Full wwPDB X-ray Structure Validation Report (i)

Oct 11, 2023 – 09:47 AM EDT

PDB ID	:	7N3M
Title	:	Co-complex CYP46A1 with 0431 (compound 17)
Authors	:	Lane, W.; Yano, J.
Deposited on	:	2021-06-01
Resolution	:	1.70 Å(reported)

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

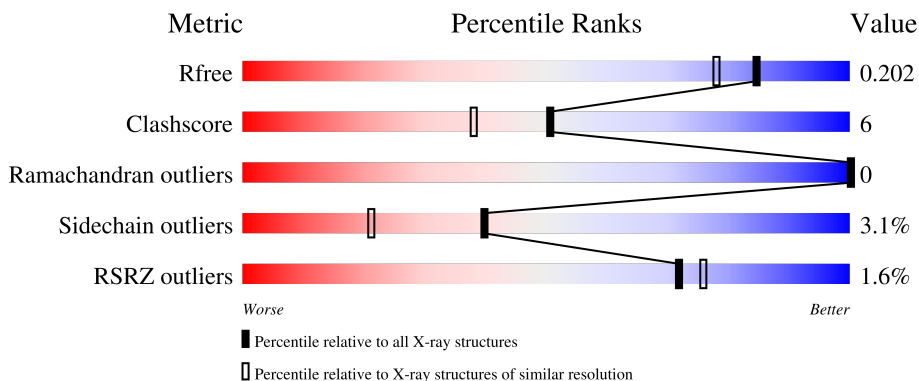|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & 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 1.70 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	Percentile Ranks	Value
Rfree		0.202
Clashscore		6
Ramachandran outliers		0
Sidechain outliers		3.1%
RSRZ outliers		1.6%

Worse *Better*

▮ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	130704	4298 (1.70-1.70)
Clashscore	141614	4695 (1.70-1.70)
Ramachandran outliers	138981	4610 (1.70-1.70)
Sidechain outliers	138945	4610 (1.70-1.70)
RSRZ outliers	127900	4222 (1.70-1.70)

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	AAA	474	% 79% 12% • 8%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 4041 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 Cholesterol 24-hydroxylase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	AAA	437	3585	2296	625	646	18	0	11	0

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	21	MET	-	initiating methionine	UNP Q9Y6A2
AAA	22	HIS	-	expression tag	UNP Q9Y6A2
AAA	23	HIS	-	expression tag	UNP Q9Y6A2
AAA	24	HIS	-	expression tag	UNP Q9Y6A2
AAA	25	HIS	-	expression tag	UNP Q9Y6A2
AAA	26	HIS	-	expression tag	UNP Q9Y6A2
AAA	27	HIS	-	expression tag	UNP Q9Y6A2

- Molecule 2 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (formula: $C_{34}H_{32}FeN_4O_4$).

HEM
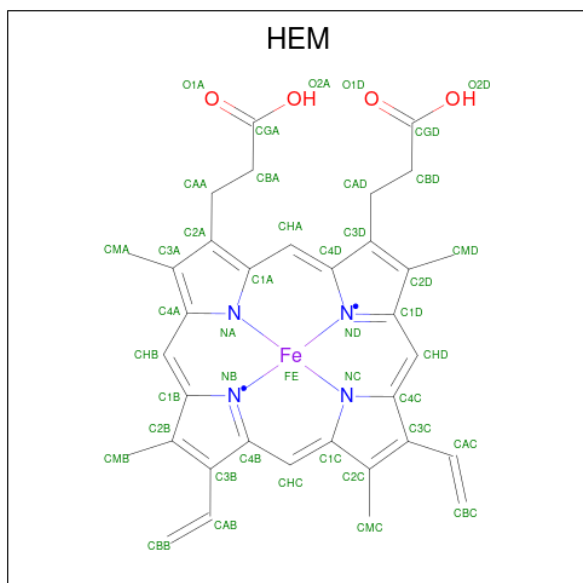
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	Fe	N	O		
2	AAA	1	43	34	1	4	4	0	0

- Molecule 3 is N,N-dimethyl-1-[4-(4-methyl-1H-pyrazol-1-yl)pyridin-3-yl]piperidine-4-carbo xamide (three-letter code: 05D) (formula: $C_{17}H_{23}N_5O$) (labeled as "Ligand of Interest" by depositor).
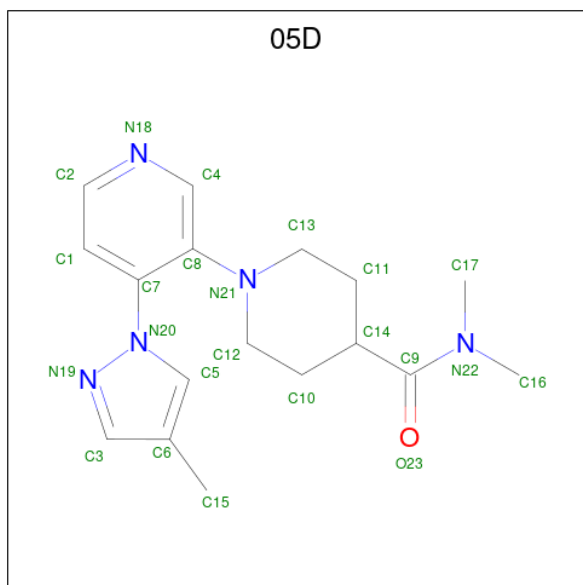
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
3	AAA	1	23	17	5	1	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	AAA	390	390	390	0	0

3 Residue-property plots ⓘ

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: Cholesterol 24-hydroxylase
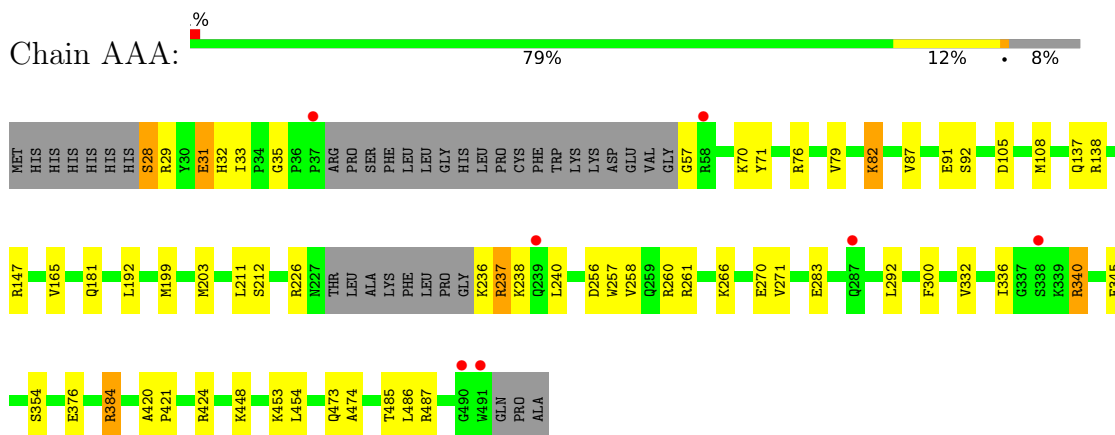
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	58.57Å 63.72Å 125.06Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	44.67 – 1.70 44.63 – 1.70	Depositor EDS
% Data completeness (in resolution range)	98.6 (44.67-1.70) 98.6 (44.63-1.70)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.11	Depositor
$< I/\sigma(I) >$ [1]	2.32 (at 1.70Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.164 , 0.192 0.175 , 0.202	Depositor DCC
R_{free} test set	2532 reflections (4.89%)	wwPDB-VP
Wilson B-factor (Å^2)	18.1	Xtriage
Anisotropy	0.311	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 45.3	EDS
L-test for twinning[2]	$< \lvert L \rvert > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	4041	wwPDB-VP
Average B, all atoms (Å^2)	24.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.52% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1] Intensities estimated from amplitudes.

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

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: HEM, 05D

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	AAA	0.91	4/3690 (0.1%)	0.97	4/4979 (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	AAA	0	1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	AAA	270	GLU	CD-OE1	10.66	1.37	1.25
1	AAA	376	GLU	CD-OE2	6.43	1.32	1.25
1	AAA	91	GLU	CD-OE2	5.84	1.32	1.25
1	AAA	31	GLU	CD-OE1	5.03	1.31	1.25

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	AAA	138	ARG	NE-CZ-NH1	6.50	123.55	120.30
1	AAA	226	ARG	NE-CZ-NH1	5.95	123.27	120.30
1	AAA	29	ARG	NE-CZ-NH1	-5.91	117.35	120.30
1	AAA	487	ARG	NE-CZ-NH1	-5.83	117.38	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	AAA	384	ARG	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	AAA	3585	0	3648	42	0
2	AAA	43	0	30	0	0
3	AAA	23	0	0	0	0
4	AAA	390	0	0	12	0
All	All	4041	0	3678	42	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 (42) 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:AAA:57:GLY:N	4:AAA:603:HOH:O	2.12	0.83
1:AAA:181[B]:GLN:OE1	4:AAA:601:HOH:O	2.02	0.77
1:AAA:236:LYS:N	4:AAA:607:HOH:O	2.20	0.73
1:AAA:256:ASP:OD2	4:AAA:602:HOH:O	2.06	0.72
1:AAA:33:ILE:O	1:AAA:76:ARG:NH1	2.26	0.67
1:AAA:203:MET:CE	1:AAA:212:SER:HB2	2.26	0.66
1:AAA:340:ARG:O	1:AAA:453:LYS:NZ	2.30	0.64
1:AAA:147:ARG:NE	4:AAA:606:HOH:O	2.17	0.63
1:AAA:473[A]:GLN:O	4:AAA:604:HOH:O	2.16	0.63
1:AAA:165[B]:VAL:HG12	1:AAA:454:LEU:HD23	1.82	0.60
1:AAA:332:VAL:HG13	1:AAA:336:ILE:HD11	1.83	0.60
1:AAA:105:ASP:OD2	4:AAA:605:HOH:O	2.17	0.60
1:AAA:70:LYS:HE3	1:AAA:71:TYR:CZ	2.38	0.58
1:AAA:199[B]:MET:HB3	1:AAA:257[B]:TRP:CZ2	2.39	0.58
1:AAA:203:MET:HE3	1:AAA:212:SER:HB2	1.86	0.57
1:AAA:70:LYS:HE3	1:AAA:71:TYR:OH	2.04	0.57
1:AAA:332:VAL:HG13	1:AAA:336:ILE:CD1	2.36	0.55
1:AAA:87[A]:VAL:CG1	1:AAA:92:SER:CB	2.85	0.55
1:AAA:181[B]:GLN:HG3	4:AAA:775:HOH:O	2.07	0.54

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:AAA:87[A]:VAL:CG1	1:AAA:92:SER:HB3	2.38	0.54
1:AAA:87[A]:VAL:HG11	1:AAA:92:SER:HB3	1.90	0.53
1:AAA:35:GLY:N	1:AAA:76:ARG:HH12	2.07	0.53
1:AAA:203:MET:HE1	1:AAA:212:SER:HB2	1.92	0.52
1:AAA:32:HIS:CE1	1:AAA:384:ARG:NH1	2.78	0.51
1:AAA:236:LYS:HB3	1:AAA:237:ARG:NH1	2.26	0.50
1:AAA:199[A]:MET:HG3	1:AAA:257[A]:TRP:CD2	2.47	0.49
1:AAA:211:LEU:HD21	1:AAA:300:PHE:HZ	1.78	0.47
1:AAA:35:GLY:N	1:AAA:76:ARG:NH1	2.63	0.47
1:AAA:87[A]:VAL:HG13	4:AAA:643:HOH:O	2.14	0.47
1:AAA:424:ARG:HA	1:AAA:424:ARG:HD2	1.71	0.46
1:AAA:485:THR:C	1:AAA:486:LEU:HD12	2.37	0.46
1:AAA:354[C]:SER:HG	1:AAA:448:LYS:HZ1	1.58	0.45
1:AAA:261:ARG:CD	1:AAA:271:VAL:HG11	2.47	0.45
1:AAA:82:LYS:HE3	1:AAA:108:MET:SD	2.57	0.45
1:AAA:28:SER:N	4:AAA:631:HOH:O	2.50	0.44
1:AAA:258:VAL:HG11	1:AAA:292:LEU:HD21	2.00	0.43
1:AAA:192:LEU:C	1:AAA:192:LEU:HD13	2.40	0.42
1:AAA:283:GLU:HG3	1:AAA:292:LEU:HB2	2.02	0.41
1:AAA:57:GLY:N	4:AAA:640:HOH:O	2.54	0.41
1:AAA:420:ALA:HA	1:AAA:421:PRO:HD3	1.98	0.41
1:AAA:473[A]:GLN:O	1:AAA:474[A]:ALA:HB3	2.21	0.41
1:AAA:237:ARG:O	4:AAA:608:HOH:O	2.22	0.41

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	AAA	443/474 (94%)	435 (98%)	8 (2%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains ⓘ

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	AAA	393/414 (95%)	381 (97%)	12 (3%)	40	21

All (12) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	AAA	28	SER
1	AAA	31	GLU
1	AAA	79	VAL
1	AAA	82	LYS
1	AAA	137	GLN
1	AAA	237	ARG
1	AAA	238	LYS
1	AAA	240	LEU
1	AAA	260	ARG
1	AAA	266	LYS
1	AAA	340	ARG
1	AAA	345	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

2 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	05D	AAA	502	2	23,25,25	3.53	6 (26%)	31,35,35	2.34	11 (35%)
2	HEM	AAA	501	3,1	41,50,50	1.40	5 (12%)	45,82,82	1.52	7 (15%)

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	05D	AAA	502	2	-	0/14/26/26	0/3/3/3
2	HEM	AAA	501	3,1	-	2/12/54/54	-

All (11) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	AAA	502	05D	C9-N22	15.07	1.56	1.34
3	AAA	502	05D	C16-N22	-4.39	1.31	1.46
3	AAA	502	05D	C12-N21	4.12	1.53	1.46
2	AAA	501	HEM	C1B-NB	-3.70	1.33	1.40
2	AAA	501	HEM	C1D-C2D	2.79	1.50	1.44
2	AAA	501	HEM	CBD-CGD	2.68	1.56	1.50
2	AAA	501	HEM	O2D-CGD	-2.54	1.22	1.30
2	AAA	501	HEM	CHB-C1B	2.28	1.40	1.35
3	AAA	502	05D	C8-N21	2.24	1.46	1.41
3	AAA	502	05D	C2-N18	2.13	1.40	1.33
3	AAA	502	05D	C7-C8	2.05	1.41	1.39

All (18) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	AAA	502	05D	O23-C9-N22	-5.63	114.63	121.92
3	AAA	502	05D	C2-N18-C4	-5.52	107.29	116.85
3	AAA	502	05D	C1-C7-C8	-4.25	116.68	120.00
2	AAA	501	HEM	CBA-CAA-C2A	-4.01	105.77	112.62
3	AAA	502	05D	O23-C9-C14	3.93	126.81	120.81
3	AAA	502	05D	C8-C4-N18	3.79	130.78	122.32
2	AAA	501	HEM	C4B-C3B-C2B	-3.61	104.25	107.11
2	AAA	501	HEM	CHD-C1D-ND	3.19	127.89	124.43
2	AAA	501	HEM	CHD-C1D-C2D	-3.10	120.14	124.98
3	AAA	502	05D	C7-C8-N21	-3.07	118.61	123.87
3	AAA	502	05D	C16-N22-C17	2.98	125.00	115.77
3	AAA	502	05D	C1-C2-N18	2.82	128.52	123.62
3	AAA	502	05D	C5-N20-N19	-2.60	110.08	112.72
2	AAA	501	HEM	C3B-C2B-C1B	2.48	108.33	106.49
3	AAA	502	05D	C12-C10-C14	2.39	114.39	110.41
2	AAA	501	HEM	C4B-CHC-C1C	2.31	125.60	122.56
2	AAA	501	HEM	CBD-CAD-C3D	-2.28	106.31	112.63
3	AAA	502	05D	C16-N22-C9	-2.09	114.48	121.89

There are no chirality outliers.

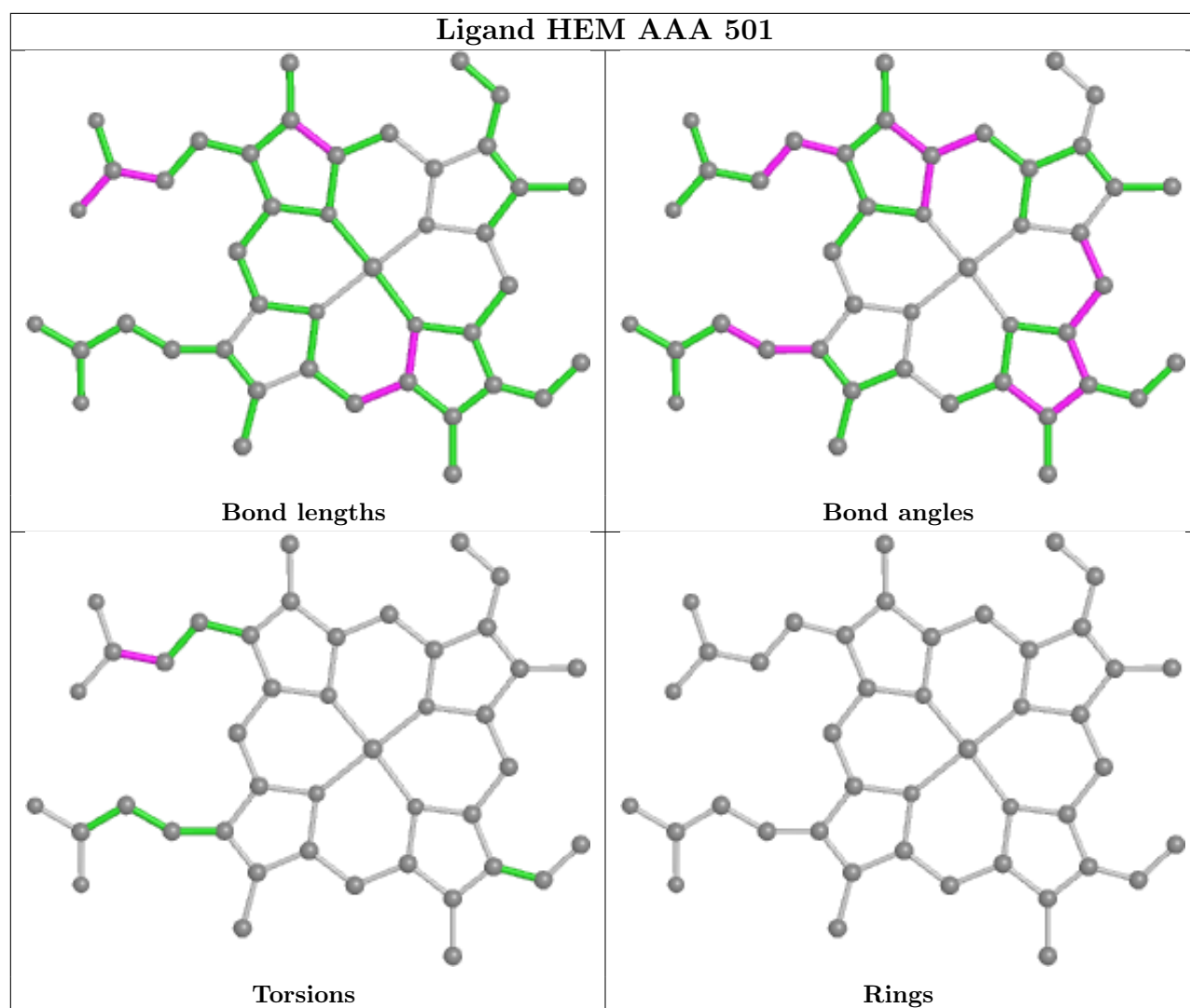All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	AAA	501	HEM	CAD-CBD-CGD-O1D
2	AAA	501	HEM	CAD-CBD-CGD-O2D

There are no ring outliers.

No monomer is involved in short contacts.

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 05D AAA 502

Bond lengths

Bond angles

Torsions

Rings

Ligand HEM AAA 501

Bond lengths

Bond angles

Torsions

Rings
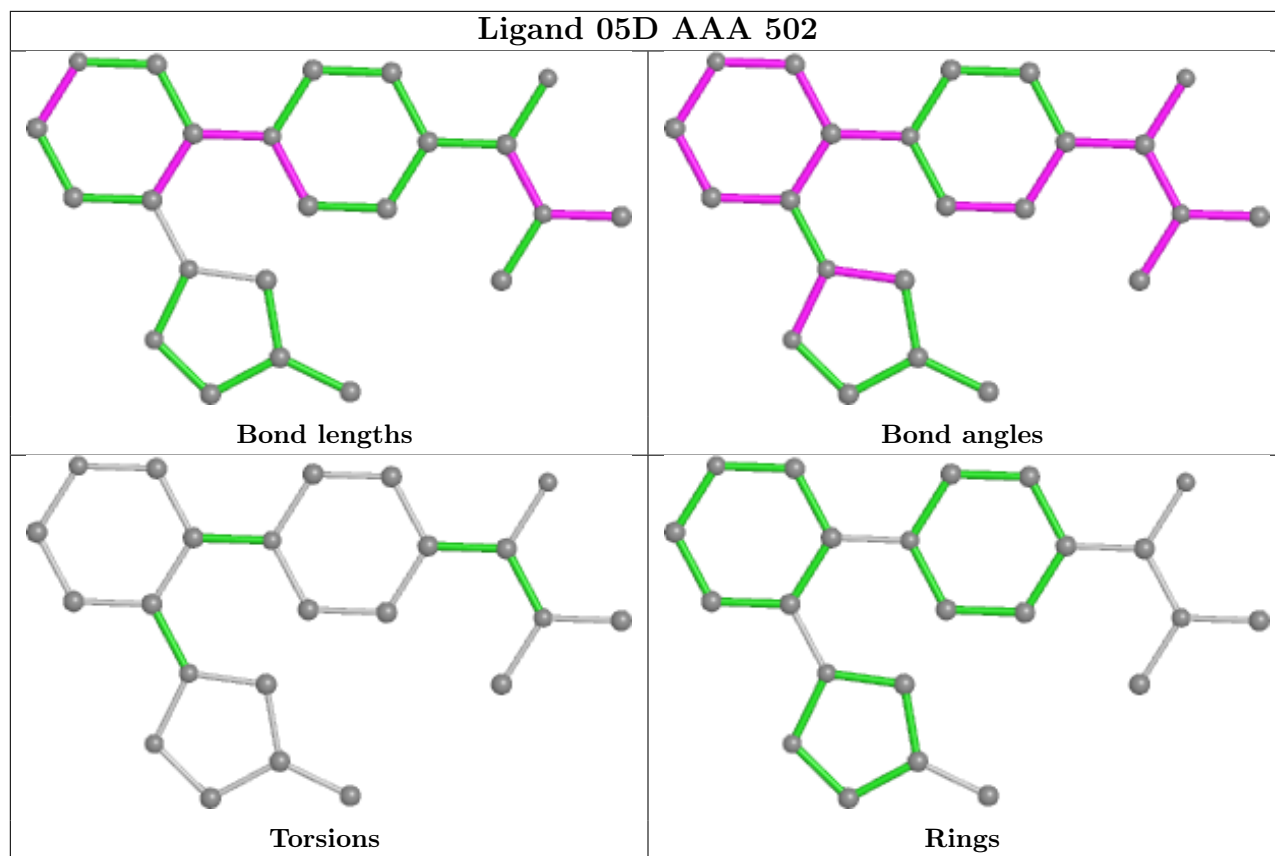
5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 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	AAA	437/474 (92%)	-0.20	7 (1%)	72	76	10, 20, 44, 68	1 (0%)

All (7) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	AAA	491	TRP	6.3
1	AAA	490	GLY	6.3
1	AAA	338	SER	2.8
1	AAA	287	GLN	2.6
1	AAA	37	PRO	2.6
1	AAA	239	GLN	2.4
1	AAA	58	ARG	2.4

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

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

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

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.
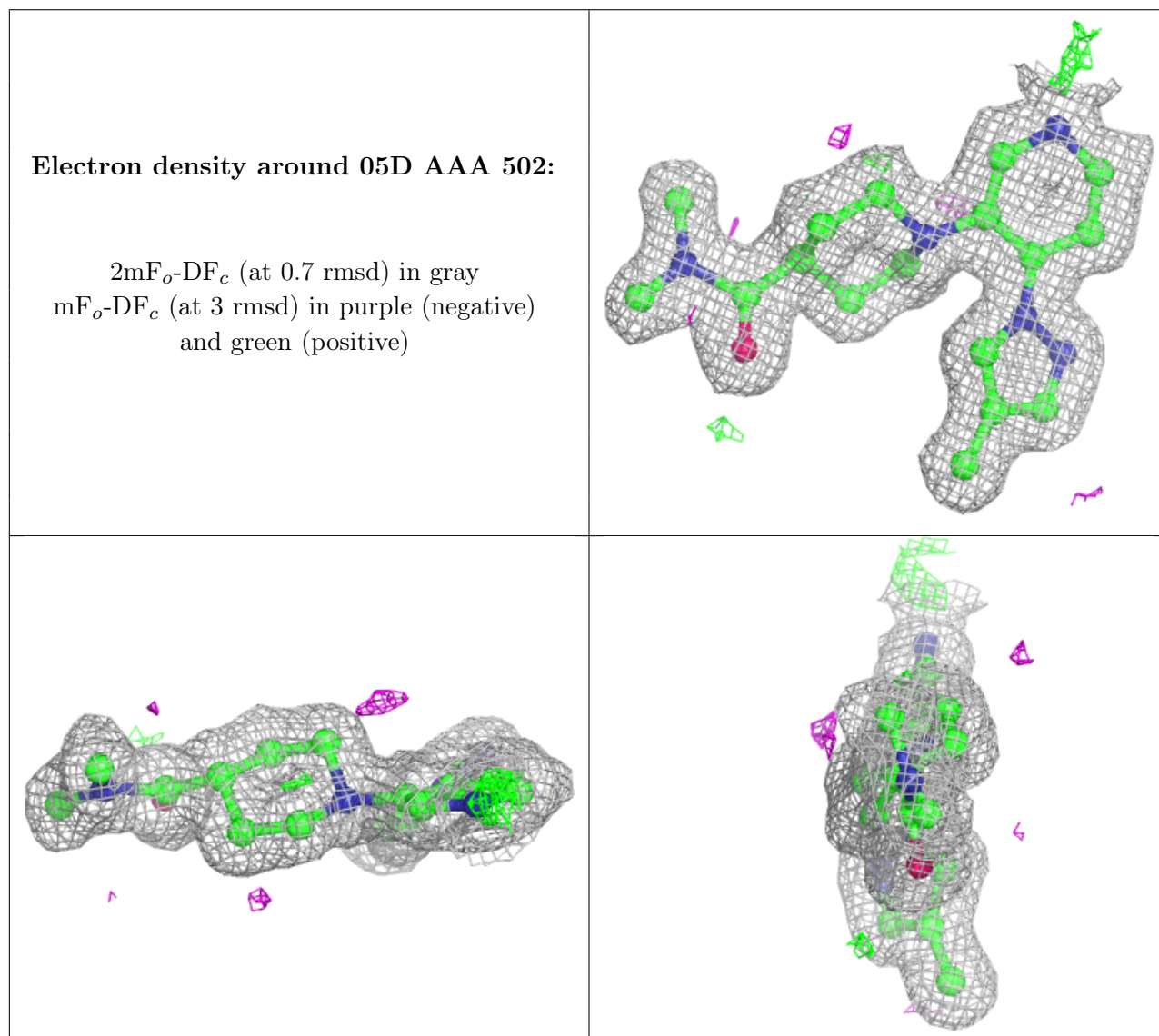
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	05D	AAA	502	23/23	0.97	0.07	9,11,14,16	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	HEM	AAA	501	43/43	0.99	0.07	9,10,13,15	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 05D AAA 502:

$2mF_o$-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 AAA 501:

$2mF_o$-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 ⓘ

There are no such residues in this entry.