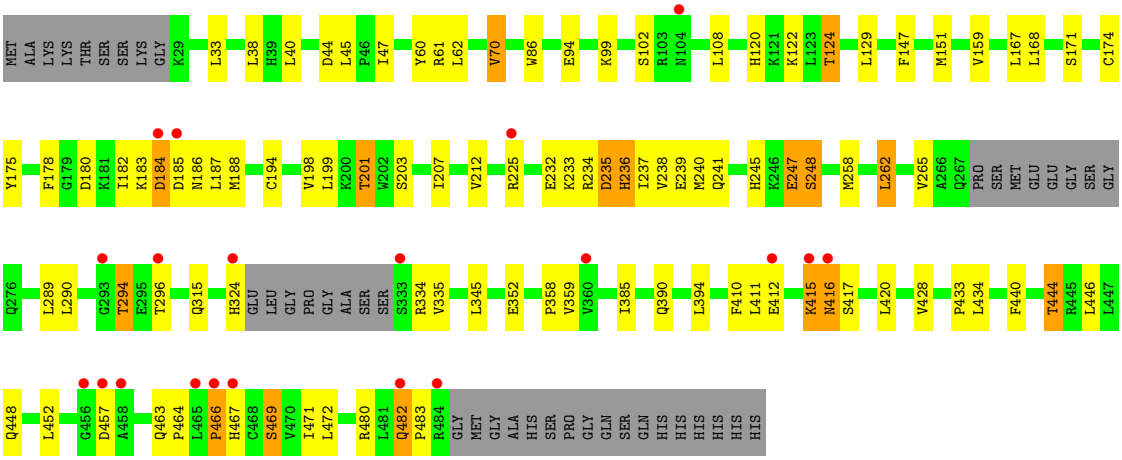
• Molecule 1: Cytochrome P450 21-hydroxylase

Chain A:



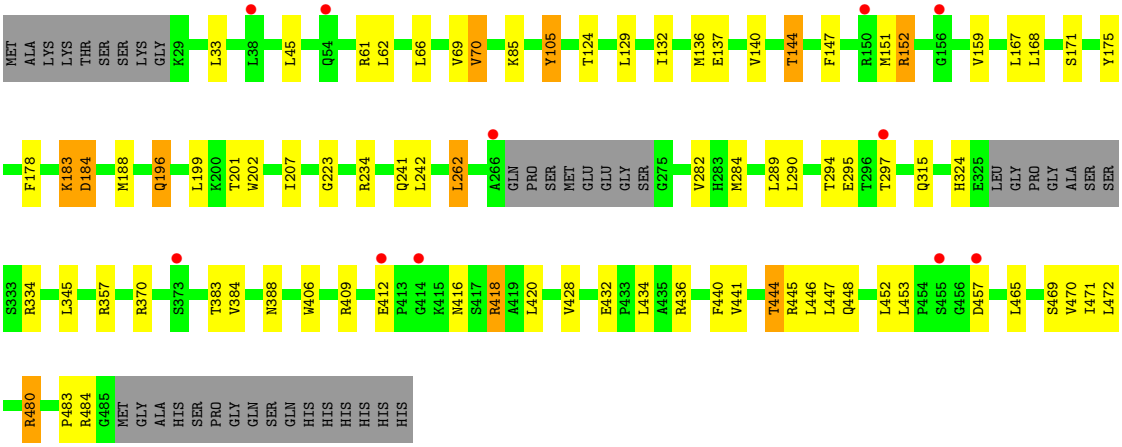
• Molecule 1: Cytochrome P450 21-hydroxylase

Chain B:



• Molecule 1: Cytochrome P450 21-hydroxylase

Chain C:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	150.38Å 86.86Å 108.92Å 90.00° 102.11° 90.00°	Depositor
Resolution (Å)	30.00 – 2.64 29.63 – 2.64	Depositor EDS
% Data completeness (in resolution range)	99.5 (30.00-2.64) 99.6 (29.63-2.64)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.40 (at 2.64Å)	Xtriage
Refinement program	REFMAC 5.8.0103	Depositor
R, R_{free}	0.231 , 0.281 0.231 , 0.281	Depositor DCC
R_{free} test set	3087 reflections (7.66%)	wwPDB-VP
Wilson B-factor (Å^2)	44.1	Xtriage
Anisotropy	0.057	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 48.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	10832	wwPDB-VP
Average B, all atoms (Å^2)	42.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.62% 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: HEM, STR, SO4

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	0.45	0/3581	0.76	8/4875 (0.2%)
1	B	0.41	0/3603	0.73	2/4905 (0.0%)
1	C	0.41	0/3621	0.68	0/4927
All	All	0.42	0/10805	0.72	10/14707 (0.1%)

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
1	A	0	1
1	B	0	2
All	All	0	3

There are no bond length outliers.

All (10) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	465	LEU	CA-CB-CG	6.65	130.60	115.30
1	A	108	LEU	CA-CB-CG	5.85	128.76	115.30
1	A	106	PRO	CA-N-CD	-5.84	103.32	111.50
1	A	38	LEU	CA-CB-CG	5.83	128.70	115.30
1	B	38	LEU	CA-CB-CG	5.62	128.22	115.30
1	A	33	LEU	CA-CB-CG	5.46	127.85	115.30
1	A	242	LEU	CA-CB-CG	5.22	127.31	115.30
1	A	234	ARG	N-CA-C	5.17	124.94	111.00
1	B	247	GLU	C-N-CA	5.08	134.41	121.70
1	A	105	TYR	CA-CB-CG	5.07	123.03	113.40

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	232	GLU	Peptide
1	B	235	ASP	Peptide
1	B	466	PRO	Peptide

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	3493	0	3528	55	0
1	B	3513	0	3563	46	0
1	C	3531	0	3586	35	0
2	A	10	0	0	0	0
2	B	5	0	0	0	0
2	C	5	0	0	0	0
3	A	43	0	30	1	0
3	B	43	0	30	0	0
3	C	43	0	30	2	0
4	A	23	0	30	1	0
4	B	23	0	30	1	0
4	C	23	0	30	0	0
5	A	22	0	0	0	0
5	B	25	0	0	0	0
5	C	30	0	0	0	0
All	All	10832	0	10857	138	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 6.

All (138) 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:178:PHE:HB2	1:A:182:ILE:HD11	1.42	0.96
1:B:315:GLN:HE22	1:B:452:LEU:H	1.14	0.92
1:A:315:GLN:HE22	1:A:452:LEU:H	1.16	0.87

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:418:ARG:HD2	1:A:418:ARG:H	1.43	0.84
1:C:315:GLN:HE22	1:C:452:LEU:H	1.31	0.79
1:A:105:TYR:HB3	1:A:106:PRO:CA	2.14	0.78
1:C:178:PHE:CD1	1:C:241:GLN:HG2	2.18	0.77
1:B:70:VAL:HG22	1:B:385:ILE:HG12	1.70	0.74
1:C:61:ARG:HG2	1:C:70:VAL:HG12	1.70	0.73
1:B:174:CYS:HB2	1:B:188:MET:HE1	1.73	0.70
1:A:151:MET:HE1	1:A:159:VAL:HG11	1.74	0.69
1:B:315:GLN:HE22	1:B:452:LEU:N	1.90	0.69
1:B:315:GLN:NE2	1:B:452:LEU:H	1.90	0.68
1:A:252:GLY:H	1:A:263:GLN:HE22	1.40	0.67
1:A:105:TYR:HB3	1:A:106:PRO:HA	1.76	0.66
1:B:236:HIS:O	1:B:238:VAL:N	2.29	0.66
1:B:415:LYS:HB2	1:B:416:ASN:HB3	1.78	0.66
1:B:440:PHE:O	1:B:444:THR:HB	1.96	0.65
1:B:290:LEU:O	1:B:294:THR:HG23	1.97	0.65
1:A:151:MET:CE	1:A:159:VAL:HG11	2.28	0.63
1:C:440:PHE:O	1:C:444:THR:HB	2.00	0.62
1:B:171:SER:HA	1:B:188:MET:HE1	1.82	0.62
1:C:140:VAL:O	1:C:144:THR:HG22	1.98	0.62
1:B:194:CYS:O	1:B:198:VAL:HG23	1.99	0.62
1:C:196:GLN:HE21	1:C:196:GLN:HA	1.65	0.61
1:A:178:PHE:CD1	1:A:241:GLN:HG2	2.35	0.61
1:A:184:ASP:H	1:A:186:ASN:H	1.50	0.60
1:B:247:GLU:H	1:B:248:SER:HB3	1.66	0.60
1:C:151:MET:HE1	1:C:159:VAL:HG11	1.82	0.60
1:B:296:THR:HA	1:B:359:VAL:HG11	1.84	0.59
1:C:406:TRP:O	1:C:409:ARG:HB3	2.03	0.58
1:B:203:SER:HB2	1:B:469:SER:HB2	1.86	0.58
1:A:138:PRO:HA	1:A:141:GLU:HG2	1.86	0.57
1:A:202:TRP:HZ3	1:A:470:VAL:HG21	1.69	0.57
1:A:290:LEU:O	1:A:294:THR:HG23	2.03	0.57
1:C:188:MET:HA	1:C:188:MET:HE2	1.87	0.57
1:B:175:TYR:OH	1:B:183:LYS:HB2	2.04	0.56
1:A:105:TYR:HB3	1:A:106:PRO:CB	2.36	0.55
3:A:603:HEM:HBC2	3:A:603:HEM:HMC2	1.89	0.55
1:C:290:LEU:O	1:C:294:THR:HG23	2.07	0.54
1:B:178:PHE:CD1	1:B:241:GLN:HG2	2.42	0.54
1:C:201:THR:O	1:C:207:ILE:HD12	2.07	0.54
1:B:40:LEU:HD11	1:B:212:VAL:HG11	1.90	0.53
1:A:61:ARG:HG2	1:A:70:VAL:HG12	1.91	0.53

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:188:MET:HA	1:C:188:MET:CE	2.39	0.53
1:B:178:PHE:CG	1:B:241:GLN:HG2	2.43	0.53
1:A:418:ARG:H	1:A:418:ARG:CD	2.17	0.52
1:B:182:ILE:HA	1:B:187:LEU:HD12	1.90	0.52
1:C:69:VAL:HG22	1:C:384:VAL:HB	1.91	0.51
1:A:339:ASP:O	1:A:341:ALA:N	2.43	0.51
1:C:406:TRP:CE3	1:C:409:ARG:HB2	2.45	0.51
1:C:171:SER:O	1:C:175:TYR:HD1	1.93	0.51
1:B:415:LYS:HD3	1:B:416:ASN:HB3	1.93	0.50
1:C:178:PHE:CG	1:C:241:GLN:HG2	2.47	0.50
1:C:418:ARG:HD3	1:C:418:ARG:H	1.77	0.50
1:A:252:GLY:H	1:A:263:GLN:NE2	2.07	0.50
1:A:105:TYR:HB3	1:A:106:PRO:HB3	1.94	0.50
1:B:184:ASP:H	1:B:186:ASN:H	1.60	0.50
1:A:212:VAL:HG23	1:A:213:ILE:HG13	1.94	0.49
1:A:140:VAL:O	1:A:144:THR:HG23	2.12	0.49
1:C:202:TRP:HZ3	1:C:470:VAL:HG21	1.78	0.49
1:A:89:PHE:O	1:A:91:GLY:N	2.46	0.48
1:B:171:SER:HA	1:B:188:MET:CE	2.41	0.48
1:A:69:VAL:HG22	1:A:384:VAL:HB	1.96	0.48
1:C:175:TYR:OH	1:C:183:LYS:HB2	2.13	0.48
1:C:105:TYR:HB2	1:C:284:MET:HG2	1.96	0.48
1:A:38:LEU:HD23	1:A:40:LEU:HD12	1.95	0.48
1:C:262:LEU:HD13	1:C:282:VAL:HG11	1.95	0.48
1:C:453:LEU:HB2	1:C:480:ARG:HB3	1.96	0.48
1:B:61:ARG:HG2	1:B:70:VAL:HG12	1.96	0.48
1:A:106:PRO:HD2	1:A:106:PRO:O	2.14	0.47
1:C:334:ARG:HD2	1:C:445:ARG:HH22	1.78	0.47
1:A:125:ARG:NH1	1:A:430:LEU:O	2.47	0.47
1:A:206:SER:O	1:A:207:ILE:HB	2.15	0.47
1:A:340:ARG:NH1	1:B:412:GLU:O	2.47	0.47
1:A:247:GLU:O	1:A:248:SER:HB3	2.15	0.47
1:A:315:GLN:NE2	1:A:452:LEU:H	1.98	0.47
1:C:409:ARG:HG3	1:C:416:ASN:HD21	1.80	0.47
1:A:194:CYS:O	1:A:198:VAL:HG23	2.15	0.47
1:A:175:TYR:OH	1:A:183:LYS:HB2	2.16	0.46
1:A:66:LEU:HD13	1:A:213:ILE:HG12	1.98	0.46
3:C:602:HEM:HMB2	3:C:602:HEM:HBB2	1.97	0.46
1:C:412:GLU:H	1:C:412:GLU:CD	2.20	0.45
1:B:147:PHE:O	1:B:151:MET:HG2	2.17	0.45
1:B:463:GLN:HA	1:B:464:PRO:HD3	1.87	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:178:PHE:CG	1:A:241:GLN:HG2	2.51	0.44
1:A:199:LEU:HD21	1:A:471:ILE:HD12	1.98	0.44
1:A:360:VAL:HG21	4:A:604:STR:H211	1.99	0.44
1:B:198:VAL:HG12	4:B:603:STR:H21	2.00	0.44
1:B:201:THR:HG22	1:B:207:ILE:CD1	2.47	0.44
1:B:258:MET:O	1:B:262:LEU:HB2	2.17	0.44
1:A:245:HIS:O	1:A:249:LEU:HB2	2.17	0.44
3:C:602:HEM:HMC1	3:C:602:HEM:HBC2	1.99	0.44
1:C:202:TRP:CZ3	1:C:470:VAL:HG21	2.52	0.44
1:A:147:PHE:O	1:A:151:MET:HG2	2.18	0.44
1:A:147:PHE:HE1	1:A:164:GLU:HB3	1.83	0.44
1:A:440:PHE:O	1:A:444:THR:HB	2.17	0.44
1:B:174:CYS:HB2	1:B:188:MET:CE	2.44	0.44
1:B:108:LEU:HD22	1:B:124:THR:HG21	1.99	0.44
1:B:482:GLN:HA	1:B:483:PRO:HD3	1.86	0.44
1:A:129:LEU:HD11	1:A:433:PRO:HG2	1.99	0.44
1:B:60:TYR:HE1	1:B:62:LEU:HD22	1.81	0.44
1:A:31:PRO:HG2	1:A:61:ARG:HG3	2.00	0.43
1:A:145:GLN:O	1:A:149:GLU:HG2	2.18	0.43
1:B:44:ASP:HB3	1:B:47:ILE:HB	2.00	0.43
1:B:129:LEU:HD11	1:B:433:PRO:HG2	2.00	0.43
1:B:247:GLU:N	1:B:248:SER:HB3	2.32	0.43
1:C:199:LEU:HD23	1:C:295:GLU:HG3	2.00	0.43
1:A:207:ILE:HD11	1:A:223:GLY:O	2.19	0.43
1:C:144:THR:HG21	1:C:441:VAL:HG12	2.01	0.43
1:A:85:LYS:HB3	1:A:88:ASP:HB2	2.01	0.43
1:A:302:SER:O	1:A:306:VAL:HG23	2.19	0.42
1:A:303:TRP:CD2	1:A:357:ARG:HG2	2.54	0.42
1:A:221:ASN:C	1:A:221:ASN:HD22	2.22	0.42
1:B:151:MET:HE1	1:B:159:VAL:HG11	2.01	0.42
1:C:144:THR:HA	1:C:147:PHE:HB3	2.01	0.42
1:B:94:GLU:HG2	1:B:99:LYS:HD2	2.02	0.42
1:B:358:PRO:HD2	1:B:390:GLN:HG3	2.02	0.42
1:C:207:ILE:HD11	1:C:223:GLY:O	2.20	0.42
1:A:206:SER:HB2	1:A:221:ASN:HD21	1.85	0.42
1:B:199:LEU:HD21	1:B:471:ILE:HD12	2.01	0.42
1:C:483:PRO:HA	1:C:484:ARG:HA	1.79	0.42
1:A:104:ASN:HD22	1:A:104:ASN:H	1.68	0.42
1:C:152:ARG:O	1:C:152:ARG:HD2	2.19	0.41
1:C:432:GLU:O	1:C:436:ARG:HG3	2.20	0.41
1:B:410:PHE:HD1	1:B:416:ASN:HB2	1.85	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:324:HIS:O	1:C:324:HIS:ND1	2.53	0.41
1:B:466:PRO:HB2	1:B:467:HIS:CG	2.56	0.41
1:A:178:PHE:HB2	1:A:182:ILE:CD1	2.32	0.41
1:A:317:ARG:NH2	1:A:408:ASP:OD2	2.54	0.41
1:A:323:ASP:O	1:A:324:HIS:HB2	2.20	0.41
1:B:183:LYS:HD3	1:B:188:MET:HB2	2.03	0.41
1:B:233:LYS:HA	1:B:235:ASP:N	2.36	0.41
1:C:132:ILE:HD12	1:C:136:MET:HB2	2.02	0.41
1:A:37:PHE:HD2	1:A:38:LEU:HD12	1.87	0.40
1:B:241:GLN:HG3	1:B:245:HIS:CE1	2.57	0.40
1:A:345:LEU:HD23	1:A:411:LEU:HD11	2.03	0.40
1:B:120:HIS:O	1:B:124:THR:HG22	2.22	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	432/482 (90%)	399 (92%)	24 (6%)	9 (2%)	7 9
1	B	434/482 (90%)	406 (94%)	20 (5%)	8 (2%)	8 11
1	C	436/482 (90%)	422 (97%)	12 (3%)	2 (0%)	29 43
All	All	1302/1446 (90%)	1227 (94%)	56 (4%)	19 (2%)	10 14

All (19) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	90	ALA
1	A	324	HIS
1	B	184	ASP
1	B	236	HIS
1	B	237	ILE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	184	ASP
1	A	207	ILE
1	B	248	SER
1	B	417	SER
1	B	469	SER
1	C	184	ASP
1	A	484	ARG
1	A	105	TYR
1	B	335	VAL
1	A	417	SER
1	B	86	TRP
1	C	469	SER
1	A	335	VAL
1	A	106	PRO

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	379/421 (90%)	345 (91%)	34 (9%)	9 14
1	B	384/421 (91%)	346 (90%)	38 (10%)	8 10
1	C	386/421 (92%)	346 (90%)	40 (10%)	7 9
All	All	1149/1263 (91%)	1037 (90%)	112 (10%)	8 11

All (112) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	30	LEU
1	A	33	LEU
1	A	45	LEU
1	A	70	VAL
1	A	108	LEU
1	A	115	LEU
1	A	137	GLU
1	A	142	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	167	LEU
1	A	168	LEU
1	A	213	ILE
1	A	221	ASN
1	A	234	ARG
1	A	262	LEU
1	A	265	VAL
1	A	289	LEU
1	A	294	THR
1	A	313	GLU
1	A	334	ARG
1	A	345	LEU
1	A	352	GLU
1	A	357	ARG
1	A	367	ARG
1	A	417	SER
1	A	418	ARG
1	A	420	LEU
1	A	428	VAL
1	A	434	LEU
1	A	444	THR
1	A	446	LEU
1	A	447	LEU
1	A	465	LEU
1	A	472	LEU
1	A	480	ARG
1	B	33	LEU
1	B	45	LEU
1	B	70	VAL
1	B	102	SER
1	B	122	LYS
1	B	124	THR
1	B	167	LEU
1	B	168	LEU
1	B	180	ASP
1	B	185	ASP
1	B	201	THR
1	B	225	ARG
1	B	232	GLU
1	B	234	ARG
1	B	239	GLU
1	B	240	MET

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	262	LEU
1	B	265	VAL
1	B	289	LEU
1	B	294	THR
1	B	324	HIS
1	B	334	ARG
1	B	345	LEU
1	B	352	GLU
1	B	394	LEU
1	B	411	LEU
1	B	415	LYS
1	B	416	ASN
1	B	420	LEU
1	B	428	VAL
1	B	434	LEU
1	B	444	THR
1	B	446	LEU
1	B	448	GLN
1	B	457	ASP
1	B	472	LEU
1	B	480	ARG
1	B	482	GLN
1	C	33	LEU
1	C	45	LEU
1	C	62	LEU
1	C	66	LEU
1	C	70	VAL
1	C	85	LYS
1	C	105	TYR
1	C	124	THR
1	C	129	LEU
1	C	137	GLU
1	C	144	THR
1	C	152	ARG
1	C	167	LEU
1	C	168	LEU
1	C	183	LYS
1	C	184	ASP
1	C	196	GLN
1	C	234	ARG
1	C	242	LEU
1	C	262	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	289	LEU
1	C	297	THR
1	C	345	LEU
1	C	357	ARG
1	C	370	ARG
1	C	383	THR
1	C	388	ASN
1	C	418	ARG
1	C	420	LEU
1	C	428	VAL
1	C	434	LEU
1	C	444	THR
1	C	446	LEU
1	C	447	LEU
1	C	448	GLN
1	C	457	ASP
1	C	465	LEU
1	C	471	ILE
1	C	472	LEU
1	C	480	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (26) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	67	GLN
1	A	104	ASN
1	A	221	ASN
1	A	263	GLN
1	A	315	GLN
1	A	388	ASN
1	A	448	GLN
1	B	67	GLN
1	B	196	GLN
1	B	241	GLN
1	B	244	GLN
1	B	263	GLN
1	B	276	GLN
1	B	283	HIS
1	B	315	GLN
1	B	319	GLN
1	B	347	ASN
1	B	388	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	448	GLN
1	C	196	GLN
1	C	241	GLN
1	C	263	GLN
1	C	315	GLN
1	C	319	GLN
1	C	388	ASN
1	C	448	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

10 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	$\#\|Z\| > 2$
2	SO4	C	601	-	4,4,4	0.30	0	6,6,6	0.11	0
3	HEM	C	602	1	41,50,50	1.32	5 (12%)	45,82,82	1.71	10 (22%)
2	SO4	B	601	-	4,4,4	0.32	0	6,6,6	0.07	0
4	STR	B	603	-	26,26,26	0.65	0	42,42,42	1.25	5 (11%)
4	STR	C	603	-	26,26,26	0.67	0	42,42,42	1.30	6 (14%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
3	HEM	B	602	1	41,50,50	1.32	5 (12%)	45,82,82	1.61	8 (17%)
2	SO4	A	601	2	4,4,4	0.33	0	6,6,6	0.09	0
2	SO4	A	602	2	4,4,4	0.32	0	6,6,6	0.09	0
3	HEM	A	603	1	41,50,50	1.31	5 (12%)	45,82,82	1.65	10 (22%)
4	STR	A	604	-	26,26,26	0.69	0	42,42,42	1.27	5 (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
3	HEM	C	602	1	-	0/12/54/54	-
4	STR	B	603	-	-	4/4/62/62	0/4/4/4
4	STR	C	603	-	-	4/4/62/62	0/4/4/4
3	HEM	B	602	1	-	0/12/54/54	-
3	HEM	A	603	1	-	0/12/54/54	-
4	STR	A	604	-	-	4/4/62/62	0/4/4/4

All (15) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	602	HEM	C1B-NB	-3.27	1.34	1.40
3	B	602	HEM	C1B-NB	-3.26	1.34	1.40
3	A	603	HEM	C1B-NB	-3.17	1.34	1.40
3	B	602	HEM	FE-NB	3.13	2.12	1.96
3	A	603	HEM	FE-NB	3.04	2.11	1.96
3	C	602	HEM	C4D-ND	-2.95	1.35	1.40
3	C	602	HEM	FE-NB	2.92	2.11	1.96
3	A	603	HEM	C4D-ND	-2.89	1.35	1.40
3	B	602	HEM	C4D-ND	-2.87	1.35	1.40
3	C	602	HEM	CHB-C1B	2.36	1.41	1.35
3	A	603	HEM	CHB-C1B	2.32	1.40	1.35
3	A	603	HEM	C3B-C4B	2.32	1.49	1.44
3	B	602	HEM	C3B-C4B	2.25	1.49	1.44
3	C	602	HEM	C3B-C4B	2.19	1.49	1.44
3	B	602	HEM	CHB-C1B	2.13	1.40	1.35

All (44) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	603	HEM	C1B-NB-C4B	4.48	109.70	105.07
3	C	602	HEM	CHC-C4B-NB	4.23	129.03	124.43
3	B	602	HEM	C1B-NB-C4B	4.04	109.24	105.07
3	C	602	HEM	C1B-NB-C4B	4.01	109.22	105.07
3	C	602	HEM	CHD-C1D-ND	3.93	128.70	124.43
3	A	603	HEM	CHC-C4B-NB	3.82	128.58	124.43
3	B	602	HEM	CHC-C4B-NB	3.82	128.58	124.43
3	B	602	HEM	CHD-C1D-ND	3.53	128.27	124.43
3	A	603	HEM	CHA-C4D-ND	3.29	128.44	124.38
3	B	602	HEM	CHA-C4D-ND	3.26	128.41	124.38
3	C	602	HEM	CHA-C4D-ND	3.21	128.34	124.38
3	A	603	HEM	CHD-C1D-ND	3.03	127.72	124.43
3	C	602	HEM	CHD-C1D-C2D	-2.98	120.32	124.98
4	B	603	STR	C21-C20-C17	2.90	121.78	117.56
4	B	603	STR	C15-C14-C13	2.90	107.33	103.84
3	A	603	HEM	CHA-C4D-C3D	-2.84	119.99	125.33
4	C	603	STR	C15-C14-C13	2.80	107.22	103.84
4	A	604	STR	C6-C5-C4	-2.77	116.24	120.87
4	C	603	STR	C21-C20-C17	2.59	121.33	117.56
3	B	602	HEM	CHA-C4D-C3D	-2.57	120.50	125.33
4	A	604	STR	C15-C14-C13	2.54	106.91	103.84
3	B	602	HEM	CHB-C1B-NB	2.54	127.51	124.38
3	C	602	HEM	CHA-C4D-C3D	-2.52	120.59	125.33
4	A	604	STR	C21-C20-C17	2.52	121.22	117.56
4	B	603	STR	C6-C5-C4	-2.51	116.66	120.87
3	B	602	HEM	CHD-C1D-C2D	-2.49	121.08	124.98
4	C	603	STR	C2-C3-C4	2.49	120.57	116.74
4	C	603	STR	C6-C5-C4	-2.44	116.78	120.87
3	A	603	HEM	CHD-C1D-C2D	-2.42	121.20	124.98
4	A	604	STR	C6-C5-C10	2.36	121.11	116.77
4	B	603	STR	C2-C3-C4	2.25	120.21	116.74
4	B	603	STR	C6-C5-C10	2.23	120.88	116.77
3	B	602	HEM	O2D-CGD-O1D	-2.22	117.76	123.30
3	A	603	HEM	CMA-C3A-C4A	-2.22	125.05	128.46
3	A	603	HEM	CHB-C1B-NB	2.22	127.12	124.38
3	A	603	HEM	CBA-CAA-C2A	2.19	116.36	112.62
3	C	602	HEM	CMA-C3A-C4A	-2.19	125.10	128.46
3	C	602	HEM	CBA-CAA-C2A	2.17	116.32	112.62
3	C	602	HEM	C4B-C3B-C2B	-2.13	105.42	107.11
3	C	602	HEM	CMC-C2C-C3C	2.11	128.63	124.68
3	A	603	HEM	O2D-CGD-O1D	-2.08	118.12	123.30
4	C	603	STR	C16-C17-C13	2.01	106.00	104.21
4	C	603	STR	C6-C5-C10	2.01	120.48	116.77

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	604	STR	C2-C3-C4	2.00	119.82	116.74

There are no chirality outliers.

All (12) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	C	603	STR	C13-C17-C20-C21
4	C	603	STR	C16-C17-C20-C21
4	A	604	STR	C13-C17-C20-O20
4	B	603	STR	C13-C17-C20-O20
4	B	603	STR	C13-C17-C20-C21
4	A	604	STR	C16-C17-C20-C21
4	B	603	STR	C16-C17-C20-C21
4	C	603	STR	C16-C17-C20-O20
4	A	604	STR	C13-C17-C20-C21
4	C	603	STR	C13-C17-C20-O20
4	A	604	STR	C16-C17-C20-O20
4	B	603	STR	C16-C17-C20-O20

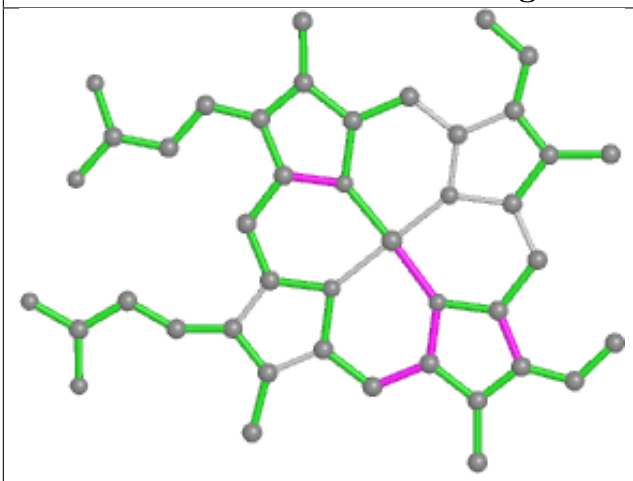
There are no ring outliers.

4 monomers are involved in 5 short contacts:

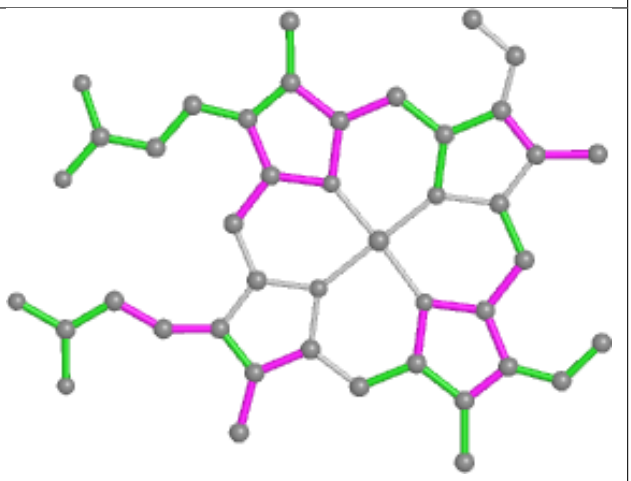
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	602	HEM	2	0
4	B	603	STR	1	0
3	A	603	HEM	1	0
4	A	604	STR	1	0

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.

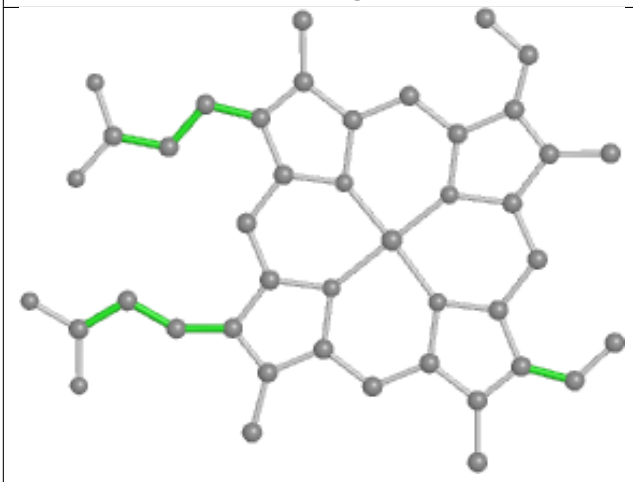
Ligand HEM C 602



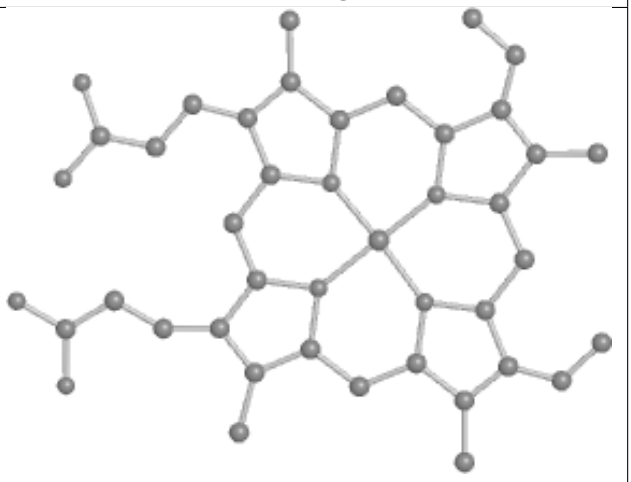
Bond lengths



Bond angles

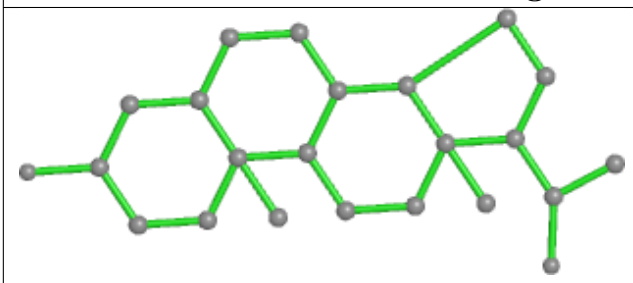


Torsions

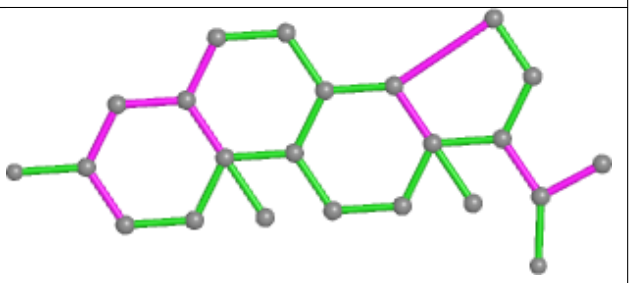


Rings

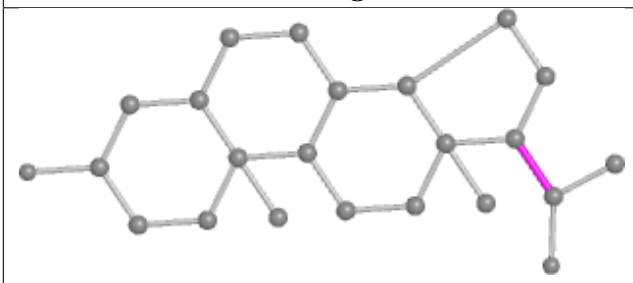
Ligand STR B 603



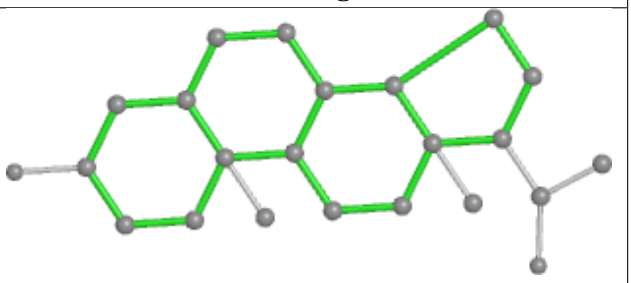
Bond lengths



Bond angles

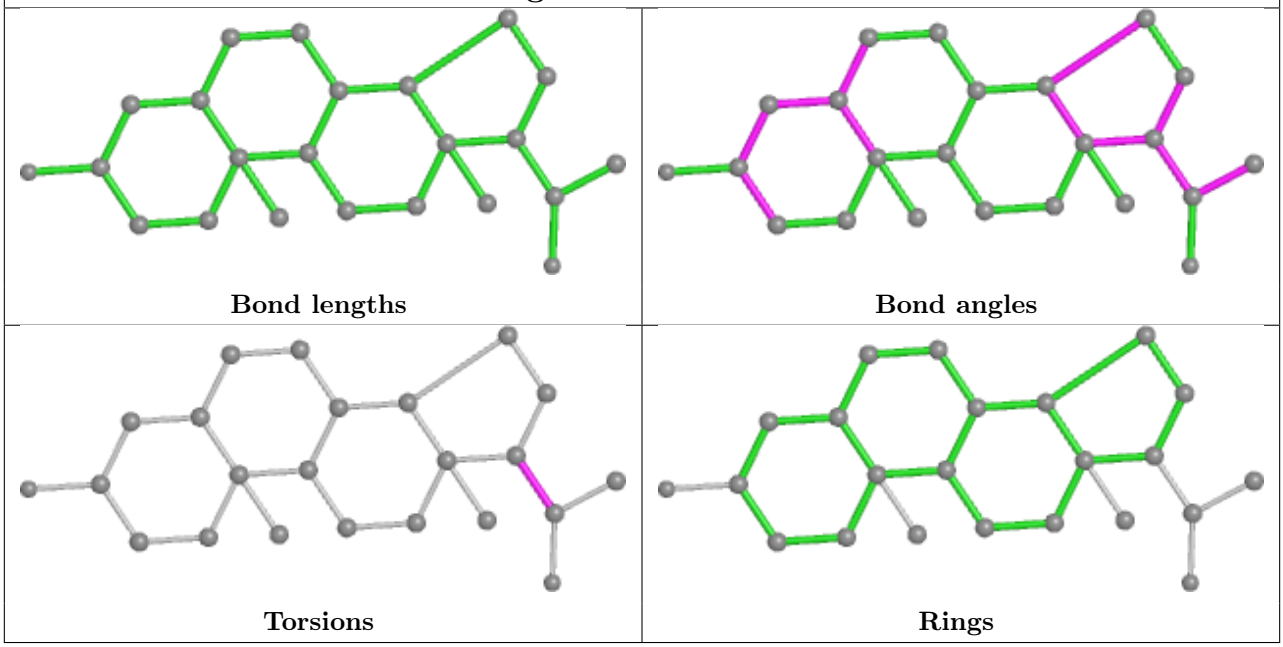


Torsions

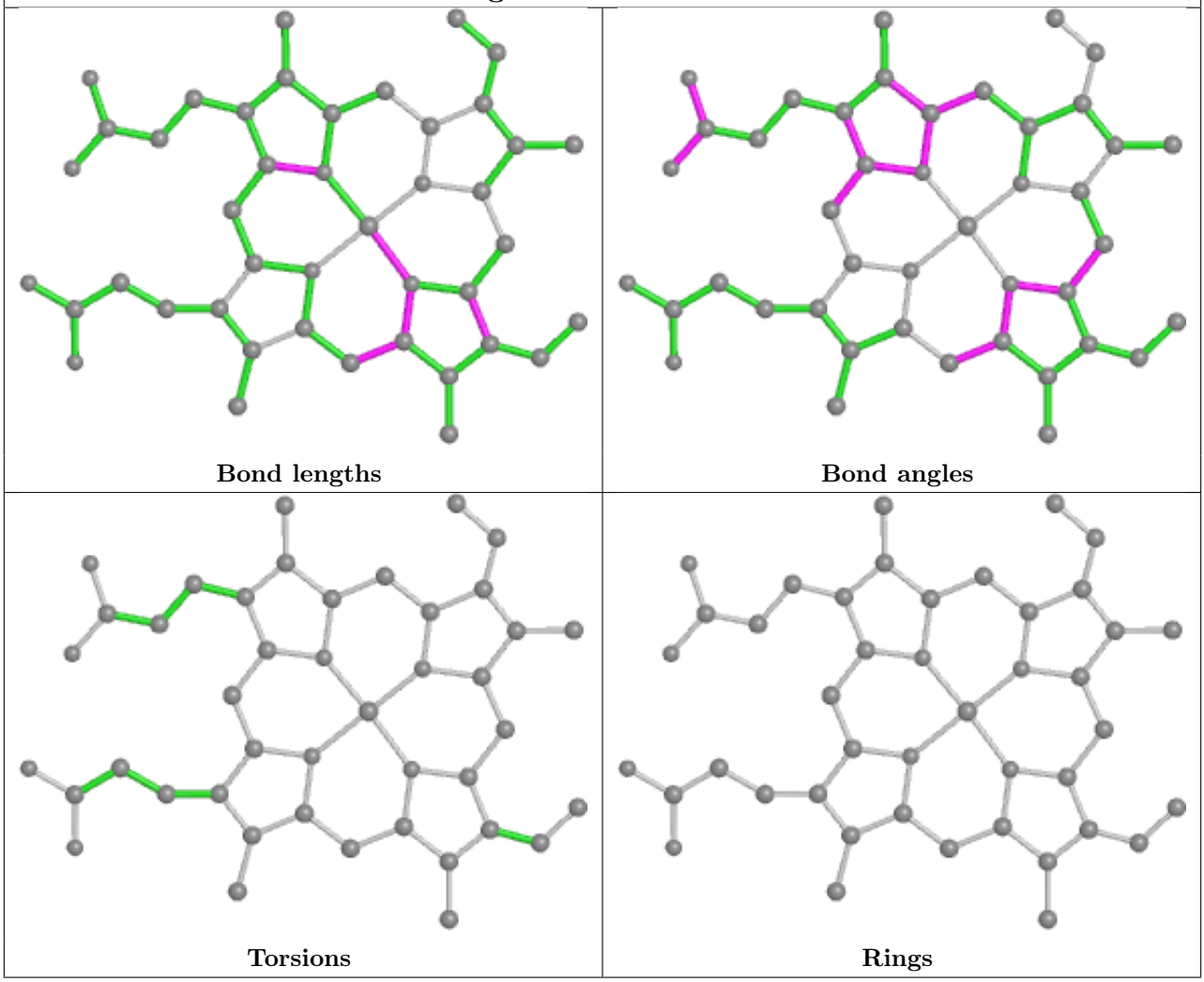


Rings

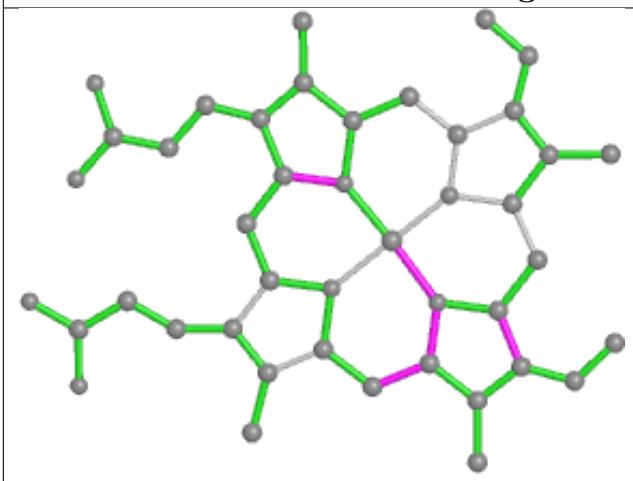
Ligand STR C 603



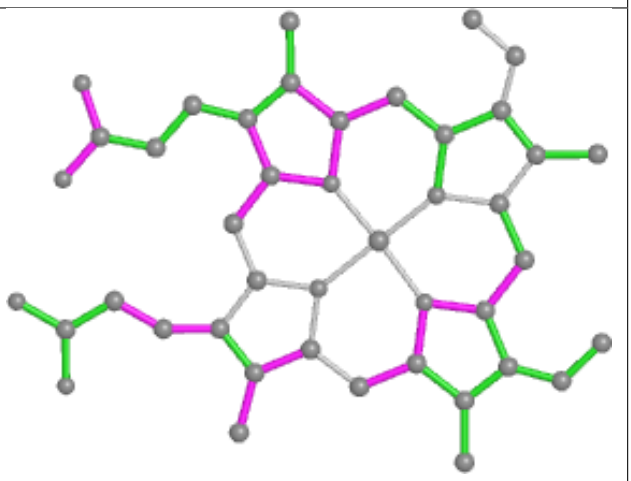
Ligand HEM B 602



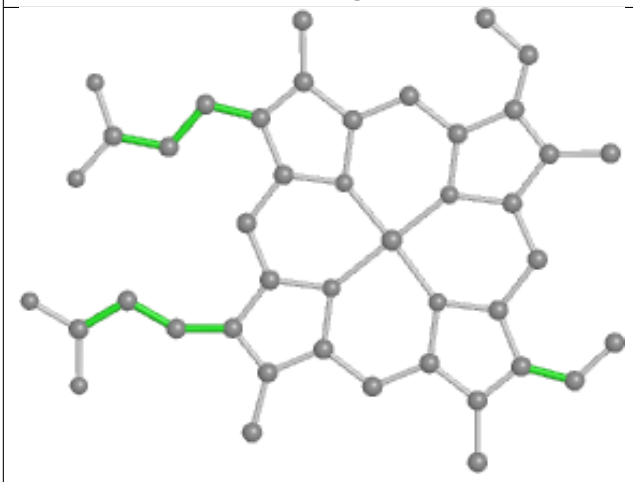
Ligand HEM A 603



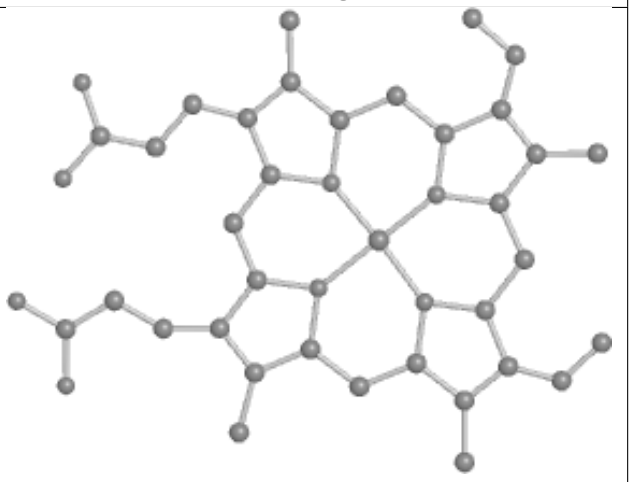
Bond lengths



Bond angles

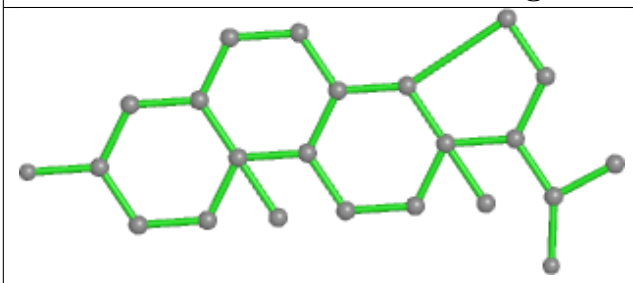


Torsions

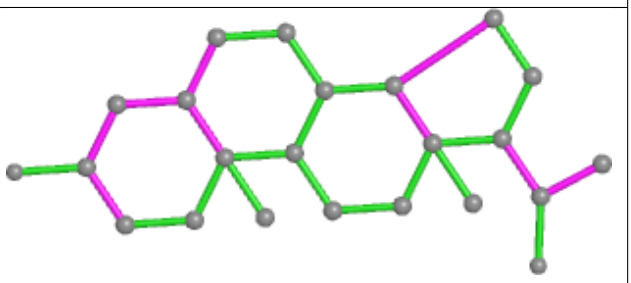


Rings

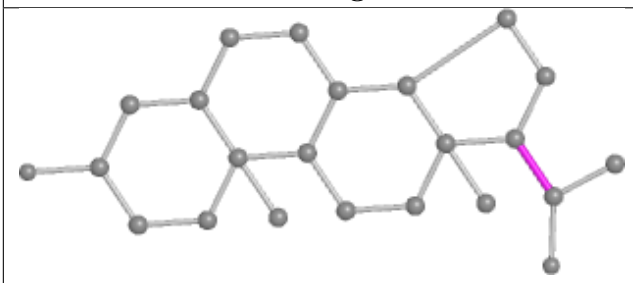
Ligand STR A 604



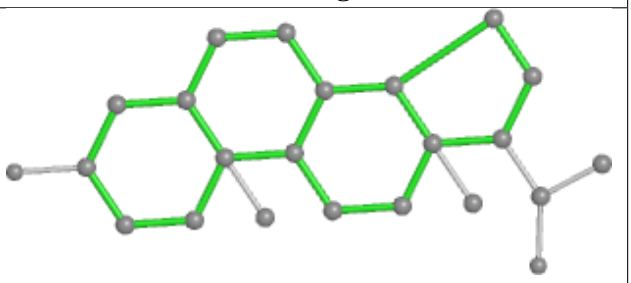
Bond lengths



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	440/482 (91%)	0.26	20 (4%)	33	30	25, 41, 58, 81	0
1	B	440/482 (91%)	0.24	20 (4%)	33	30	28, 41, 63, 82	0
1	C	442/482 (91%)	0.21	11 (2%)	57	53	27, 43, 60, 73	0
All	All	1322/1446 (91%)	0.24	51 (3%)	39	36	25, 42, 61, 82	0

All (51) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	457	ASP	4.1
1	B	324	HIS	3.7
1	B	458	ALA	3.6
1	C	54	GLN	3.4
1	A	411	LEU	3.4
1	B	466	PRO	3.3
1	B	412	GLU	3.2
1	B	456	GLY	3.0
1	A	292	GLY	3.0
1	C	266	ALA	2.9
1	A	458	ALA	2.9
1	B	482	GLN	2.9
1	A	54	GLN	2.9
1	B	333	SER	2.8
1	A	312	PRO	2.8
1	A	105	TYR	2.7
1	A	323	ASP	2.7
1	B	296	THR	2.7
1	C	38	LEU	2.6
1	B	484	ARG	2.6
1	A	293	GLY	2.5
1	A	414	GLY	2.5
1	A	316	GLN	2.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	360	VAL	2.5
1	B	185	ASP	2.5
1	C	156	GLY	2.4
1	B	225	ARG	2.4
1	B	467	HIS	2.4
1	C	457	ASP	2.4
1	B	104	ASN	2.3
1	C	412	GLU	2.3
1	B	184	ASP	2.3
1	A	294	THR	2.3
1	C	414	GLY	2.2
1	A	158	PRO	2.2
1	C	373	SER	2.2
1	A	232	GLU	2.2
1	C	150	ARG	2.2
1	A	406	TRP	2.2
1	A	324	HIS	2.2
1	B	465	LEU	2.2
1	B	416	ASN	2.2
1	C	297	THR	2.1
1	A	415	LYS	2.1
1	A	218	PHE	2.1
1	B	415	LYS	2.1
1	A	205	TRP	2.1
1	A	466	PRO	2.0
1	B	293	GLY	2.0
1	C	455	SER	2.0
1	A	266	ALA	2.0

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 monosaccharides in this entry.

6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum,

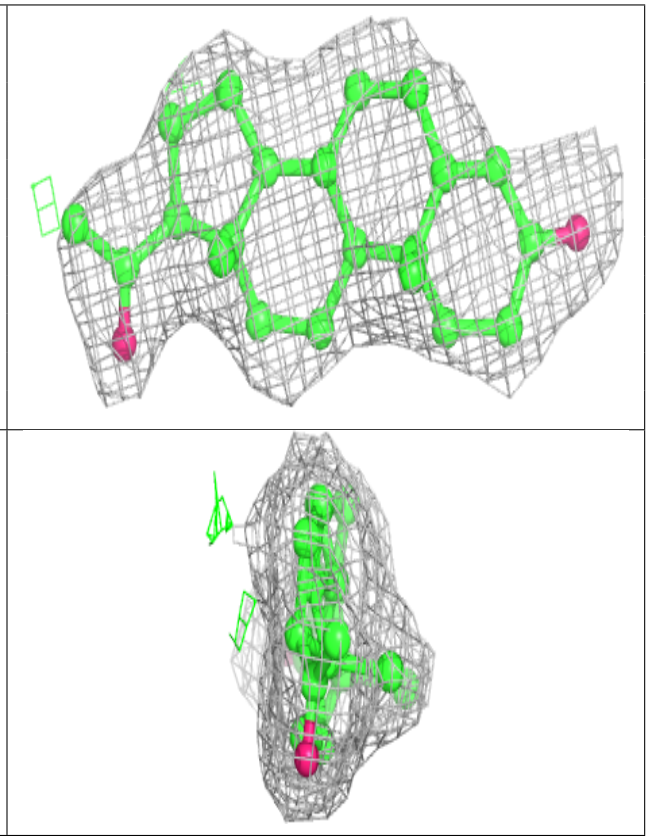
median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	STR	A	604	23/23	0.90	0.23	30,30,31,32	0
4	STR	B	603	23/23	0.93	0.23	30,30,33,35	0
4	STR	C	603	23/23	0.94	0.20	34,35,36,38	0
3	HEM	B	602	43/43	0.96	0.21	26,27,29,29	0
3	HEM	A	603	43/43	0.96	0.20	24,25,26,26	0
3	HEM	C	602	43/43	0.97	0.18	24,25,26,26	0
2	SO4	A	601	5/5	0.98	0.11	30,30,30,30	5
2	SO4	A	602	5/5	0.98	0.10	36,36,36,37	5
2	SO4	B	601	5/5	0.99	0.11	41,41,42,42	0
2	SO4	C	601	5/5	0.99	0.13	31,31,32,32	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.

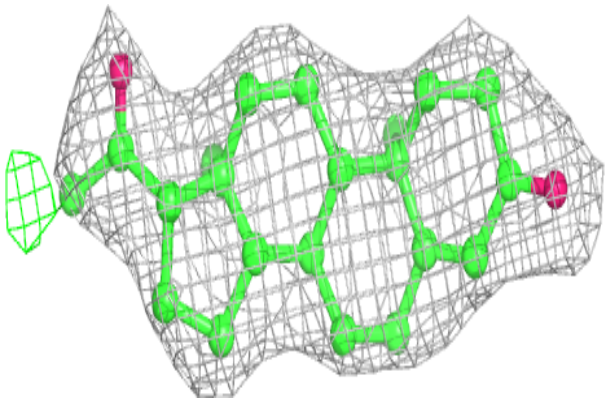
Electron density around STR A 604:

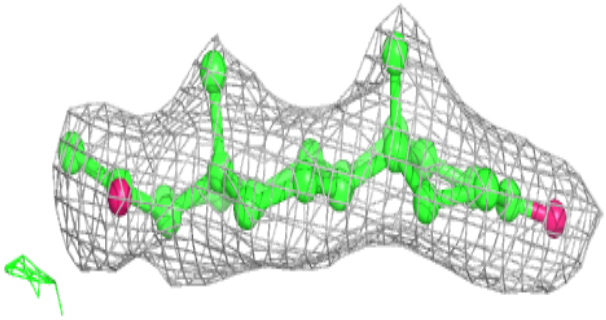
2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative)
and green (positive)

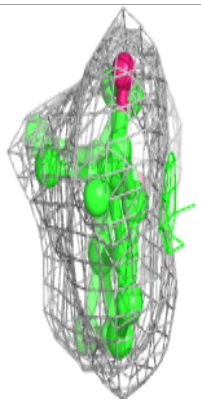


Electron density around STR B 603:

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

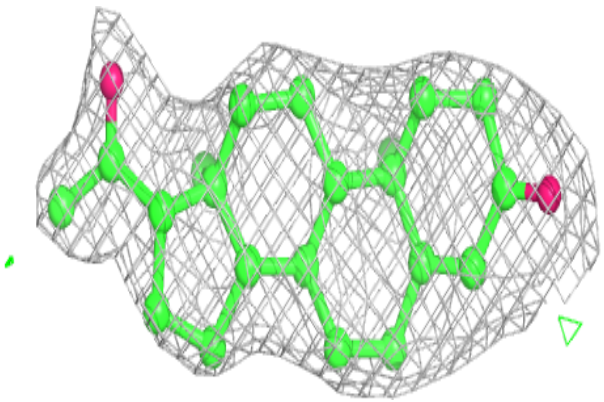


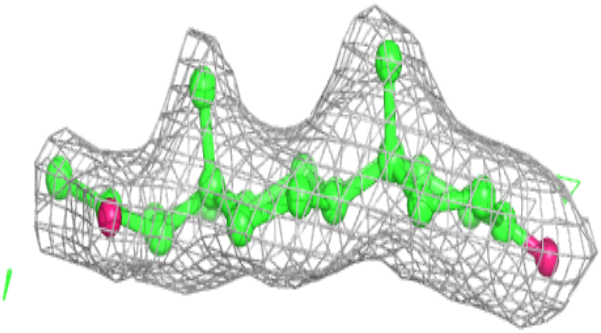


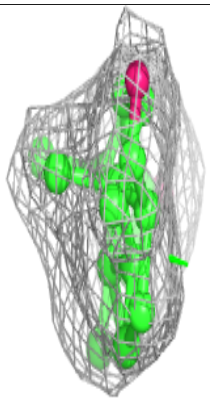


Electron density around STR C 603:

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

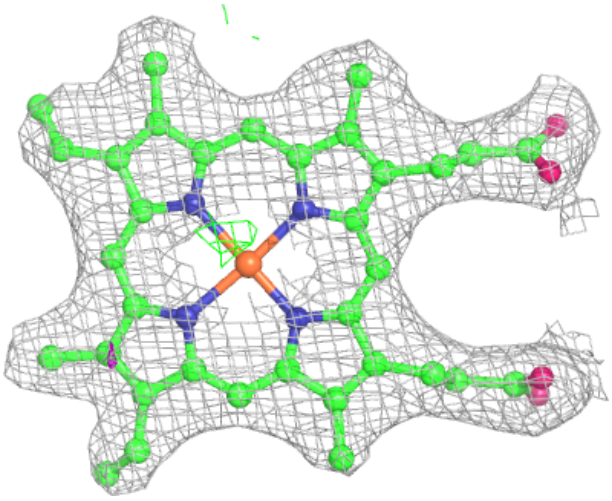


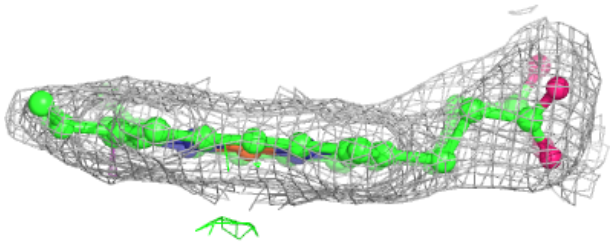


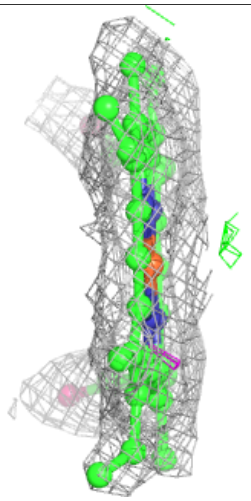


Electron density around HEM B 602:

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

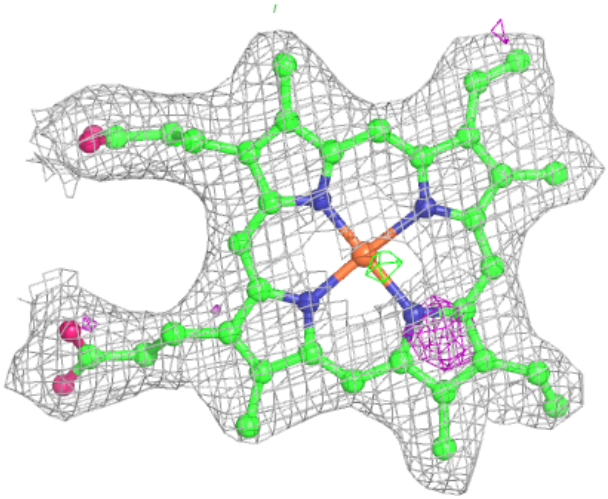


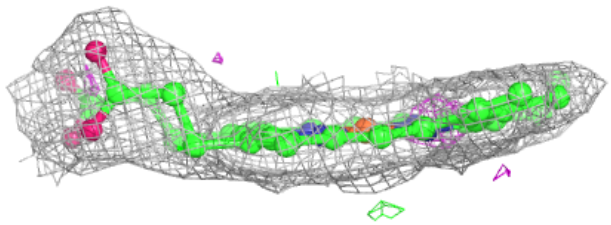


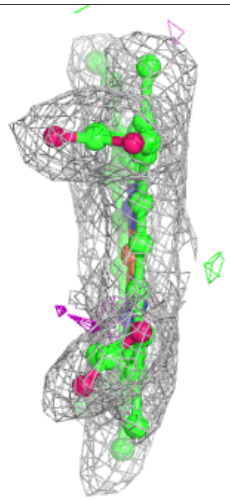


Electron density around HEM A 603:

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

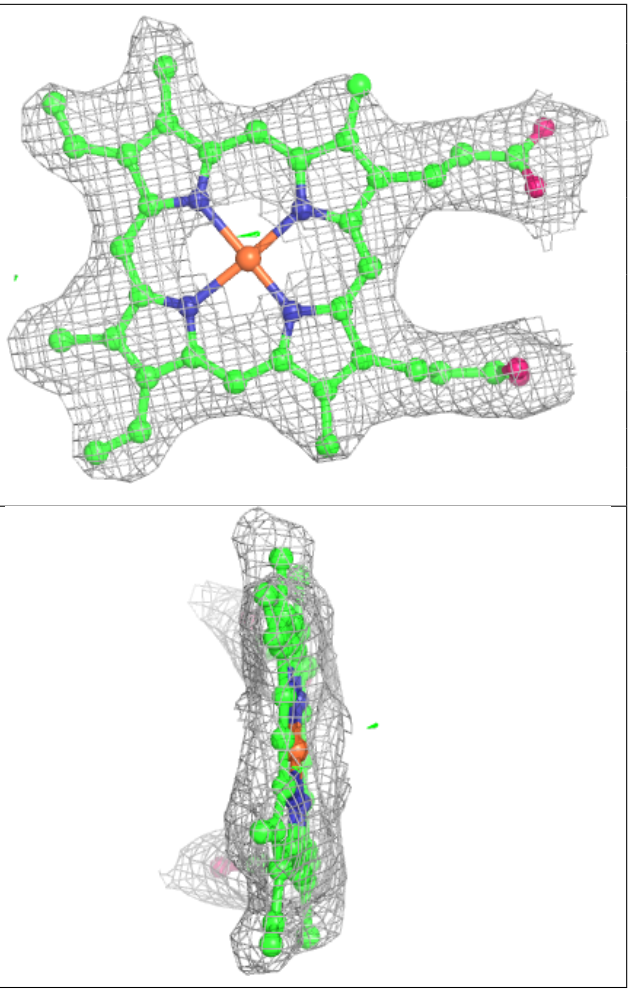






Electron density around HEM C 602:

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.