



wwPDB X-ray Structure Validation Summary Report ⓘ

May 14, 2020 – 04:14 pm BST

PDB ID : 2CA0
Title : Crystal structure of YC-17-bound cytochrome P450 PikC (CYP107L1)
Authors : Yermalitskaya, L.I.; Kim, Y.; Sherman, D.H.; Waterman, M.R.; Podust, L.M.
Deposited on : 2005-12-15
Resolution : 2.85 Å(reported)

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

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) 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 : 2.11
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) : 2.11

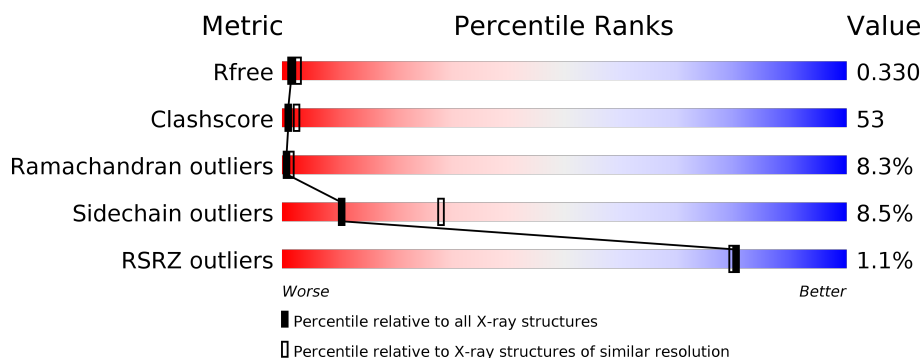
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.85 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	3168 (2.90-2.82)
Clashscore	141614	3438 (2.90-2.82)
Ramachandran outliers	138981	3348 (2.90-2.82)
Sidechain outliers	138945	3351 (2.90-2.82)
RSRZ outliers	127900	3103 (2.90-2.82)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. 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	436	<div> <div> <div></div> <div>29%</div> <div>52%</div> <div>8%</div> <div>10%</div> </div> </div>
1	B	436	<div> <div> <div></div> <div>24%</div> <div>55%</div> <div>10%</div> <div>10%</div> </div> </div>

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 6269 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 MONOOXYGENASE.

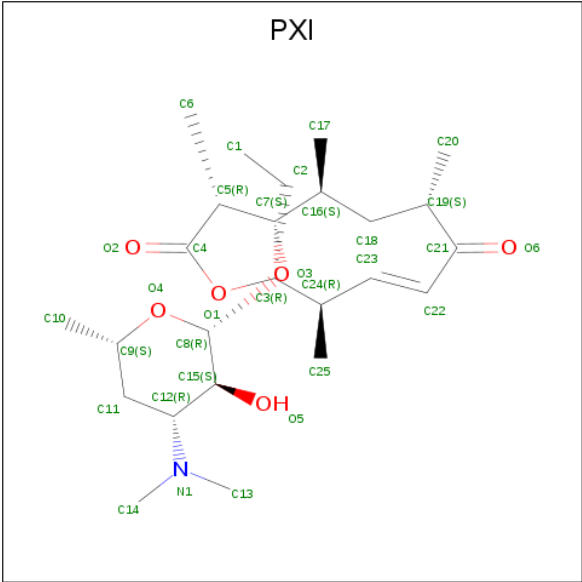
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	393	Total	C	N	O	S	0	0	0
			3028	1910	544	561	13			
1	B	393	Total	C	N	O	S	0	0	0
			3028	1910	544	561	13			

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



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

- Molecule 3 is 4-[[4-(DIMETHYLAMINO)-3-HYDROXY-6-METHYLTETRAHYDRO-2H-PYRAN-2-YL]OXY]-12-ETHYL-3,5,7,11-TETRAMETHYLOXACYCLODODEC-9-ENE-2,8-DIONE (three-letter code: PXI) (formula: $C_{25}H_{43}NO_6$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			32	25	1	6		
3	B	1	Total	C	N	O	0	0
			32	25	1	6		

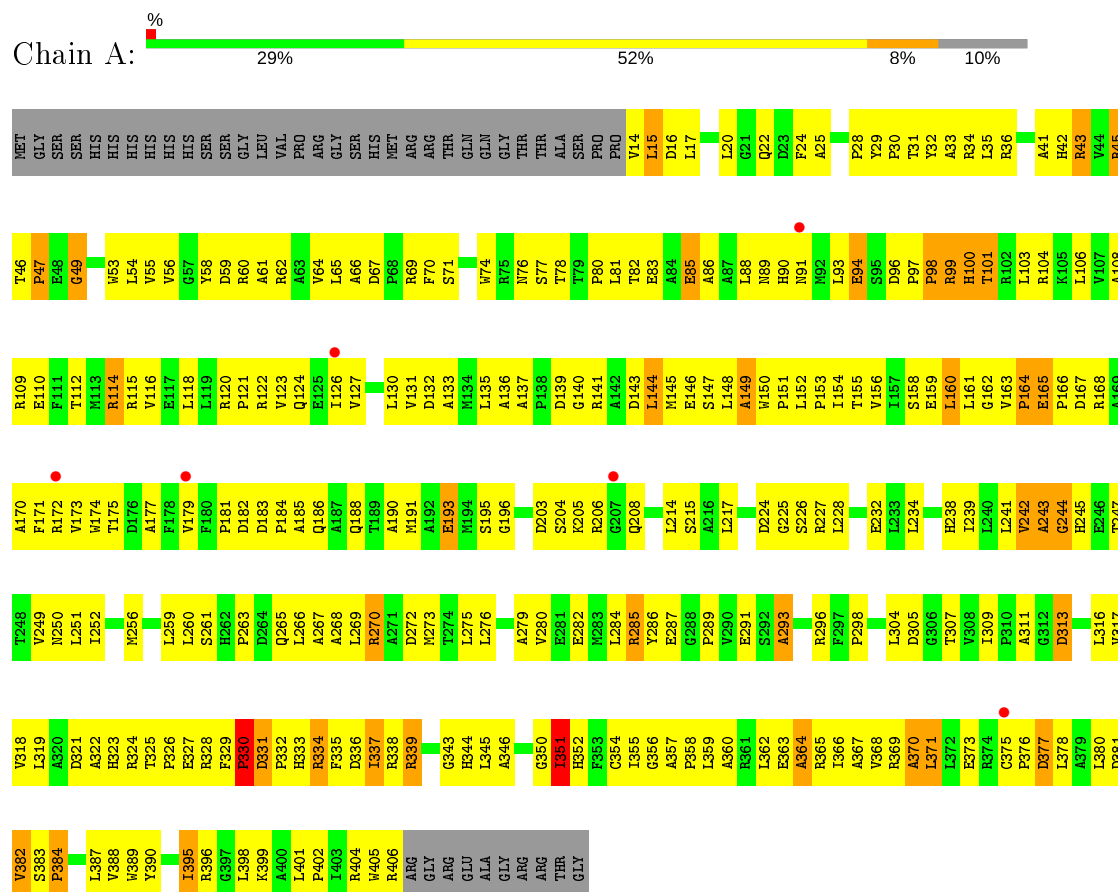
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	33	Total	O	0	0
			33	33		
4	B	30	Total	O	0	0
			30	30		

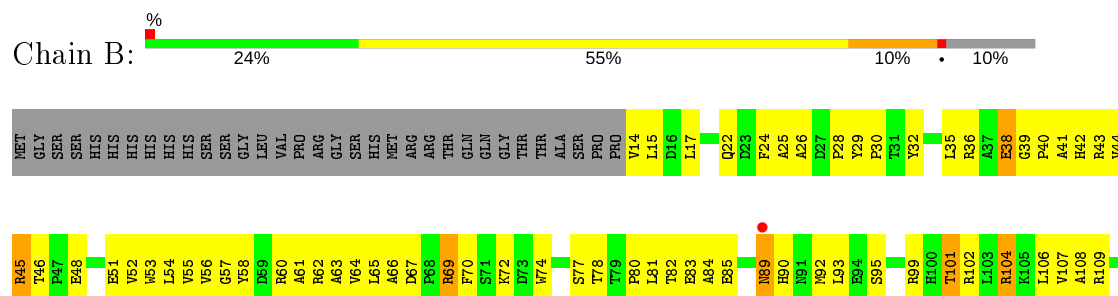
3 Residue-property plots [i](#)

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



• Molecule 1: CYTOCHROME P450 MONOOXYGENASE



R374	C375	P376	D377	L378	A379	L380	D381	V382	S383	P384	G385	E386	L387	V388	W389	Y390	P391	N392	P393	P394	I395	R396	G397	L398	K399	A400	L401	P402	T403	R404	W405	R406	ARG	GLY	ARG	ARG	GLU	ALA	GLY	ARG	ARG	THR	GLY															
T247	T248	V249	N250	L251	I252	A253	N254	G255	N256	D321	A322	H323	R324	T325	P326	E327	D284	Q285	L286	D331	P332	H333	R334	F335	D336	I337	R338	R339	D340	T341	A342	G343	H344	L345	A346	F347	G348	H349	C354	I355	G356	A357	P358	L359	A360	R361	L362	E363	A364	R365	I366	A367	V368	R369	A370	L371	L372	E373
T248	W174	T175	D176	A177	F178	V179	F180	D181	D182	D183	P184	A185	Q186	A187	Q188	T189	Q284	Q285	L286	A287	A288	L289	R270	A271	D272	M273	T274	L275	L276	D277	G278	E281	E282	N283	L284	R285	Y286	E287	G288	P289	V290	E291	S292	A293	T294	Y295	R296	F297	P298	V299	E300	D303	L304	D305	G306	T307	V308	I309
M112	M113	R114	R115	V116	E117	L118	L119	R120	P121	R122	V123	Q124	E125	I126	V127	D128	G129	L130	V131	D132	A133	M134	L135	A136	A137	P138	D139	G140	R141	A142	D143	L144	M145	E146	S147	L148	A149	W150	P151	L152	P153	I154	T155	V156	S157	S158	L161	G162	V163	P164	E165	P166	D167	R168	A169	A170	F171	R172

4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	61.09Å 91.97Å 68.90Å 90.00° 89.92° 90.00°	Depositor
Resolution (Å)	32.41 – 2.85 45.74 – 2.85	Depositor EDS
% Data completeness (in resolution range)	92.5 (32.41-2.85) 92.2 (45.74-2.85)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.51 (at 2.86Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.240 , 0.338 0.234 , 0.330	Depositor DCC
R_{free} test set	1675 reflections (10.05%)	wwPDB-VP
Wilson B-factor (Å ²)	31.9	Xtriage
Anisotropy	0.519	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 16.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.40$, $\langle L^2 \rangle = 0.22$	Xtriage
Estimated twinning fraction	0.427 for h,-k,-l	Xtriage
F_o, F_c correlation	0.89	EDS
Total number of atoms	6269	wwPDB-VP
Average B, all atoms (Å ²)	23.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.78% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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, PXI

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	0.43	0/3098	0.68	0/4229
1	B	0.41	0/3098	0.71	0/4229
All	All	0.42	0/6196	0.70	0/8458

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3028	0	2986	314	0
1	B	3028	0	2986	342	0
2	A	43	0	30	3	0
2	B	43	0	30	5	0
3	A	32	0	43	1	0
3	B	32	0	43	7	0
4	A	33	0	0	9	0
4	B	30	0	0	9	0
All	All	6269	0	6118	656	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 53.

The worst 5 of 656 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:B:55:VAL:HG11	1:B:64:VAL:HG21	1.41	1.03
1:B:285:ARG:HH11	1:B:285:ARG:HB3	1.25	1.01
1:A:285:ARG:HH21	1:A:286:TYR:HA	1.26	1.00
1:A:152:LEU:HB3	1:A:153:PRO:HD3	1.48	0.95
1:A:45:ARG:HH21	1:A:49:GLY:H	1.02	0.92

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	391/436 (90%)	285 (73%)	77 (20%)	29 (7%)	1	2
1	B	391/436 (90%)	277 (71%)	78 (20%)	36 (9%)	1	1
All	All	782/872 (90%)	562 (72%)	155 (20%)	65 (8%)	1	1

5 of 65 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	98	PRO
1	A	99	ARG
1	A	149	ALA
1	A	242	VAL
1	A	243	ALA

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	313/355 (88%)	288 (92%)	25 (8%)	12	31
1	B	313/355 (88%)	285 (91%)	28 (9%)	9	26
All	All	626/710 (88%)	573 (92%)	53 (8%)	10	28

5 of 53 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	384	PRO
1	B	114	ARG
1	B	334	ARG
1	A	395	ILE
1	B	89	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	344	HIS
1	B	89	ASN
1	B	265	GLN
1	A	238	HIS
1	B	254	ASN

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 carbohydrates in this entry.

5.6 Ligand geometry

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	PXI	B	1408	-	33,33,33	1.94	6 (18%)	40,47,47	2.04	9 (22%)
2	HEM	A	1407	1	27,50,50	1.91	9 (33%)	17,82,82	1.95	5 (29%)
3	PXI	A	1408	-	33,33,33	2.07	7 (21%)	40,47,47	1.85	10 (25%)
2	HEM	B	1407	1	27,50,50	1.99	9 (33%)	17,82,82	2.17	3 (17%)

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	PXI	B	1408	-	-	19/43/59/59	0/1/2/2
2	HEM	A	1407	1	-	4/6/54/54	-
3	PXI	A	1408	-	-	14/43/59/59	0/1/2/2
2	HEM	B	1407	1	-	2/6/54/54	-

The worst 5 of 31 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	1408	PXI	C16-C7	6.32	1.64	1.53
3	A	1408	PXI	C19-C21	5.87	1.61	1.51
3	A	1408	PXI	C16-C7	5.44	1.63	1.53
2	B	1407	HEM	C3B-CAB	-5.07	1.37	1.47
2	A	1407	HEM	C3B-CAB	-4.47	1.38	1.47

The worst 5 of 27 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	1408	PXI	C3-O1-C4	7.97	129.85	117.89

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
3	A	1408	PXI	C3-O1-C4	7.11	128.56	117.89
2	B	1407	HEM	CBD-CAD-C3D	6.76	124.94	112.48
3	B	1408	PXI	O3-C7-C5	-4.96	102.51	111.14
2	A	1407	HEM	C4A-C3A-C2A	4.30	109.98	107.00

There are no chirality outliers.

5 of 39 torsion outliers are listed below:

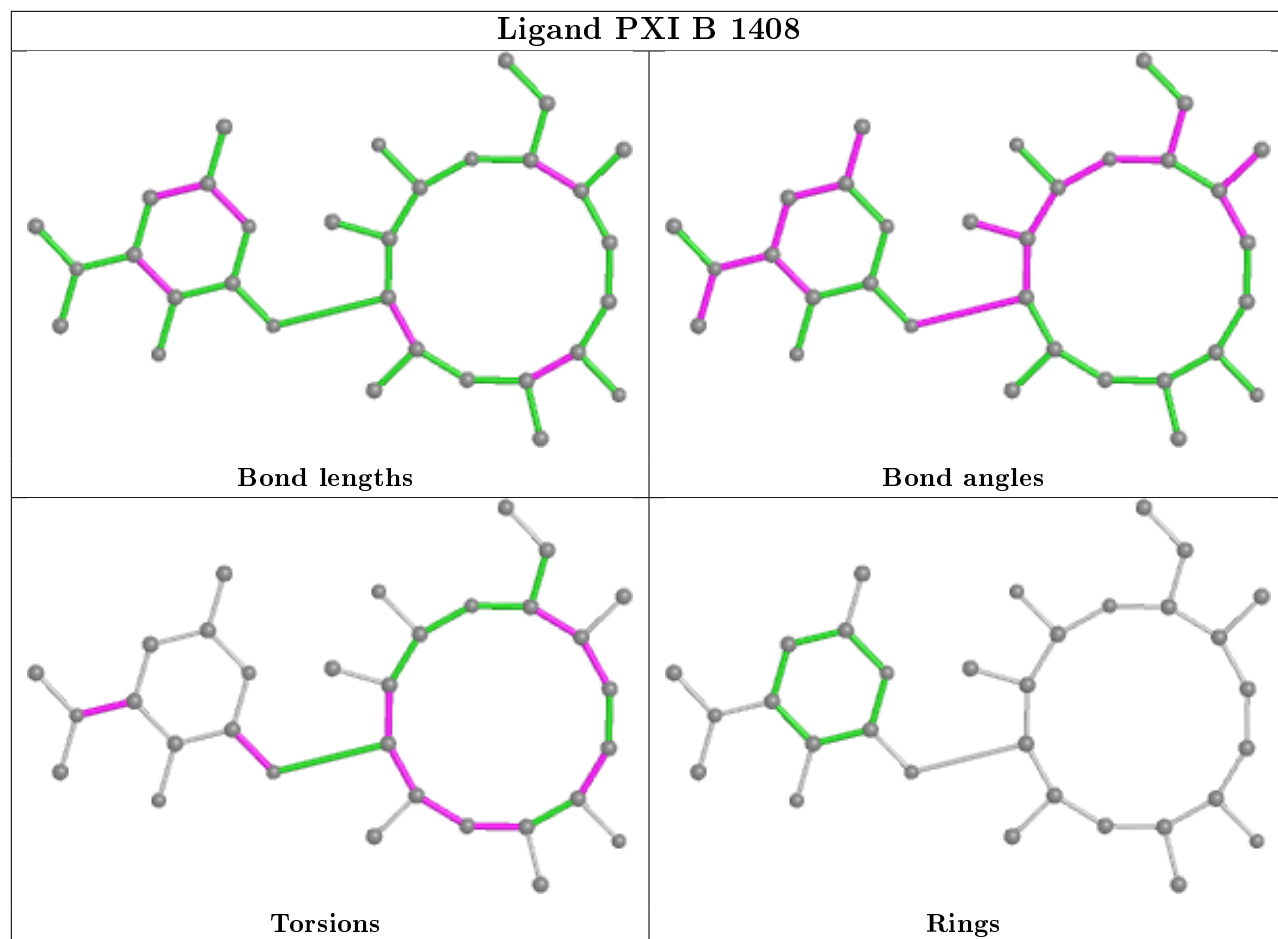
Mol	Chain	Res	Type	Atoms
3	B	1408	PXI	C23-C24-C3-O1
3	B	1408	PXI	C18-C16-C7-O3
3	B	1408	PXI	C11-C12-N1-C13
3	B	1408	PXI	C11-C12-N1-C14
3	B	1408	PXI	C7-C16-C18-C19

There are no ring outliers.

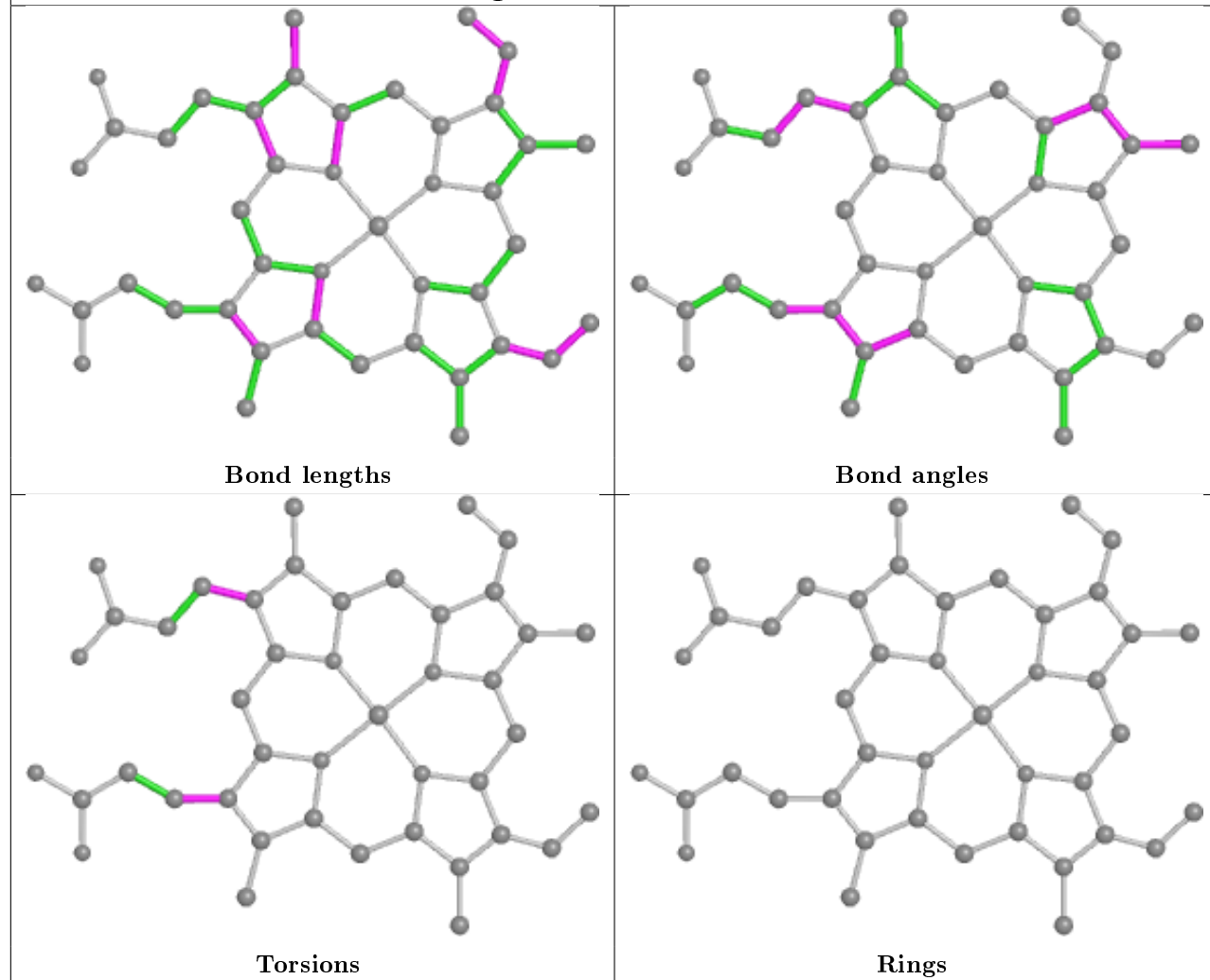
4 monomers are involved in 16 short contacts:

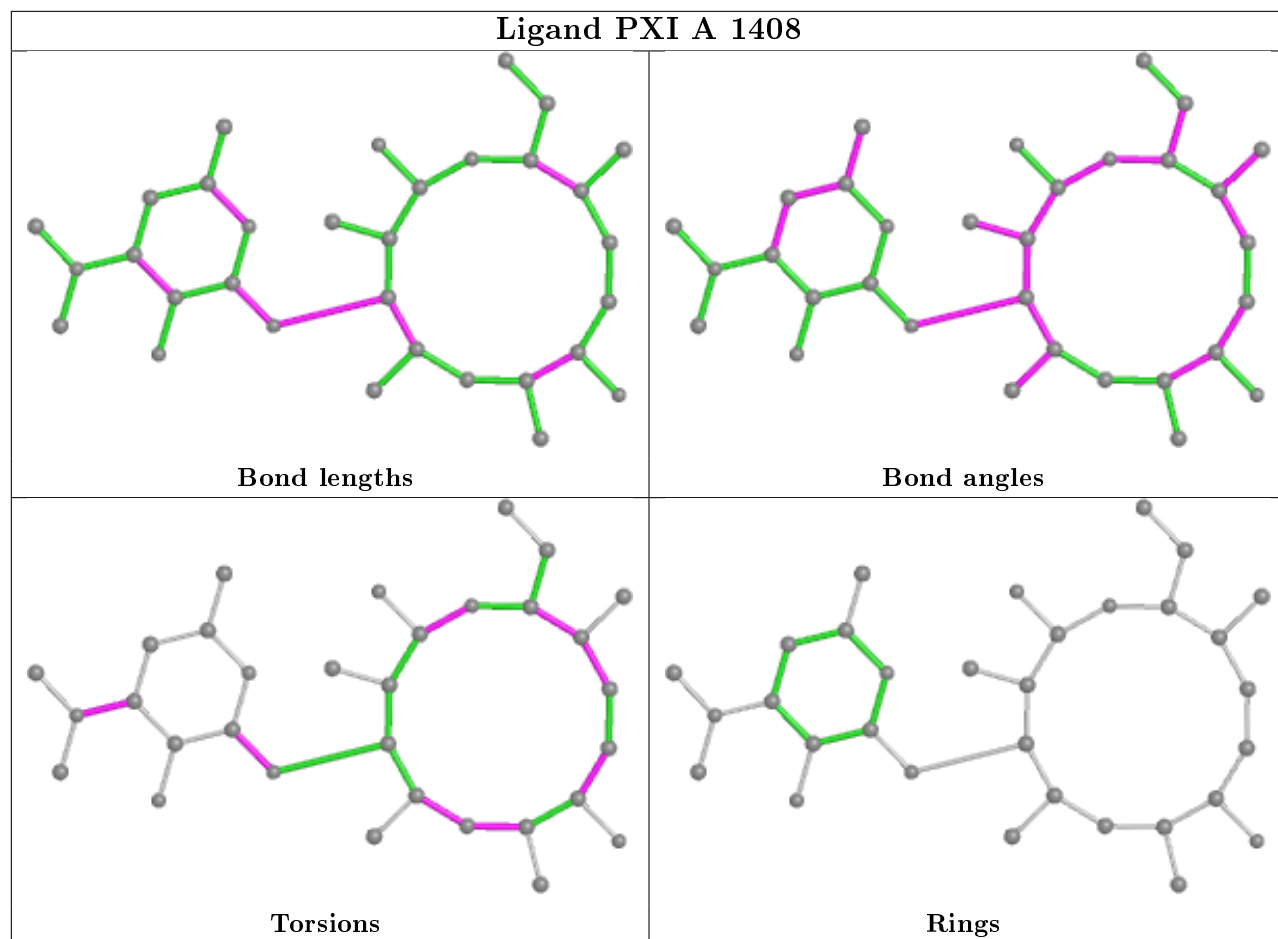
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	1408	PXI	7	0
2	A	1407	HEM	3	0
3	A	1408	PXI	1	0
2	B	1407	HEM	5	0

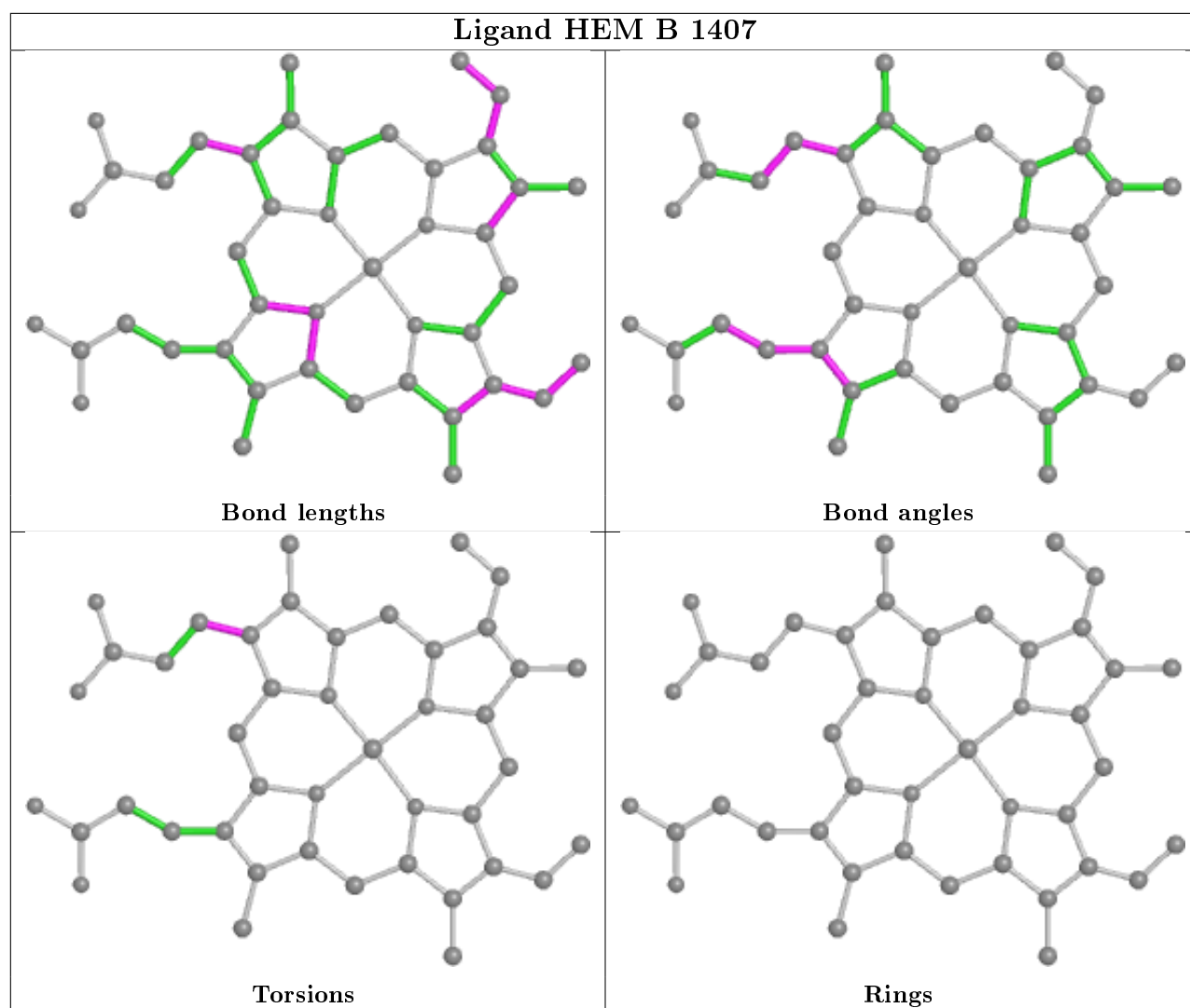
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 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 A 1407







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	393/436 (90%)	0.32	6 (1%) 73 72	4, 23, 37, 48	0
1	B	393/436 (90%)	0.34	3 (0%) 86 85	5, 22, 39, 46	0
All	All	786/872 (90%)	0.33	9 (1%) 80 80	4, 23, 38, 48	0

The worst 5 of 9 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	126	ILE	3.2
1	B	89	ASN	3.0
1	B	384	PRO	2.5
1	A	172	ARG	2.4
1	A	207	GLY	2.4

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](#)

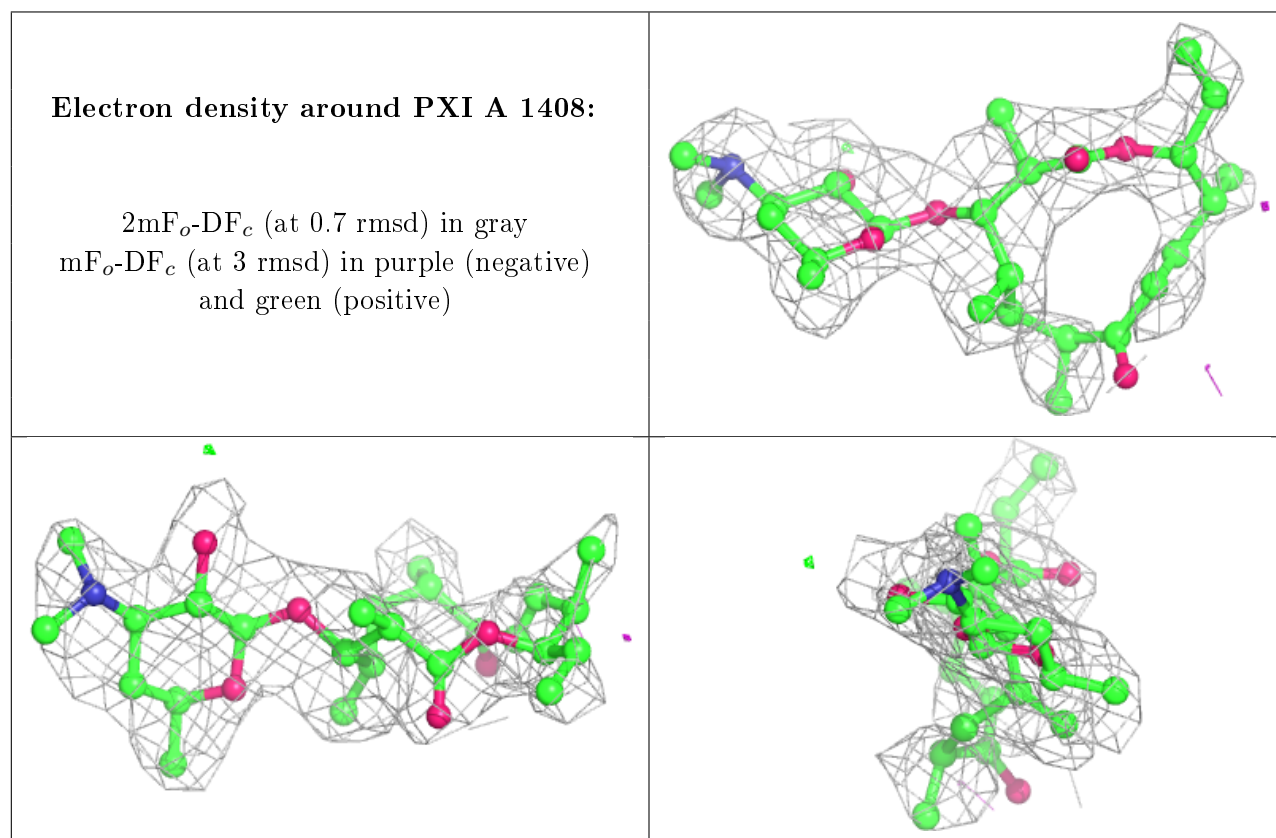
There are no carbohydrates in this entry.

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.

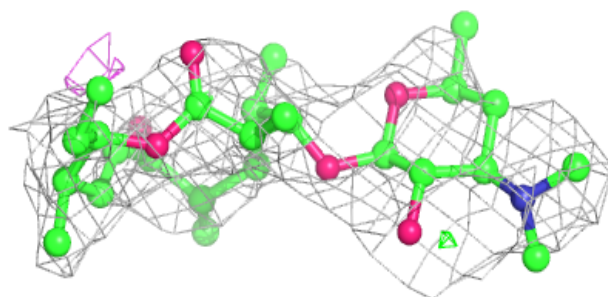
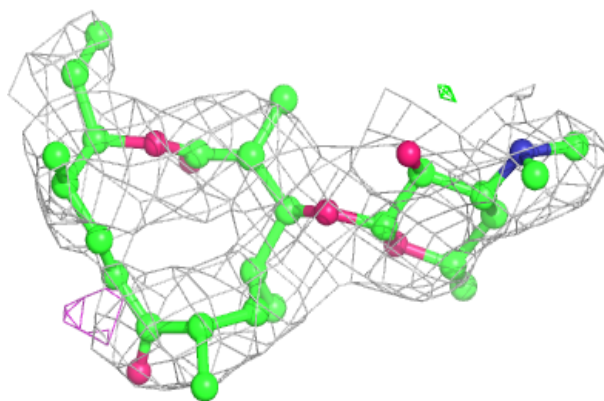
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
3	PXI	A	1408	32/32	0.82	0.29	27,31,32,35	0
3	PXI	B	1408	32/32	0.89	0.28	28,31,33,34	0
2	HEM	A	1407	43/43	0.94	0.20	3,4,12,15	0
2	HEM	B	1407	43/43	0.94	0.19	3,6,12,19	0

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.



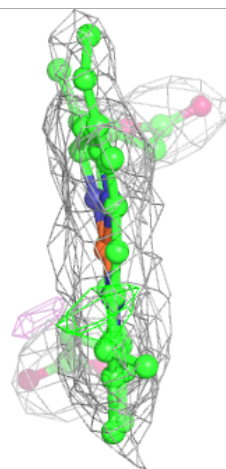
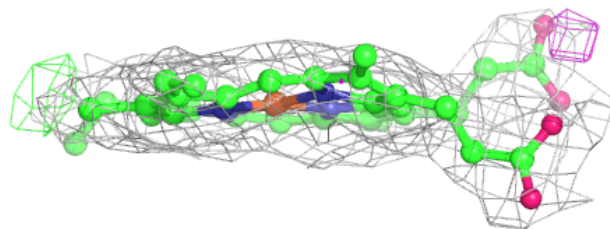
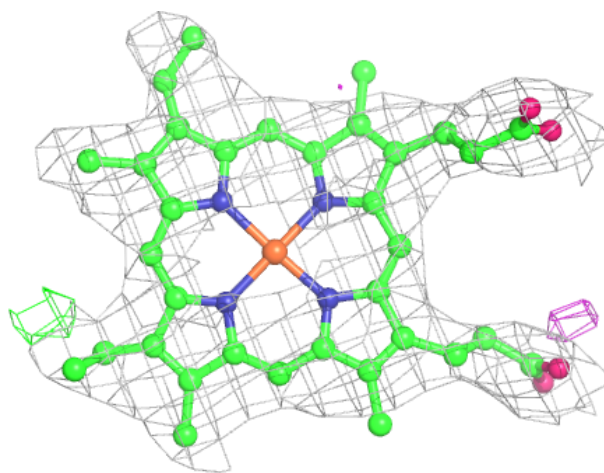
Electron density around PXI B 1408:

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



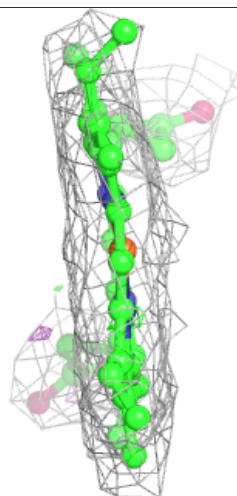
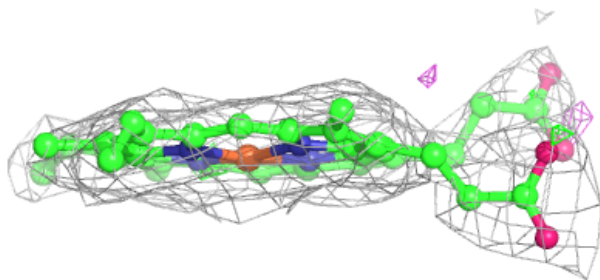
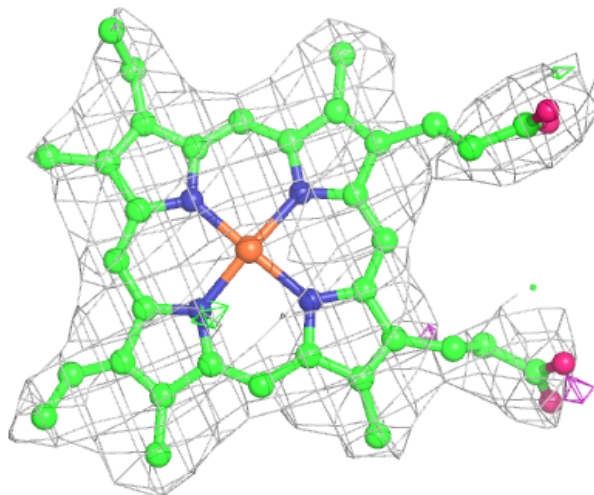
Electron density around HEM A 1407:

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



Electron density around HEM B 1407:

$2mF_o-DF_c$ (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.