W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report ⓘ

Jul 9, 2025 – 04:32 PM JST

PDB ID	:	9KPU / pdb_00009kpu
Title	:	Crystal structure of FrazP2 in complex with forazoline C
Authors	:	Zhang, F.; Chen, X.R.
Deposited on	:	2024-11-24
Resolution	:	2.30 Å(reported)

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

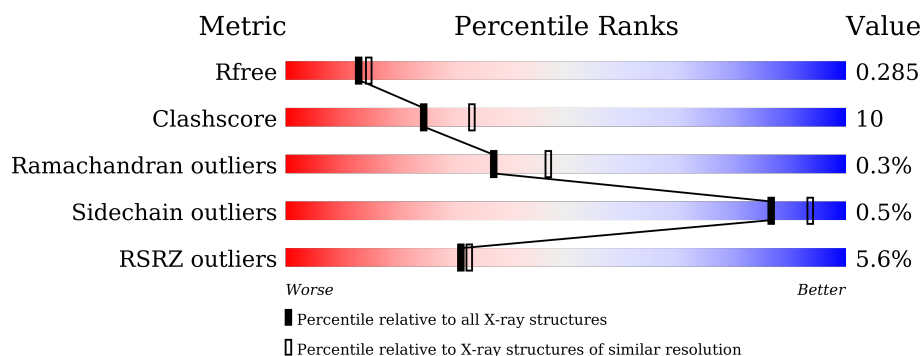|		
MolProbity	:	4-5-2 with Phenix2.0rc1
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	2.0rc1
EDS	:	3.0
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.006 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.44

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 2.30 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		0.285
Clashscore		10
Ramachandran outliers		0.3%
Sidechain outliers		0.5%
RSRZ outliers		5.6%

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}	164625	5963 (2.30-2.30)
Clashscore	180529	6698 (2.30-2.30)
Ramachandran outliers	177936	6640 (2.30-2.30)
Sidechain outliers	177891	6640 (2.30-2.30)
RSRZ outliers	164620	5963 (2.30-2.30)

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	390	6% / 80% / 17% / · ·
1	B	390	5% / 83% / 16% / ·

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 5920 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.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	380	Total	C	N	O	S	0	0	0
			2857	1808	498	537	14			
1	B	386	Total	C	N	O	S	0	0	0
			2878	1825	503	536	14			

- Molecule 2 is PROTOPORPHYRIN IX CONTAINING FE (CCD ID: HEM) (formula: $C_{34}H_{32}FeN_4O_4$).
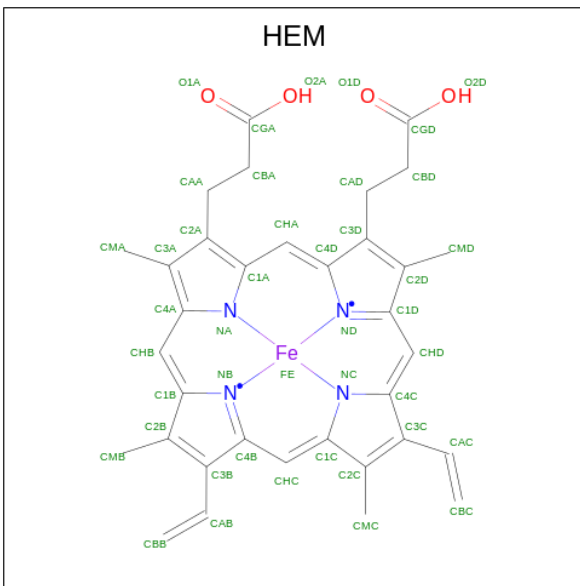
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
2	B	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		

- Molecule 3 is forazoline C (CCD ID: A1EGF) (formula: $C_{26}H_{36}N_2O_5S_3$).

A1EGF
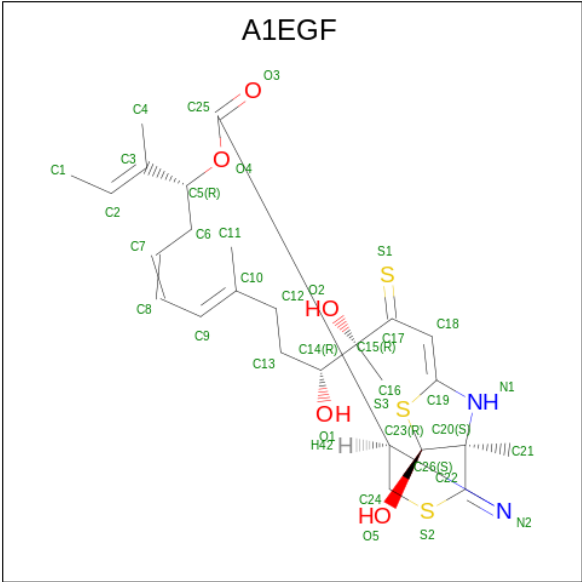
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
3	B	1	36	26	2	5	3	0	0

- Molecule 4 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	25	Total 25	O 25	0	0
4	B	38	Total 38	O 38	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: Cytochrome P450

Chain A:

• Molecule 1: Cytochrome P450

Chain B:
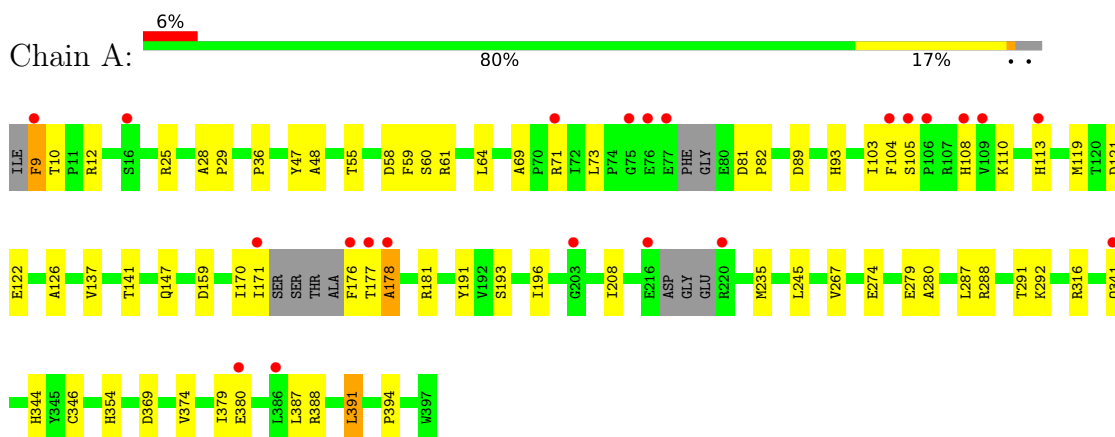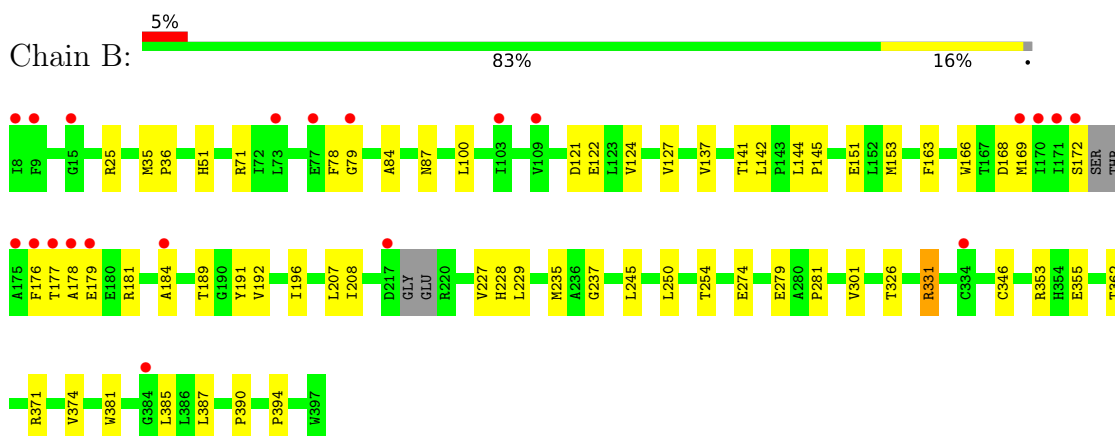
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	54.12Å 74.11Å 103.19Å 90.00° 98.13° 90.00°	Depositor
Resolution (Å)	43.42 – 2.30 43.42 – 2.30	Depositor EDS
% Data completeness (in resolution range)	97.0 (43.42-2.30) 97.0 (43.42-2.30)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.87 (at 2.29Å)	Xtriage
Refinement program	REFMAC 5.8.0430	Depositor
R, R_{free}	0.215 , 0.276 0.230 , 0.285	Depositor DCC
R_{free} test set	1753 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	42.9	Xtriage
Anisotropy	0.148	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 39.8	EDS
L-test for twinning [2]	$< |L| > = 0.54, < L^2 > = 0.37$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	5920	wwPDB-VP
Average B, all atoms (Å^2)	50.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 45.26 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.3562e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

---

[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 ⓘ

5.1   Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: A1EGF, HEM

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	0.68	1/2921 (0.0%)	1.23	6/3994 (0.2%)
1	B	0.68	0/2944	1.26	7/4030 (0.2%)
All	All	0.68	1/5865 (0.0%)	1.25	13/8024 (0.2%)

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	2
1	B	0	5
All	All	0	7

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	291	THR	C-N	-6.54	1.24	1.33

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	179	GLU	N-CA-C	-9.00	102.22	113.02
1	A	47	TYR	CA-C-O	-7.01	113.31	121.16
1	B	127	VAL	N-CA-CB	6.79	118.49	110.55
1	A	104	PHE	N-CA-C	-6.41	99.20	109.39
1	B	355	GLU	N-CA-CB	6.29	119.13	110.01

There are no chirality outliers.

5 of 7 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	25	ARG	Sidechain
1	A	316	ARG	Sidechain
1	B	25	ARG	Sidechain
1	B	331	ARG	Sidechain
1	B	71	ARG	Sidechain

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	2857	0	2806	54	0
1	B	2878	0	2821	53	0
2	A	43	0	30	1	0
2	B	43	0	30	5	0
3	B	36	0	0	3	0
4	A	25	0	0	1	0
4	B	38	0	0	3	0
All	All	5920	0	5687	112	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 10.

The worst 5 of 112 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:169:MET:HG2	1:B:176:PHE:CE1	1.91	1.05
1:B:169:MET:HE3	1:B:184:ALA:CB	2.00	0.91
1:B:169:MET:HG2	1:B:176:PHE:HE1	1.29	0.91
1:B:169:MET:CE	1:B:184:ALA:HB2	2.05	0.87
2:B:402:HEM:FE	2:B:402:HEM:NA	1.41	0.86

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	A	372/390 (95%)	358 (96%)	13 (4%)	1 (0%)	37	47
1	B	380/390 (97%)	361 (95%)	18 (5%)	1 (0%)	37	47
All	All	752/780 (96%)	719 (96%)	31 (4%)	2 (0%)	37	47

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	48	ALA
1	B	178	ALA

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	A	294/315 (93%)	291 (99%)	3 (1%)	73	85
1	B	293/315 (93%)	293 (100%)	0	100	100
All	All	587/630 (93%)	584 (100%)	3 (0%)	86	93

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	379	ILE
1	A	387	LEU
1	A	391	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	350	ASN
1	B	113	HIS
1	B	354	HIS
1	B	147	GLN
1	A	341	GLN

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 oligosaccharides in this entry.

5.6   Ligand geometry ⓘ

3 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	HEM	A	401	4	41,50,50	1.80	9 (21%)	45,82,82	1.09	4 (8%)
3	A1EGF	B	401	-	29,38,38	1.26	3 (10%)	33,56,56	1.45	4 (12%)
2	HEM	B	402	1	41,50,50	1.46	6 (14%)	45,82,82	1.07	2 (4%)

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
2	HEM	A	401	4	-	1/12/54/54	-
3	A1EGF	B	401	-	-	6/34/72/72	0/2/3/3
2	HEM	B	402	1	-	1/12/54/54	-

The worst 5 of 18 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	401	HEM	FE-NB	5.09	2.22	1.96
3	B	401	A1EGF	C6-C7	4.74	1.64	1.50
2	A	401	HEM	CAB-C3B	-4.27	1.35	1.47
2	B	402	HEM	FE-NB	3.91	2.16	1.96
2	A	401	HEM	C3C-CAC	-3.79	1.40	1.47

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	401	A1EGF	C5-C6-C7	5.92	131.25	112.69
2	B	402	HEM	C4B-CHC-C1C	3.79	127.56	122.56
2	B	402	HEM	C2C-C3C-C4C	3.21	109.14	106.90
2	A	401	HEM	C4C-CHD-C1D	2.89	126.37	122.56
3	B	401	A1EGF	C21-C20-N1	-2.79	108.56	111.68

There are no chirality outliers.

5 of 8 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	401	A1EGF	C6-C7-C8-C9
3	B	401	A1EGF	C17-C18-C19-S3
3	B	401	A1EGF	C5-C6-C7-C8
3	B	401	A1EGF	C1-C2-C3-C5
3	B	401	A1EGF	C1-C2-C3-C4

There are no ring outliers.

3 monomers are involved in 8 short contacts:

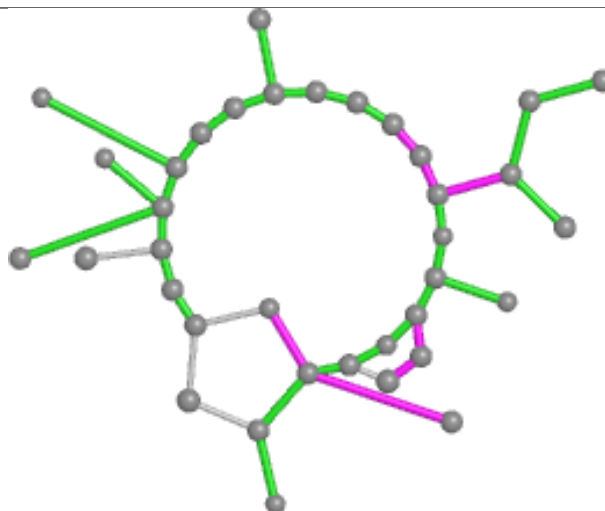| Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	401	HEM	1	0
3	B	401	A1EGF	3	0
2	B	402	HEM	5	0

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 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 A 401

Bond lengths

Bond angles
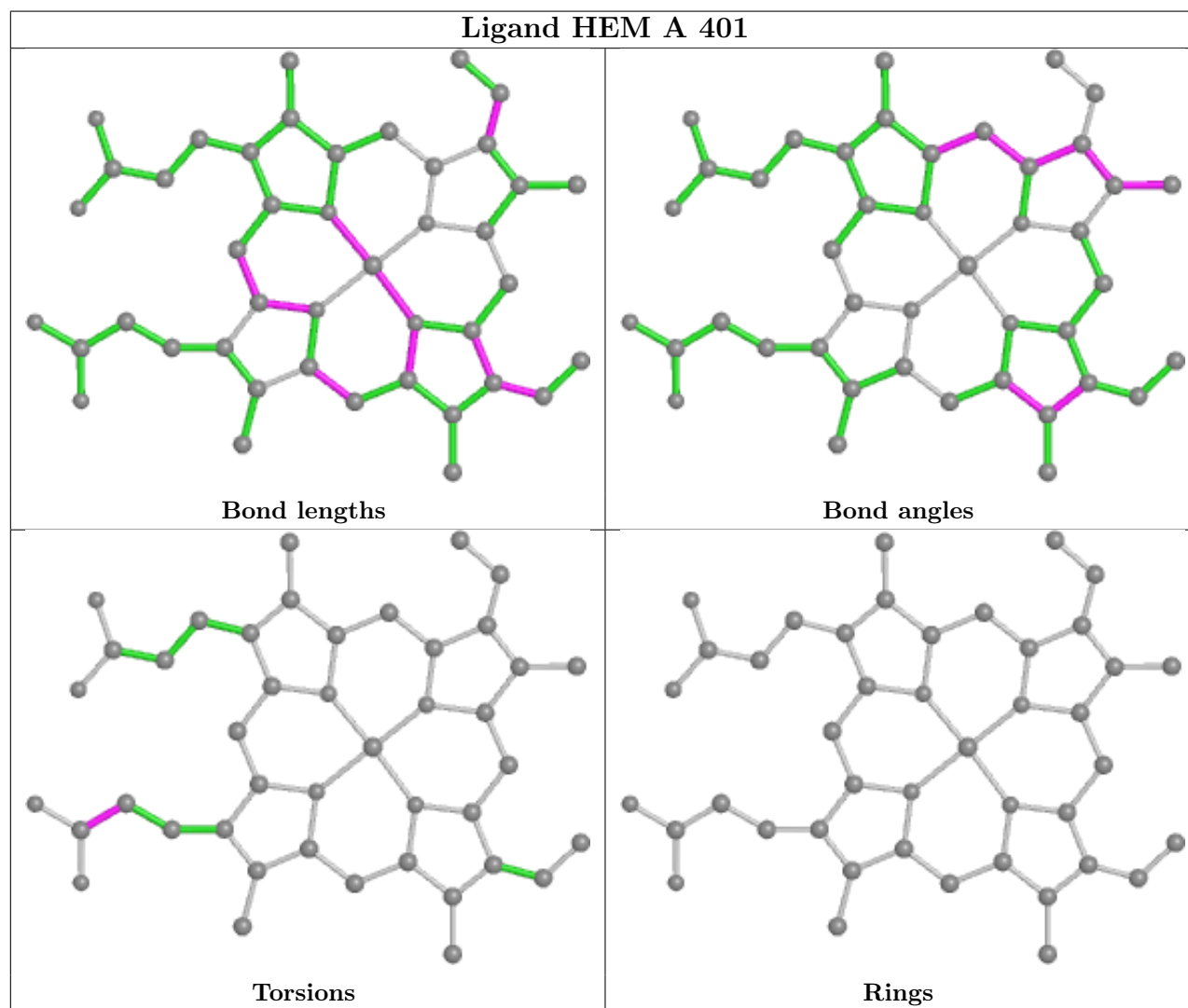
Torsions

Rings

Ligand A1EGF B 401

Bond lengths

Bond angles

Torsions

Rings

Ligand HEM B 402
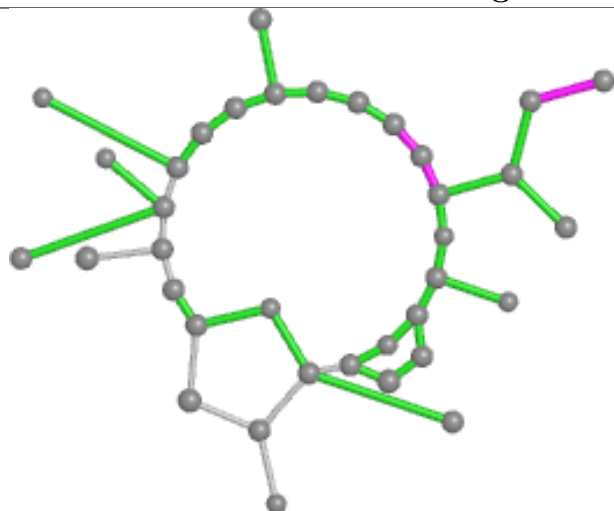
Bond lengths

Bond angles
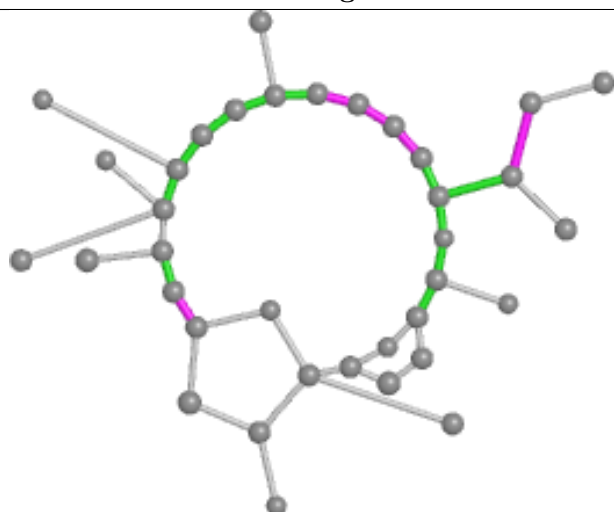
Torsions
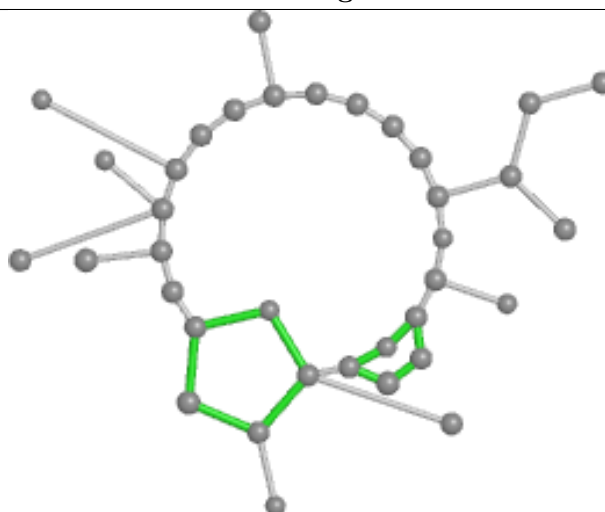
Rings
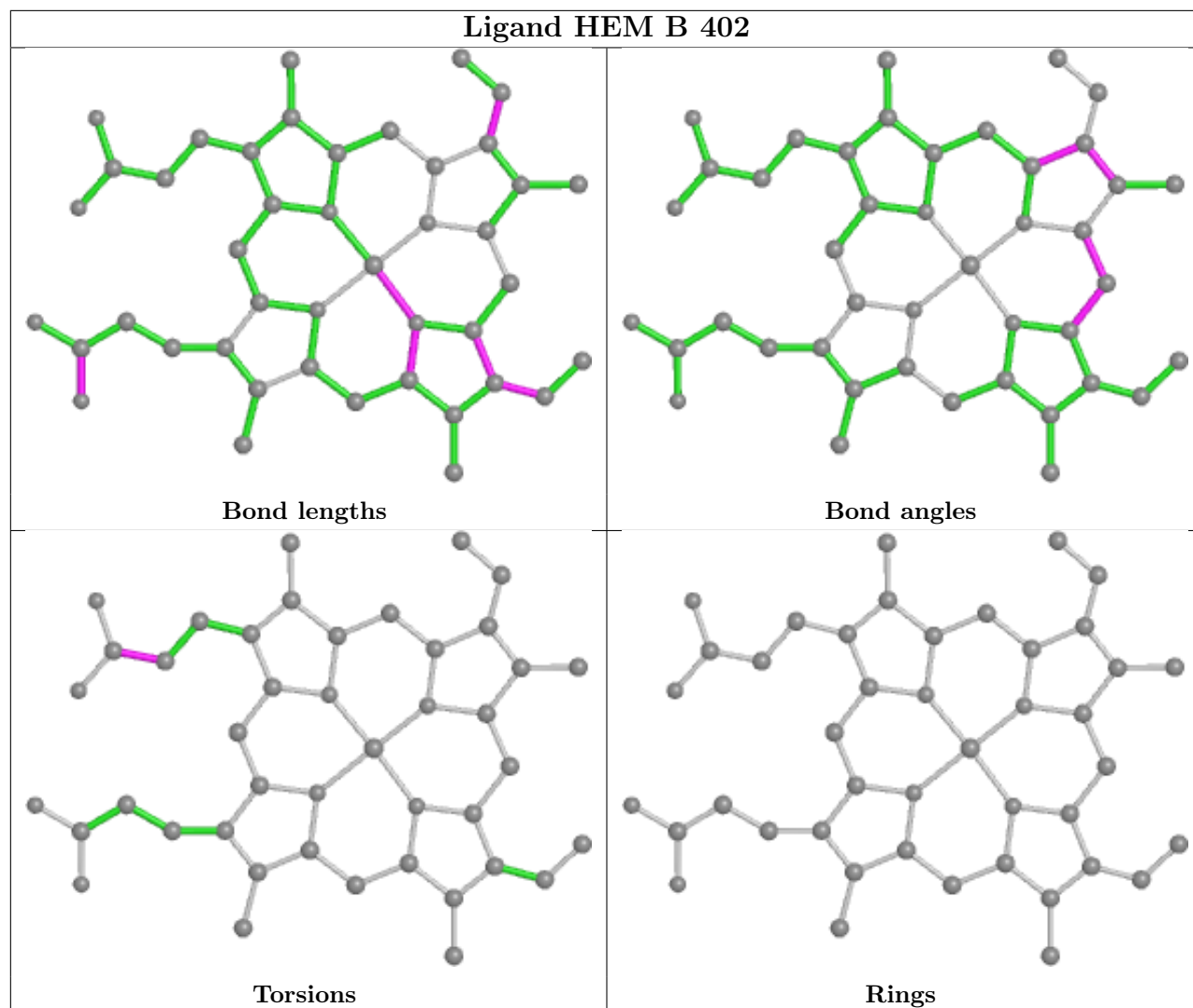
5.7 Other polymers ⓘ

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	A	380/390 (97%)	0.45	22 (5%)	30	32	32, 47, 82, 117	0
1	B	386/390 (98%)	0.39	21 (5%)	32	34	30, 46, 79, 112	0
All	All	766/780 (98%)	0.42	43 (5%)	31	33	30, 46, 81, 117	0

The worst 5 of 43 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	8	ILE	8.8
1	A	176	PHE	5.8
1	A	9	PHE	5.5
1	B	171	ILE	4.6
1	A	177	THR	4.5

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 oligosaccharides 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	A1EGF	B	401	36/36	0.86	0.15	46,69,84,87	0
2	HEM	B	402	43/43	0.87	0.12	25,35,43,47	0
2	HEM	A	401	43/43	0.91	0.10	27,35,44,53	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 A1EGF 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)
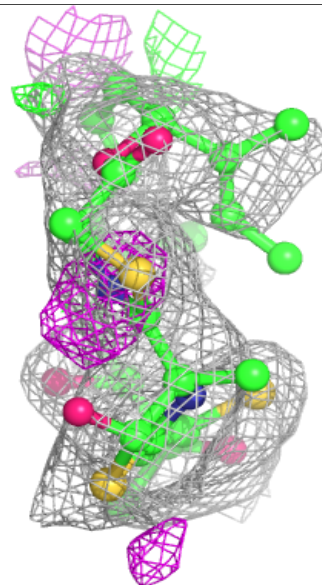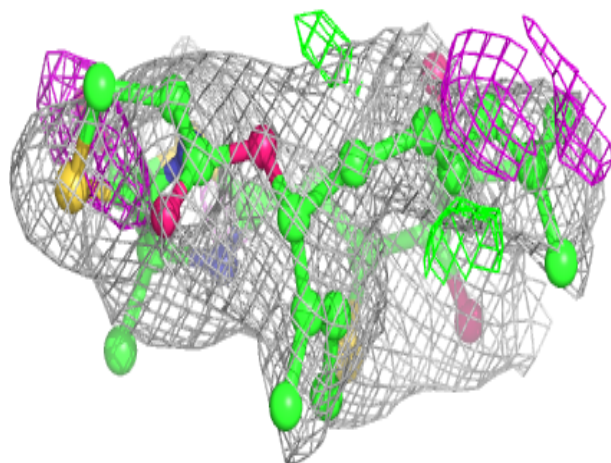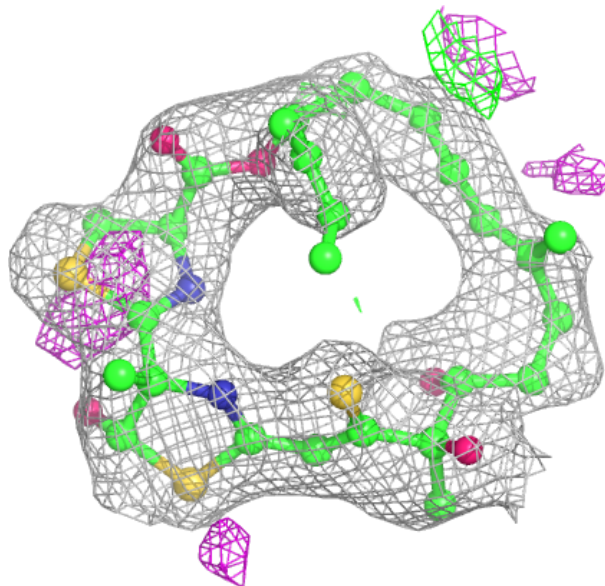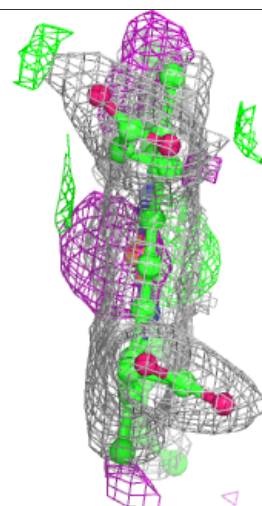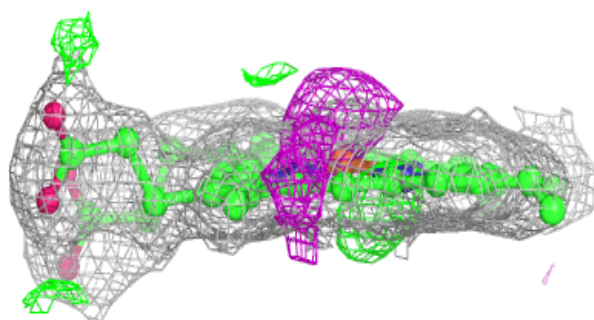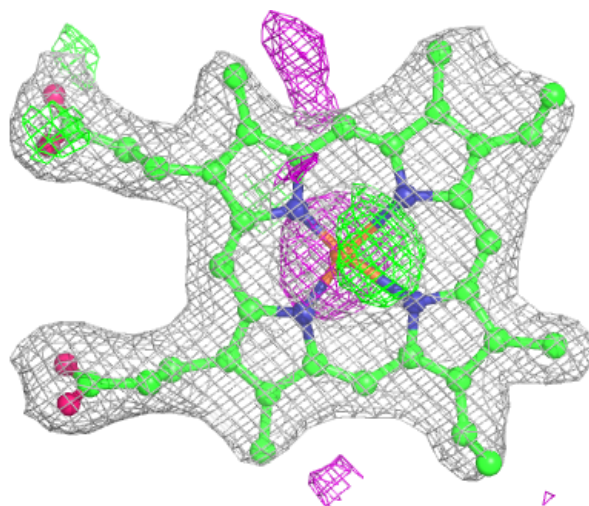
Electron density around HEM B 402:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)
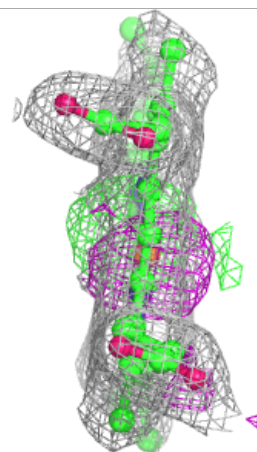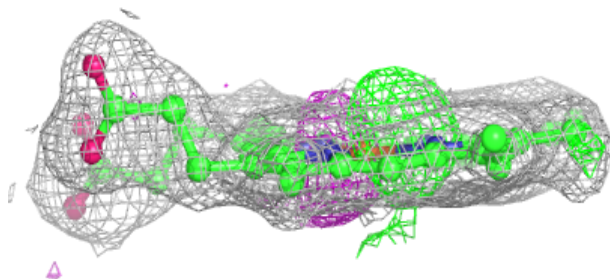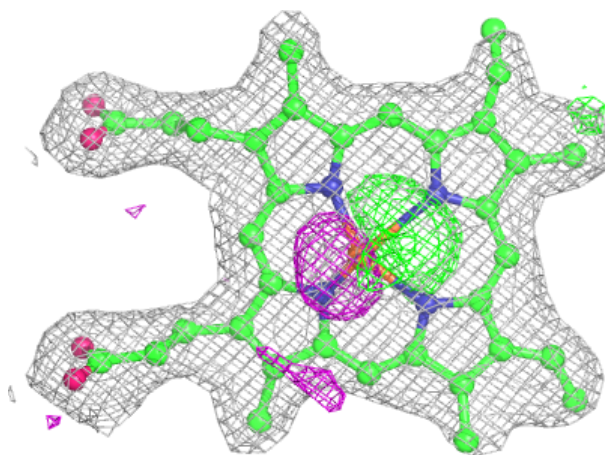
Electron density around HEM A 401:

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

6.5 Other polymers ⓘ

There are no such residues in this entry.