WORLDWIDE PDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 12, 2024 – 06:05 AM EST

PDB ID	:	1TH4
Title	:	crystal structure of NADPH depleted bovine liver catalase complexed with 3-amino-1,2,4-triazole
Authors	:	Sugadev, R.; Ponnuswamy, M.N.; Kumaran, D.; Swaminathan, S.; Sekar, K.
Deposited on	:	2004-06-01
Resolution	:	2.98 Å(reported)

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

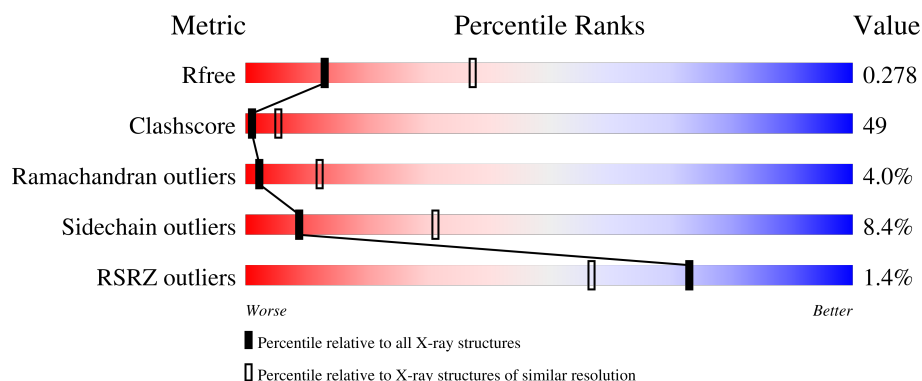|		
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

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 2.98 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		0.278
Clashscore		49
Ramachandran outliers		4.0%
Sidechain outliers		8.4%
RSRZ outliers		1.4%

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	3360 (3.00-2.96)
Clashscore	180529	3751 (3.00-2.96)
Ramachandran outliers	177936	3628 (3.00-2.96)
Sidechain outliers	177891	3631 (3.00-2.96)
RSRZ outliers	164620	3372 (3.00-2.96)

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	506	32% 58% 8% ••
1	B	506	35% 53% 9% ••
1	C	506	38% 51% 9% ••
1	D	506	39% 52% 7% •

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard

residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

| Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	HEM	A	2000	-	-	X	-
2	HEM	B	2001	-	-	X	-
2	HEM	C	2002	-	-	X	-
2	HEM	D	2003	-	-	X	-
3	3TR	D	3074	-	-	X	-

2   Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 16915 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 Catalase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	499	Total	C	N	O	S	0	0	0
			4017	2548	715	740	14			
1	B	499	Total	C	N	O	S	0	0	0
			4017	2548	715	740	14			
1	C	499	Total	C	N	O	S	1	0	0
			4017	2548	715	740	14			
1	D	499	Total	C	N	O	S	0	0	0
			4017	2548	715	740	14			

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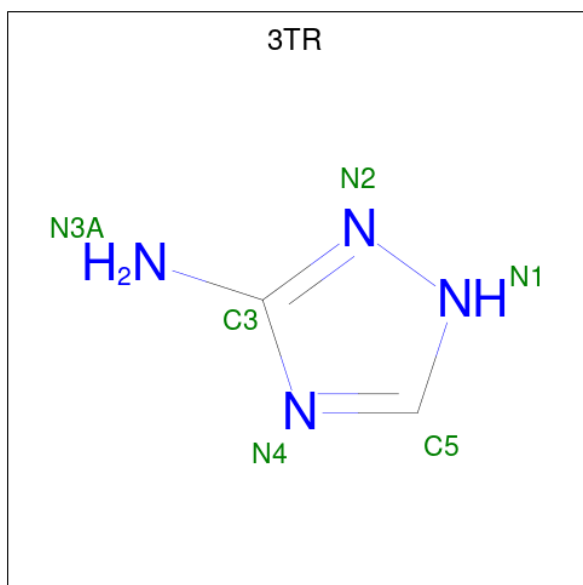

HEM
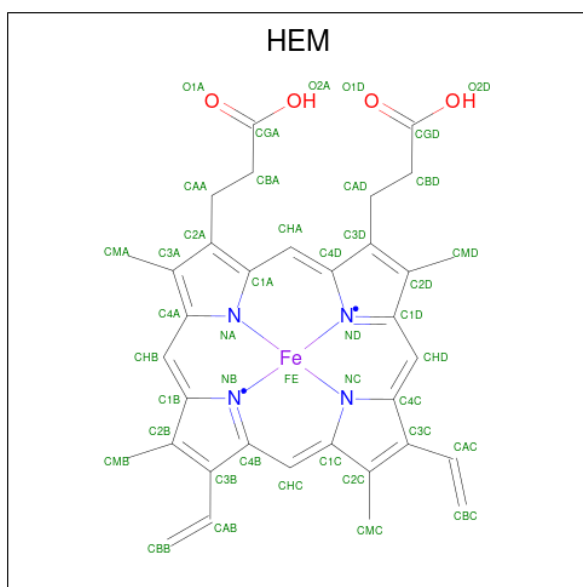
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
2	B	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		

Continued on next page...

Continued from previous page...

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

- Molecule 3 is 3-AMINO-1,2,4-TRIAZOLE (three-letter code: 3TR) (formula: $C_2H_4N_4$).

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	D	1	Total	C	N	0	0
			6	2	4		

- Molecule 4 is water.

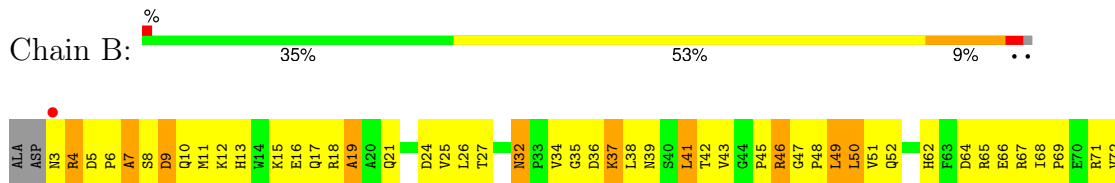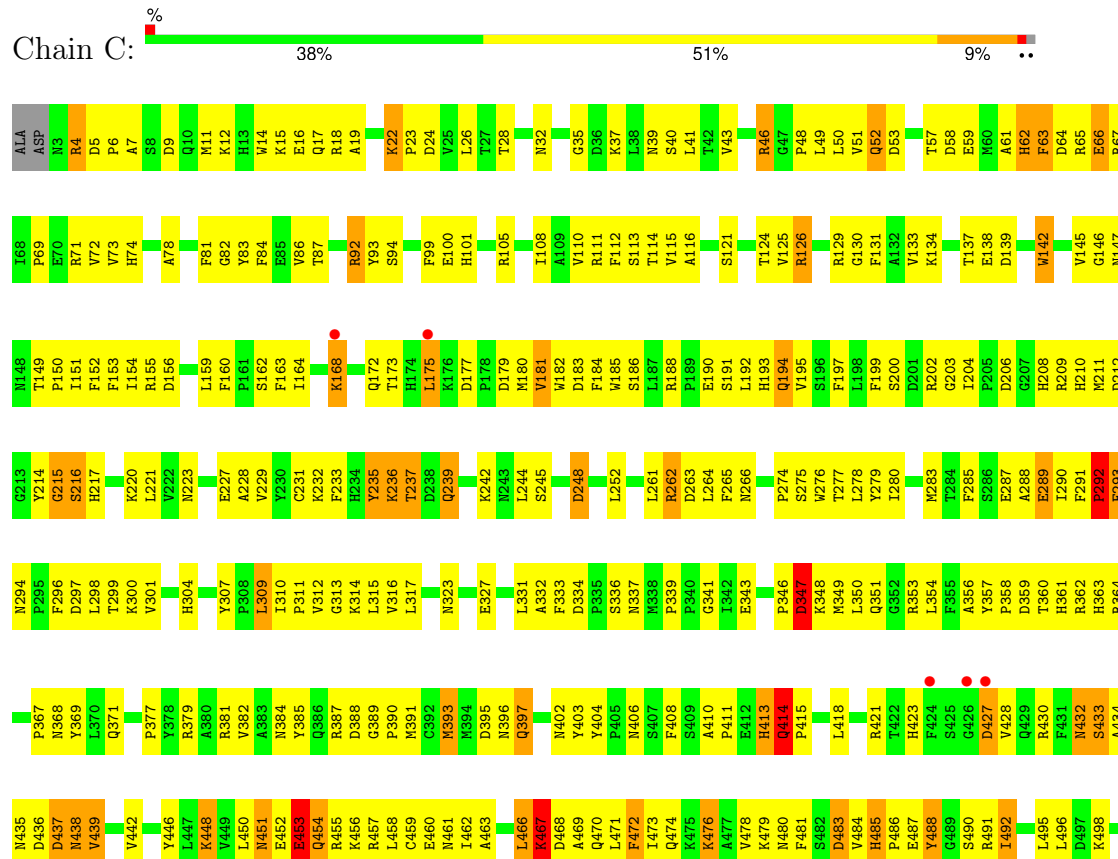| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	154	Total	O	0	0
			154	154		
4	B	145	Total	O	0	0
			145	145		
4	C	188	Total	O	0	0
			188	188		
4	D	182	Total	O	0	0
			182	182		

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.
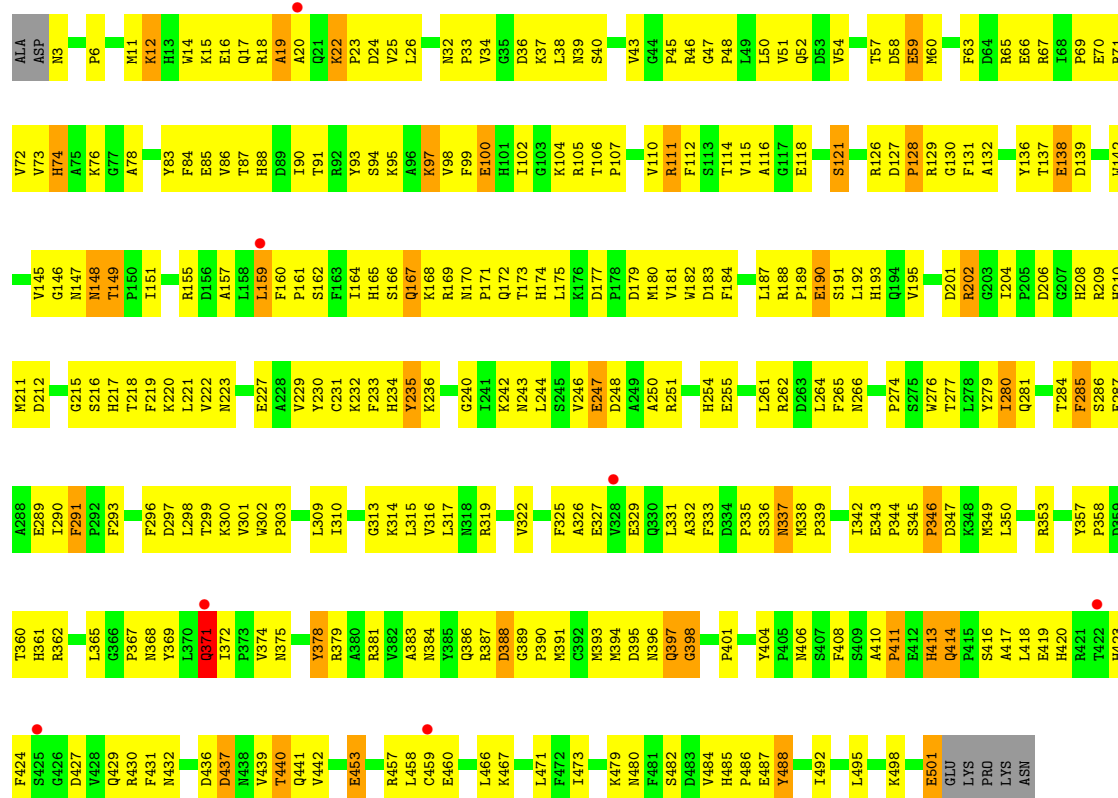
- Molecule 1: Catalase

Chain A:

- Molecule 1: Catalase

Chain B:

• Molecule 1: Catalase

E501 GLU LYS PRO LYS ASN

● Molecule 1: Catalase

%

Chain D: 39%   52%   7%   ·
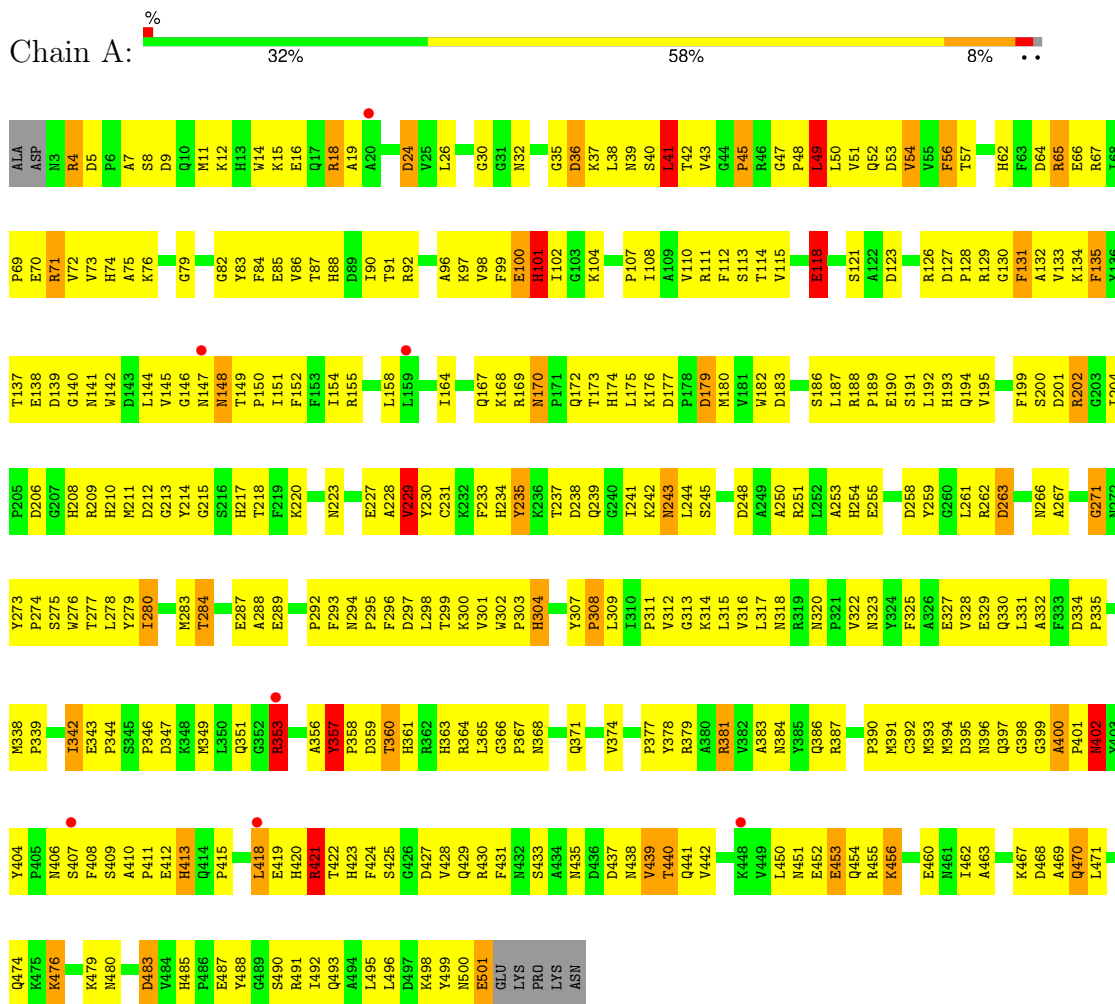
4   Data and refinement statistics   ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	85.03Å 141.09Å 231.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	39.90 – 2.98 39.90 – 2.98	Depositor EDS
% Data completeness (in resolution range)	57.6 (39.90-2.98) 57.7 (39.90-2.98)	Depositor EDS
R_{merge}	0.30	Depositor
R_{sym}	0.30	Depositor
$< I/\sigma(I) >$ [1]	0.73 (at 2.96Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.235 , 0.292 0.220 , 0.278	Depositor DCC
R_{free} test set	1000 reflections (3.00%)	wwPDB-VP
Wilson B-factor (Å^2)	53.3	Xtriage
Anisotropy	0.182	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 62.5	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	0.89	EDS
Total number of atoms	16915	wwPDB-VP
Average B, all atoms (Å^2)	46.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.54% 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 ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: 3TR, 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	0.60	5/4137 (0.1%)	1.78	61/5619 (1.1%)
1	B	0.99	6/4137 (0.1%)	1.45	31/5619 (0.6%)
1	C	0.66	8/4137 (0.2%)	1.49	33/5619 (0.6%)
1	D	0.61	1/4137 (0.0%)	0.92	9/5619 (0.2%)
All	All	0.73	20/16548 (0.1%)	1.44	134/22476 (0.6%)

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	1	2
1	B	1	3
1	C	2	2
1	D	1	0
All	All	5	7

The worst 5 of 20 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	319	ARG	CD-NE	49.16	2.30	1.46
1	D	413	HIS	CA-CB	-20.41	1.09	1.53
1	C	202	ARG	NE-CZ	16.91	1.55	1.33
1	B	319	ARG	NE-CZ	16.04	1.53	1.33
1	A	304	HIS	CA-CB	12.45	1.81	1.53

The worst 5 of 134 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	357	TYR	CB-CG-CD2	-68.04	80.18	121.00
1	C	202	ARG	NE-CZ-NH2	-52.67	93.97	120.30
1	B	319	ARG	NE-CZ-NH1	-46.60	97.00	120.30
1	C	202	ARG	NE-CZ-NH1	44.10	142.35	120.30
1	B	319	ARG	CG-CD-NE	-41.12	25.46	111.80

All (5) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	395	ASP	CA
1	B	395	ASP	CA
1	C	453	GLU	CA
1	C	501	GLU	CA
1	D	371	GLN	CA

5 of 7 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	357	TYR	Sidechain
1	A	421	ARG	Sidechain
1	B	126	ARG	Sidechain
1	B	319	ARG	Sidechain
1	B	394	MET	Peptide

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	4017	0	3838	473	1
1	B	4017	0	3839	431	0
1	C	4017	0	3840	429	0
1	D	4017	0	3839	389	0
2	A	43	0	30	79	0
2	B	43	0	30	39	0
2	C	43	0	30	49	0
2	D	43	0	30	25	0
3	D	6	0	3	12	0
4	A	154	0	0	30	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	145	0	0	28	0
4	C	188	0	0	17	1
4	D	182	0	0	14	0
All	All	16915	0	15479	1562	1

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 49.

The worst 5 of 1562 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:304:HIS:CB	1:A:304:HIS:CA	1.81	1.58
1:A:100:GLU:HG2	1:A:101:HIS:CE1	1.43	1.52
2:A:2000:HEM:CAC	2:A:2000:HEM:CBC	1.87	1.47
1:A:353:ARG:HD2	2:A:2000:HEM:CHC	1.53	1.37
1:C:111:ARG:NE	2:C:2002:HEM:O1D	1.57	1.36

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:289:GLU:OE1	4:C:2061:HOH:O[3_655]	2.18	0.02

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	497/506 (98%)	400 (80%)	80 (16%)	17 (3%)	3	15
1	B	497/506 (98%)	408 (82%)	65 (13%)	24 (5%)	2	9
1	C	497/506 (98%)	398 (80%)	82 (16%)	17 (3%)	3	15

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	D	497/506 (98%)	402 (81%)	74 (15%)	21 (4%)	2	11
All	All	1988/2024 (98%)	1608 (81%)	301 (15%)	79 (4%)	2	12

5 of 79 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	100	GLU
1	B	19	ALA
1	B	100	GLU
1	B	331	LEU
1	B	394	MET

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	431/437 (99%)	396 (92%)	35 (8%)	9	32
1	B	431/437 (99%)	387 (90%)	44 (10%)	6	23
1	C	431/437 (99%)	394 (91%)	37 (9%)	8	30
1	D	431/437 (99%)	403 (94%)	28 (6%)	14	41
All	All	1724/1748 (99%)	1580 (92%)	144 (8%)	9	31

5 of 144 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	501	GLU
1	D	501	GLU
1	D	131	PHE
1	D	314	LYS
1	B	144	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 74 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	52	GLN
1	D	480	ASN
1	D	167	GLN
1	D	330	GLN
1	B	234	HIS

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 ⓘ

5 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	D	2003	1,3	42,50,50	2.62	10 (23%)	46,82,82	4.88	7 (15%)
3	3TR	D	3074	1,2	5,6,6	2.75	3 (60%)	0,7,7	-	-
2	HEM	C	2002	1	42,50,50	1.45	6 (14%)	46,82,82	1.04	2 (4%)
2	HEM	A	2000	1	42,50,50	2.18	8 (19%)	46,82,82	1.24	3 (6%)
2	HEM	B	2001	1	42,50,50	1.88	8 (19%)	46,82,82	3.24	6 (13%)

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	D	2003	1,3	-	9/12/54/54	-
3	3TR	D	3074	1,2	-	-	0/1/1/1
2	HEM	C	2002	1	-	6/12/54/54	-
2	HEM	A	2000	1	-	6/12/54/54	-
2	HEM	B	2001	1	-	7/12/54/54	-

The worst 5 of 35 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	2003	HEM	CHC-C4B	-10.89	1.10	1.40
2	A	2000	HEM	CBC-CAC	9.32	1.87	1.29
2	D	2003	HEM	CHD-C1D	-8.26	1.17	1.40
2	B	2001	HEM	C4A-CHB	6.10	1.58	1.41
2	A	2000	HEM	FE-NB	5.14	2.26	1.98

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	D	2003	HEM	CBA-CAA-C2A	29.86	162.74	112.54
2	B	2001	HEM	CHD-C1D-C2D	-12.76	104.89	125.03
2	B	2001	HEM	CHD-C1D-ND	11.76	137.09	124.44
2	B	2001	HEM	C4C-CHD-C1D	-10.66	108.49	122.56
2	D	2003	HEM	C4B-CHC-C1C	8.06	133.19	122.56

There are no chirality outliers.

5 of 28 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	2000	HEM	C1A-C2A-CAA-CBA
2	A	2000	HEM	C3A-C2A-CAA-CBA
2	A	2000	HEM	C2B-C3B-CAB-CBB
2	A	2000	HEM	C4B-C3B-CAB-CBB
2	C	2002	HEM	C2B-C3B-CAB-CBB

There are no ring outliers.

5 monomers are involved in 192 short contacts:

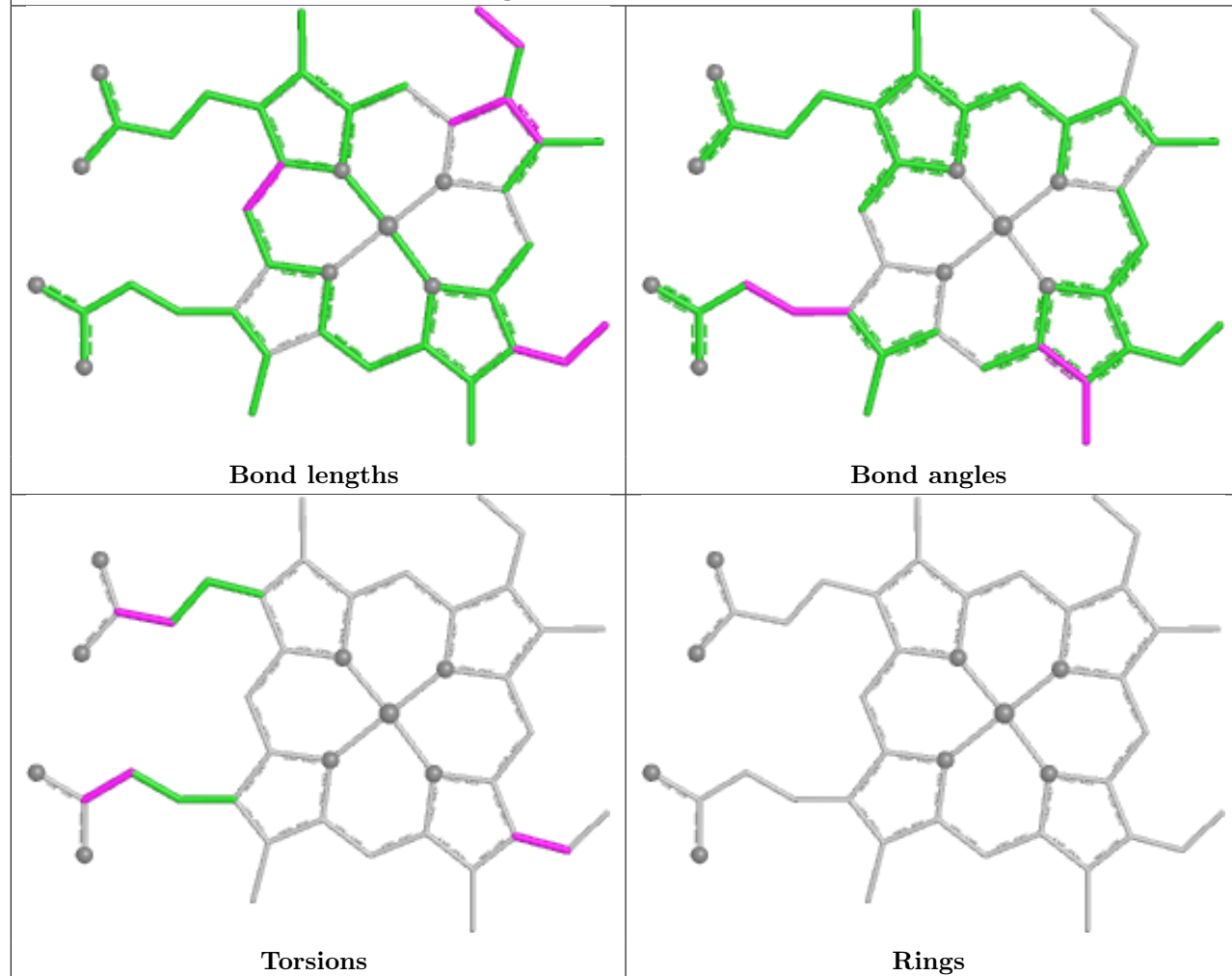| Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	2003	HEM	25	0
3	D	3074	3TR	12	0
2	C	2002	HEM	49	0
2	A	2000	HEM	79	0
2	B	2001	HEM	39	0

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 D 2003
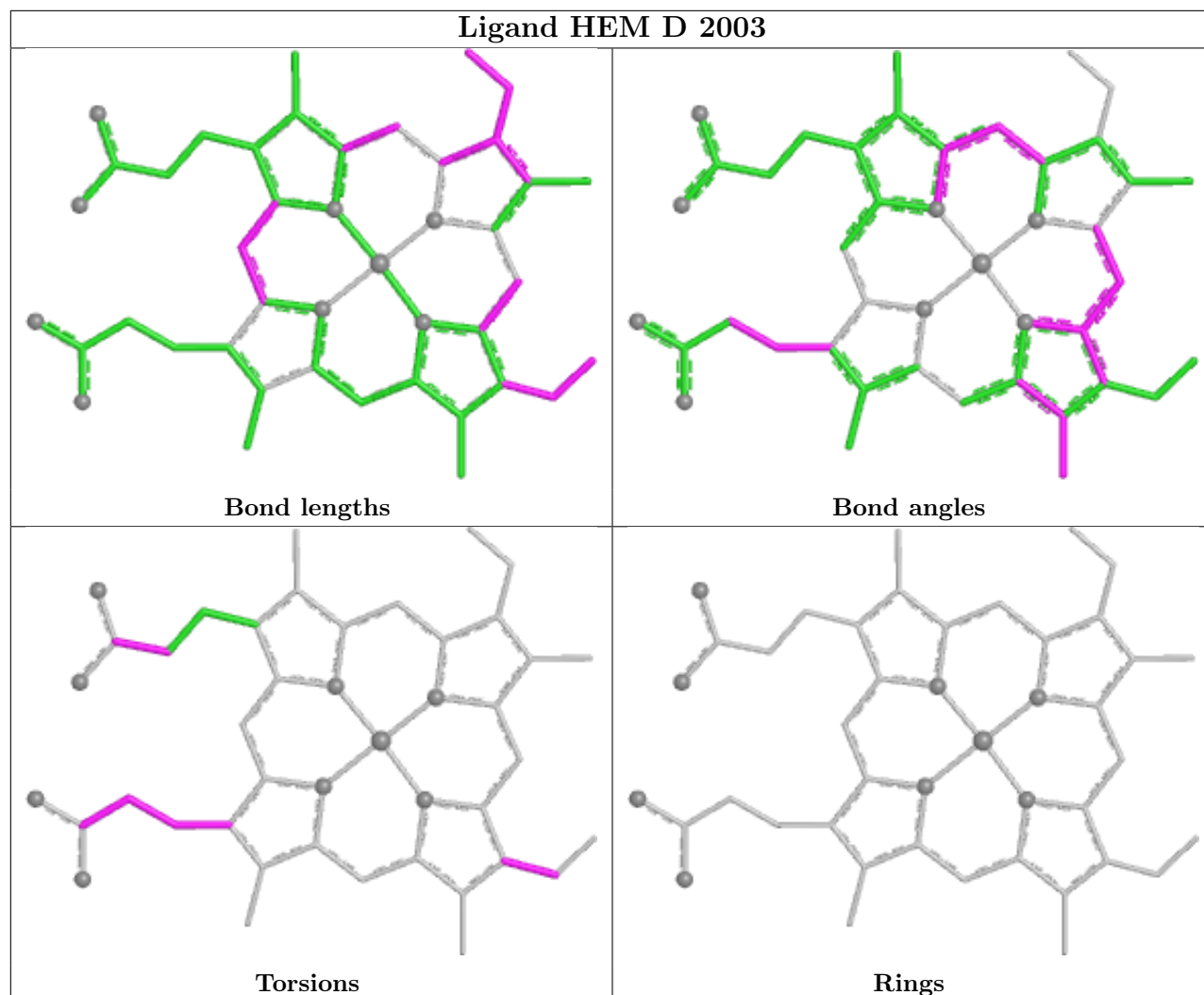
Bond lengths

Bond angles

Torsions

Rings

Ligand HEM C 2002

Bond lengths

Bond angles

Torsions

Rings

Ligand HEM A 2000

Bond lengths

Bond angles

Torsions

Rings

Ligand HEM B 2001

Bond lengths

Bond angles

Torsions

Rings
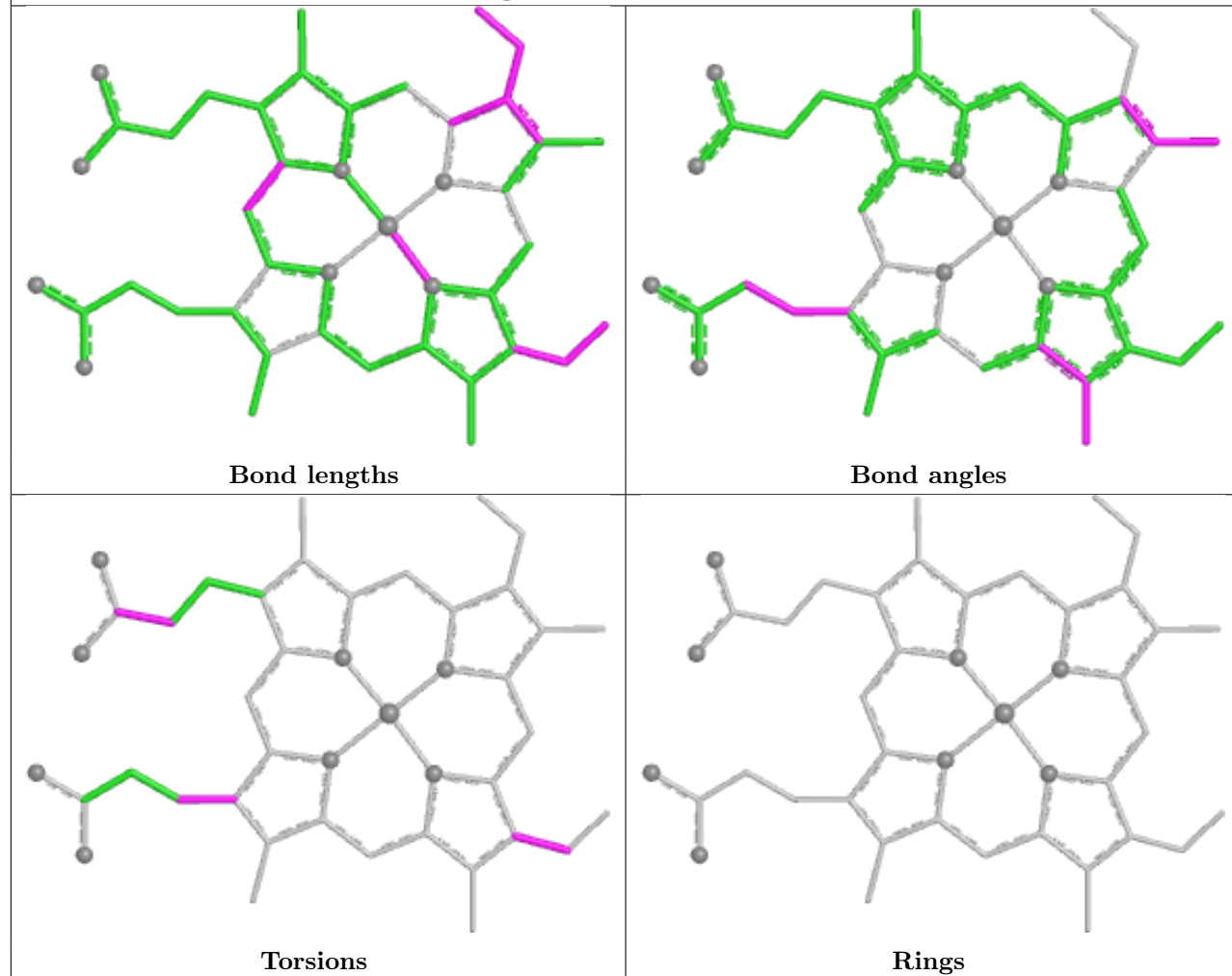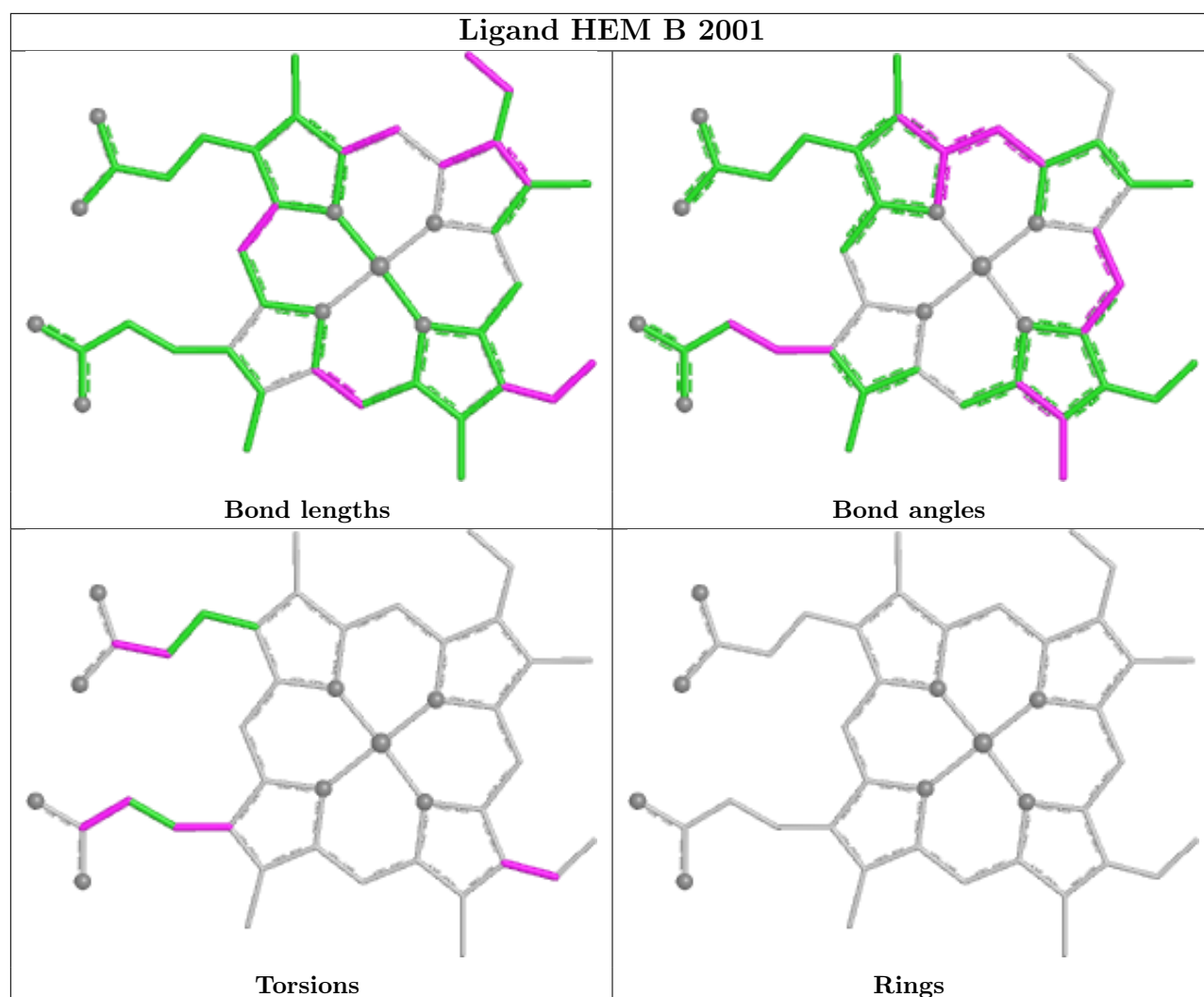
5.7 Other polymers ⓘ

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 (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	499/506 (98%)	0.13	7 (1%)	73	57	10, 40, 76, 97	0
1	B	499/506 (98%)	0.17	7 (1%)	73	57	16, 44, 78, 102	0
1	C	499/506 (98%)	0.12	6 (1%)	76	60	16, 41, 81, 100	1 (0%)
1	D	499/506 (98%)	0.18	7 (1%)	73	57	15, 45, 84, 101	0
All	All	1996/2024 (98%)	0.15	27 (1%)	73	57	10, 43, 80, 102	1 (0%)

The worst 5 of 27 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	353	ARG	5.8
1	D	371	GLN	3.7
1	C	424	PHE	3.1
1	C	501	GLU	2.9
1	C	175	LEU	2.8

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
2	HEM	A	2000	43/43	0.77	0.30	80,94,101,101	0
2	HEM	C	2002	43/43	0.81	0.23	41,58,72,92	0
2	HEM	B	2001	43/43	0.88	0.20	21,56,83,101	0
3	3TR	D	3074	6/6	0.89	0.26	89,96,99,101	0
2	HEM	D	2003	43/43	0.95	0.11	20,38,56,59	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.
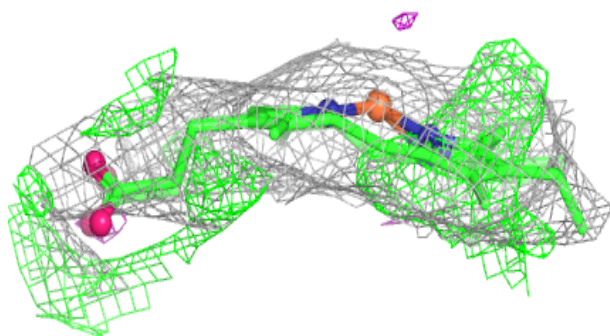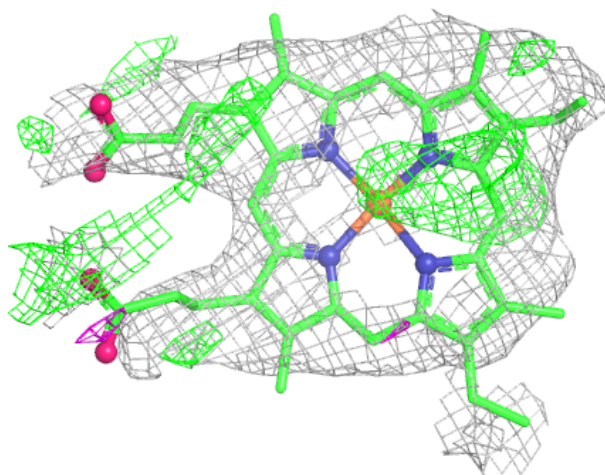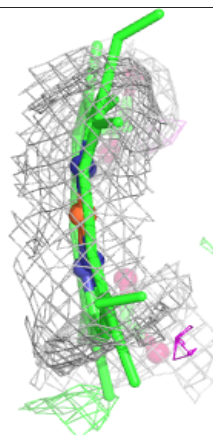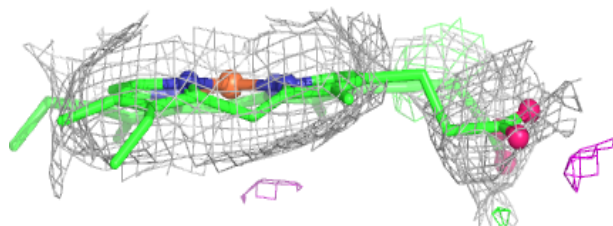

Electron density around HEM A 2000:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)
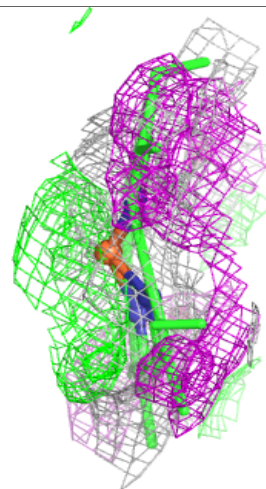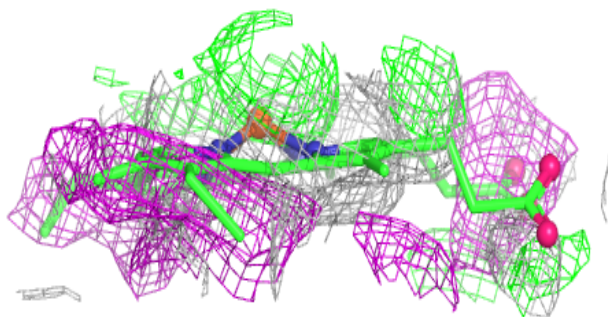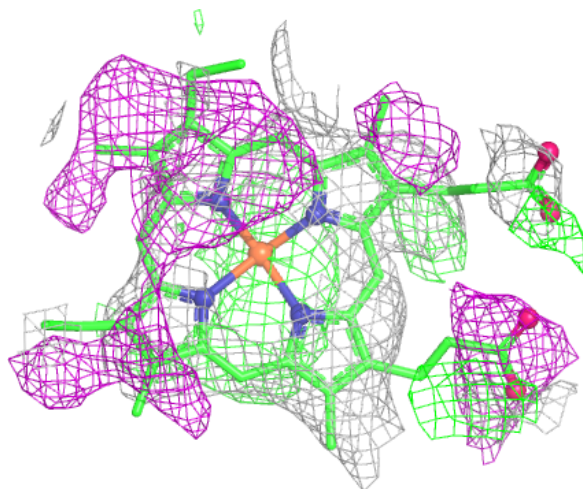
Electron density around HEM C 2002:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
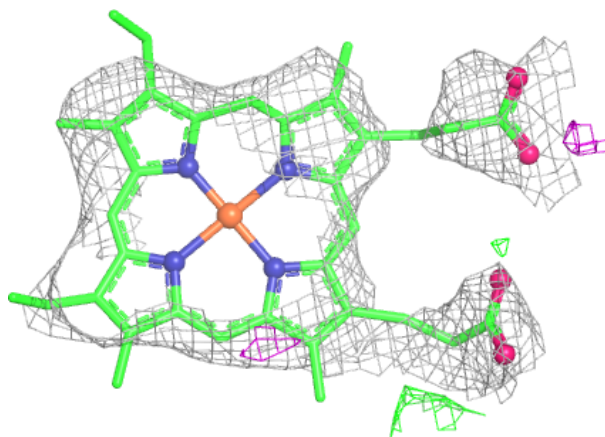and green (positive)
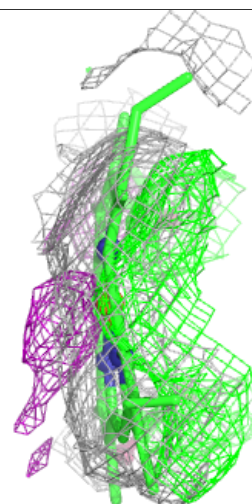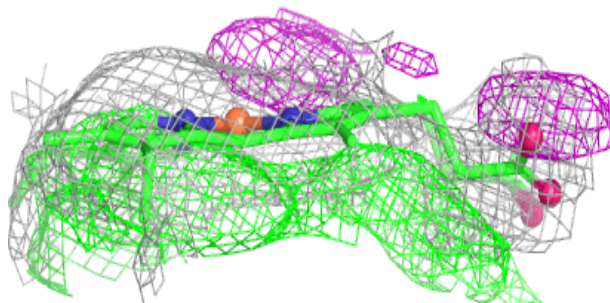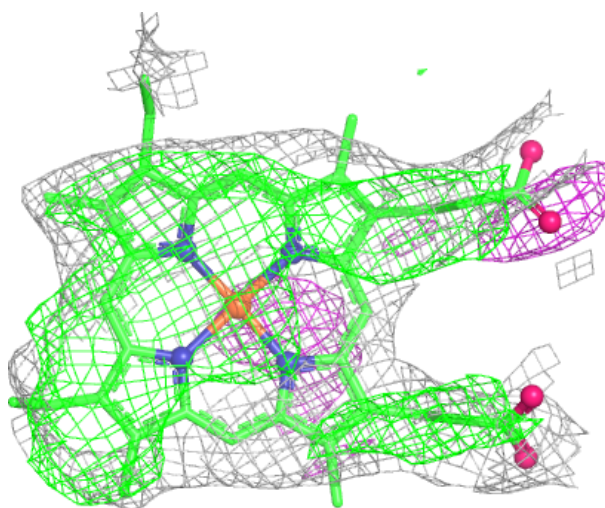
Electron density around HEM B 2001:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)
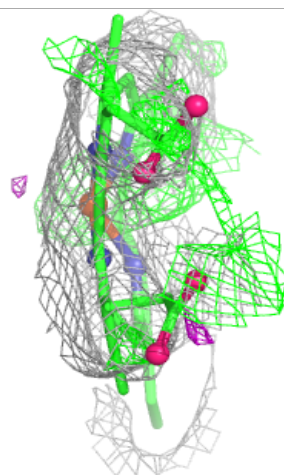
Electron density around HEM D 2003:

$2mF_o$-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.