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 (i)

Feb 27, 2021 – 08:36 PM EST

PDB ID	:	1SP3
Title	:	Crystal structure of octaheme cytochrome c from Shewanella oneidensis
Authors	:	Mowat, C.G.; Rothery, E.; Miles, C.S.; McIver, L.; Doherty, M.K.; Drewette, K.; Taylor, P.; Walkinshaw, M.D.; Chapman, S.K.; Reid, G.A.
Deposited on	:	2004-03-16
Resolution	:	2.20 Å(reported)

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

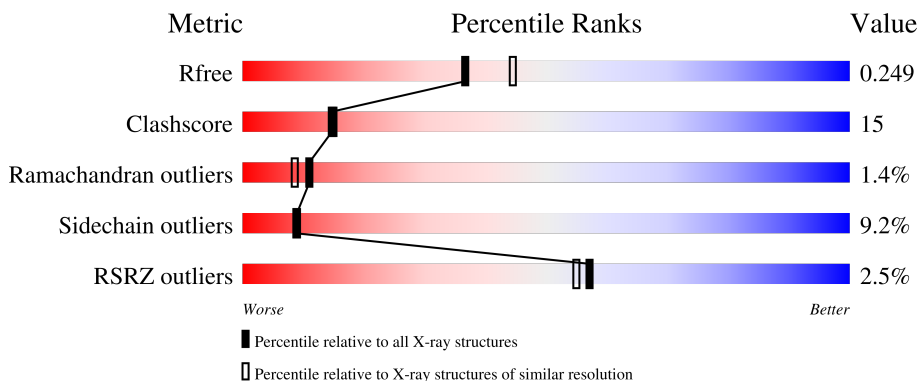|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.17.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.17.1

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 2.20 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 |

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	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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	443	2% 71% 21% 6% ··

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 4243 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.
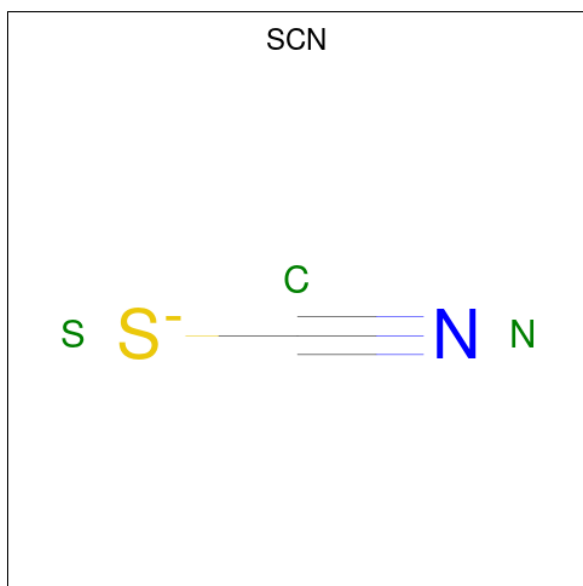
- Molecule 1 is a protein called cytochrome c, putative.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	436	Total 3426	C 2140	N 598	O 655	S 33	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	211	ASN	GLN	conflict	UNP Q8E9W8

- Molecule 2 is THIOCYANATE ION (three-letter code: SCN) (formula: CNS).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total 3	C 1	N 1	S 1	0	0

- Molecule 3 is HEME C (three-letter code: HEC) (formula: $C_{34}H_{34}FeN_4O_4$).

HEC

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

• Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	470	Total	O	0	0
			470	470		

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 c, putative
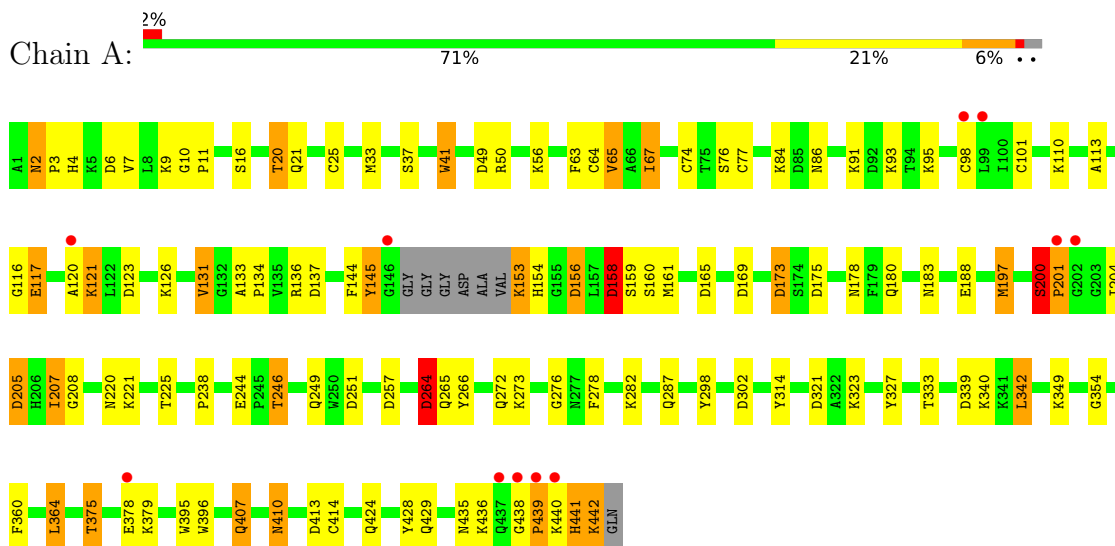
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	40.59Å 61.77Å 185.24Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	23.77 – 2.20 23.76 – 2.20	Depositor EDS
% Data completeness (in resolution range)	100.0 (23.77-2.20) 96.6 (23.76-2.20)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	3.33 (at 2.19Å)	Xtriage
Refinement program	REFMAC 5.1.24	Depositor
R, R_{free}	0.155 , 0.252 0.158 , 0.249	Depositor DCC
R_{free} test set	1217 reflections (5.14%)	wwPDB-VP
Wilson B-factor (Å^2)	35.5	Xtriage
Anisotropy	0.044	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 41.8	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	0.97	EDS
Total number of atoms	4243	wwPDB-VP
Average B, all atoms (Å^2)	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.08% 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: HEC, SCN

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	1.20	4/3519 (0.1%)	1.15	22/4760 (0.5%)

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	4

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	314	TYR	CD2-CE2	7.79	1.51	1.39
1	A	327	TYR	CD2-CE2	5.74	1.48	1.39
1	A	37	SER	CB-OG	-5.40	1.35	1.42
1	A	41	TRP	CB-CG	5.23	1.59	1.50

The worst 5 of 22 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	342	LEU	C-N-CA	-8.74	99.85	121.70
1	A	165	ASP	CB-CG-OD2	7.25	124.82	118.30
1	A	207	ILE	C-N-CA	-7.02	107.56	122.30
1	A	156	ASP	CB-CG-OD2	6.89	124.50	118.30
1	A	205	ASP	CB-CG-OD2	6.84	124.45	118.30

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	10	GLY	Peptide
1	A	153	LYS	Peptide
1	A	264	ASP	Peptide
1	A	441	HIS	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	3426	0	3243	99	0
2	A	3	0	0	0	0
3	A	344	0	243	28	0
4	A	470	0	0	24	4
All	All	4243	0	3486	106	4

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 15.

The worst 5 of 106 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:74:CYS:SG	3:A:802:HEC:CAB	2.10	1.38
1:A:25:CYS:SG	3:A:801:HEC:HAC	1.77	1.15
1:A:429:GLN:H	1:A:439:PRO:HG3	1.17	1.07
1:A:101:CYS:SG	3:A:803:HEC:HAC	1.94	1.04
1:A:74:CYS:SG	3:A:802:HEC:HAB	2.00	1.00

All (4) 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)
4:A:975:HOH:O	4:A:1274:HOH:O[1_455]	1.49	0.71
4:A:971:HOH:O	4:A:1276:HOH:O[4_455]	1.53	0.67
4:A:1075:HOH:O	4:A:1123:HOH:O[3_545]	2.04	0.16
4:A:1044:HOH:O	4:A:1123:HOH:O[3_545]	2.07	0.13

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	432/443 (98%)	406 (94%)	20 (5%)	6 (1%)	11	8

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	200	SER
1	A	201	PRO
1	A	11	PRO
1	A	173	ASP
1	A	439	PRO

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	370/373 (99%)	336 (91%)	34 (9%)	9	9

5 of 34 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	410	ASN
1	A	424	GLN
1	A	440	LYS
1	A	158	ASP
1	A	153	LYS

Sometimes 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	277	ASN
1	A	363	ASN
1	A	441	HIS
1	A	407	GLN
1	A	178	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 monosaccharides in this entry.

5.6   Ligand geometry ⓘ

9 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	HEC	A	808	1	26,50,50	2.53	8 (30%)	18,82,82	1.93	7 (38%)
3	HEC	A	806	1	26,50,50	2.16	6 (23%)	18,82,82	2.68	7 (38%)
3	HEC	A	803	1	26,50,50	2.92	13 (50%)	18,82,82	1.91	5 (27%)
3	HEC	A	801	1	26,50,50	2.36	8 (30%)	18,82,82	2.73	7 (38%)
3	HEC	A	804	1	26,50,50	2.68	9 (34%)	18,82,82	3.02	7 (38%)
2	SCN	A	810	3	1,2,2	0.53	0	0,1,1	0.00	-
3	HEC	A	805	1	26,50,50	2.31	10 (38%)	18,82,82	2.99	9 (50%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#|Z| > 2	Counts	RMSZ	#|Z| > 2
3	HEC	A	807	1	26,50,50	2.69	8 (30%)	18,82,82	2.88	11 (61%)
3	HEC	A	802	1,2	26,50,50	2.50	14 (53%)	18,82,82	2.37	7 (38%)

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	HEC	A	808	1	-	0/6/54/54	-
3	HEC	A	806	1	-	0/6/54/54	-
3	HEC	A	803	1	-	0/6/54/54	-
3	HEC	A	801	1	-	0/6/54/54	-
3	HEC	A	804	1	-	1/6/54/54	-
3	HEC	A	805	1	-	0/6/54/54	-
3	HEC	A	807	1	-	1/6/54/54	-
3	HEC	A	802	1,2	-	0/6/54/54	-

The worst 5 of 76 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	807	HEC	C3B-C2B	-8.83	1.31	1.40
3	A	803	HEC	C3B-C2B	-7.58	1.32	1.40
3	A	808	HEC	C3C-C2C	-7.01	1.33	1.40
3	A	804	HEC	C3B-C2B	-6.98	1.33	1.40
3	A	803	HEC	C3C-C2C	-6.72	1.33	1.40

The worst 5 of 60 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	806	HEC	CBD-CAD-C3D	-6.91	99.74	112.49
3	A	801	HEC	CBA-CAA-C2A	-6.50	100.50	112.48
3	A	804	HEC	CMB-C2B-C3B	6.25	133.17	125.82
3	A	804	HEC	CMB-C2B-C1B	-6.02	119.22	128.46
3	A	804	HEC	CBD-CAD-C3D	-5.99	101.44	112.49

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	804	HEC	C2A-CAA-CBA-CGA

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
3	A	807	HEC	C3D-CAD-CBD-CGD

There are no ring outliers.

7 monomers are involved in 28 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	808	HEC	2	0
3	A	806	HEC	1	0
3	A	803	HEC	5	0
3	A	801	HEC	7	0
3	A	804	HEC	2	0
3	A	807	HEC	1	0
3	A	802	HEC	10	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 HEC A 808

Bond lengths

Bond angles

Torsions

Rings

Ligand HEC A 806

Bond lengths

Bond angles

Torsions

Rings

Ligand HEC A 803

Bond lengths

Bond angles

Torsions

Rings

Ligand HEC A 801

Bond lengths

Bond angles

Torsions

Rings

Ligand HEC A 804

Bond lengths

Bond angles

Torsions

Rings

Ligand HEC A 805

Bond lengths

Bond angles

Torsions

Rings

Ligand HEC A 807

Bond lengths

Bond angles

Torsions

Rings

Ligand HEC A 802

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	436/443 (98%)	-0.17	11 (2%)	57	55	23, 35, 55, 79	0

The worst 5 of 11 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	202	GLY	10.9
1	A	439	PRO	4.2
1	A	437	GLN	4.1
1	A	146	GLY	3.8
1	A	99	LEU	3.6

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	HEC	A	801	43/43	0.94	0.13	24,32,46,59	0
3	HEC	A	802	43/43	0.94	0.12	24,32,47,61	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	HEC	A	803	43/43	0.94	0.15	22,29,45,49	0
3	HEC	A	806	43/43	0.94	0.14	17,29,45,56	0
3	HEC	A	807	43/43	0.94	0.13	17,27,43,45	0
2	SCN	A	810	3/3	0.95	0.15	39,39,41,42	0
3	HEC	A	808	43/43	0.96	0.12	21,28,44,53	0
3	HEC	A	805	43/43	0.98	0.11	17,25,35,40	0
3	HEC	A	804	43/43	0.98	0.13	20,24,33,34	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.

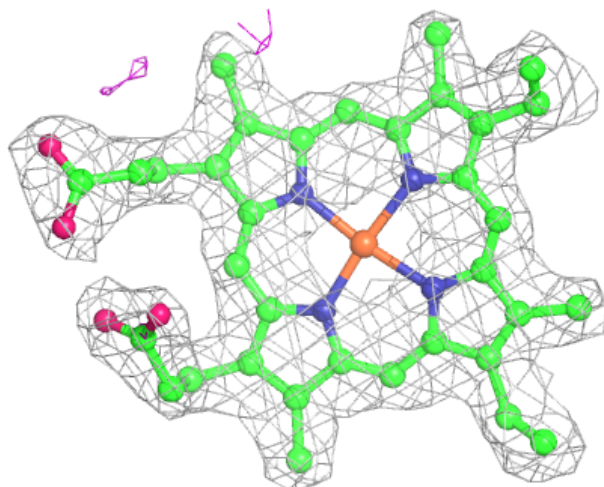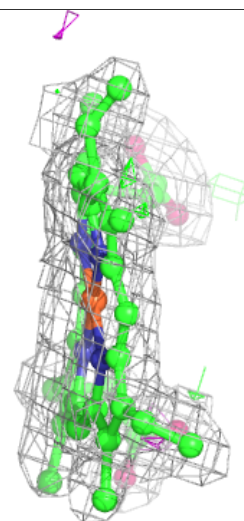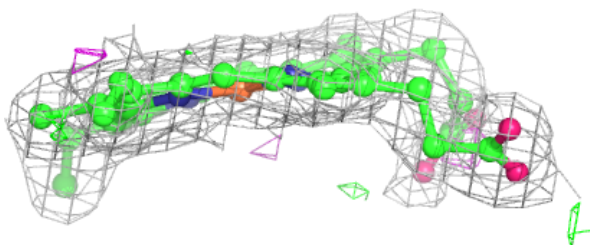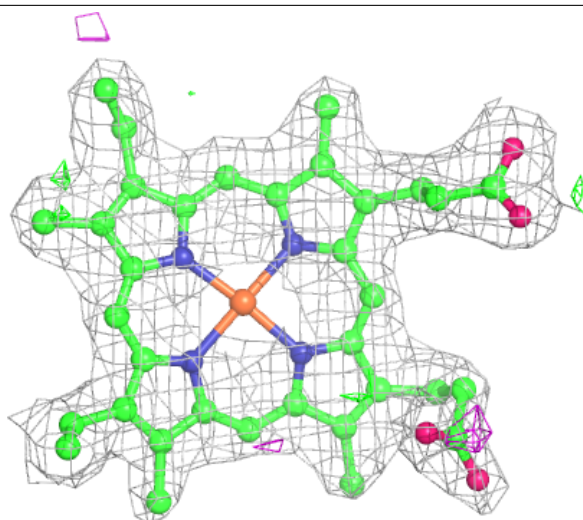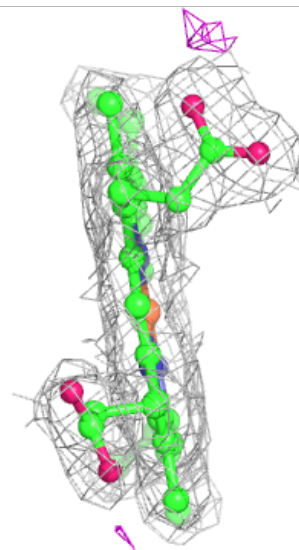**Electron density around HEC A 801:**

2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative)
and green (positive)
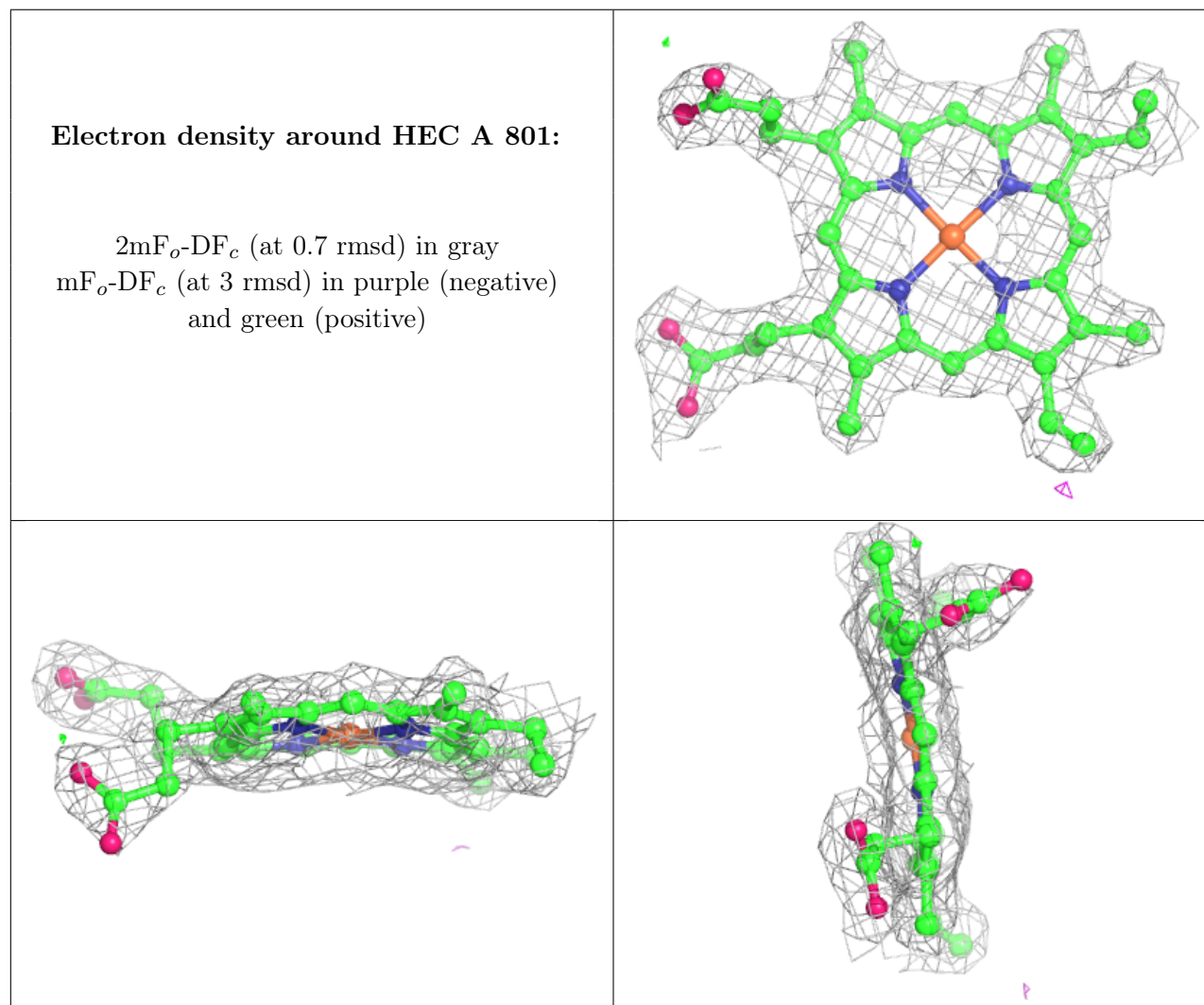
Electron density around HEC A 802:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)
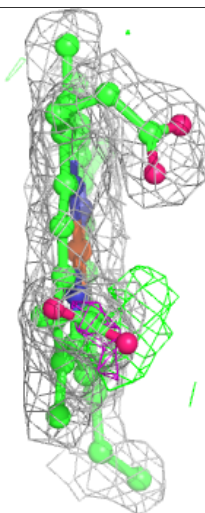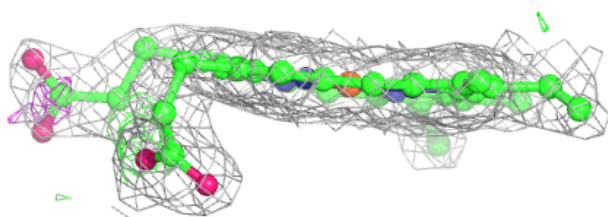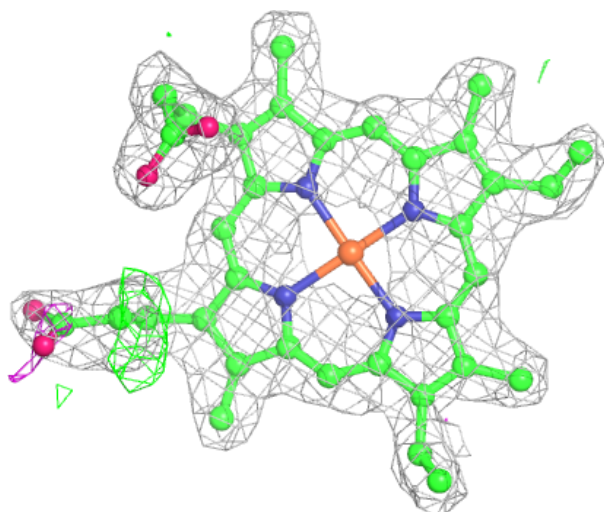
Electron density around HEC A 803:

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

Electron density around HEC A 806:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
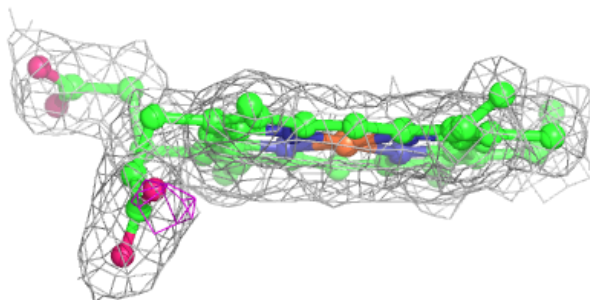$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
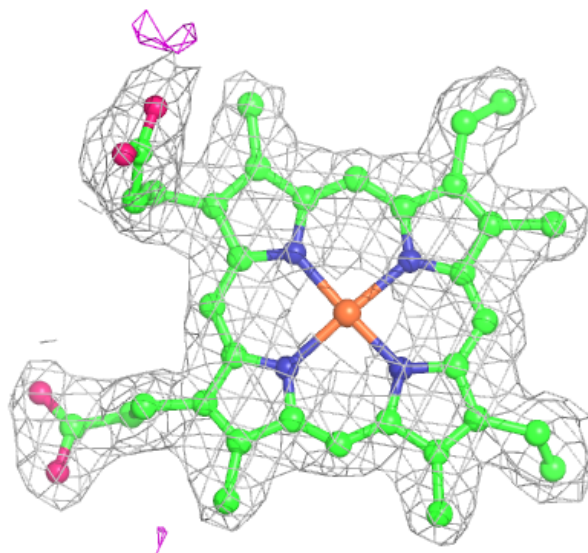and green (positive)
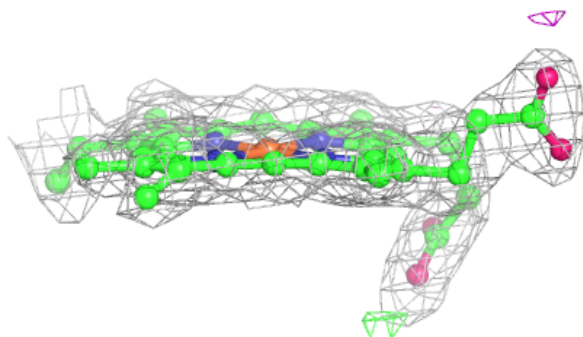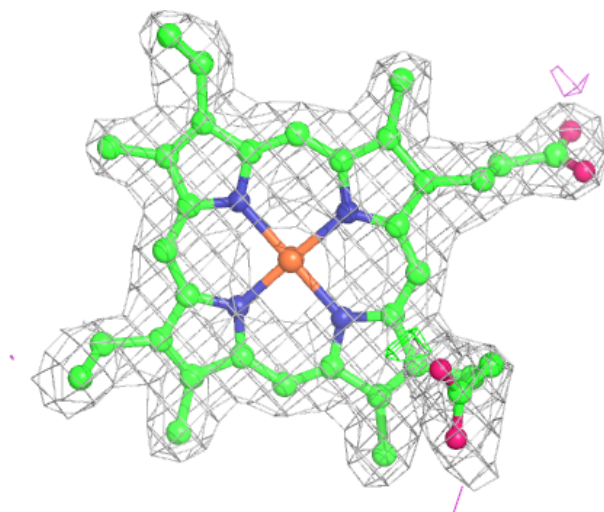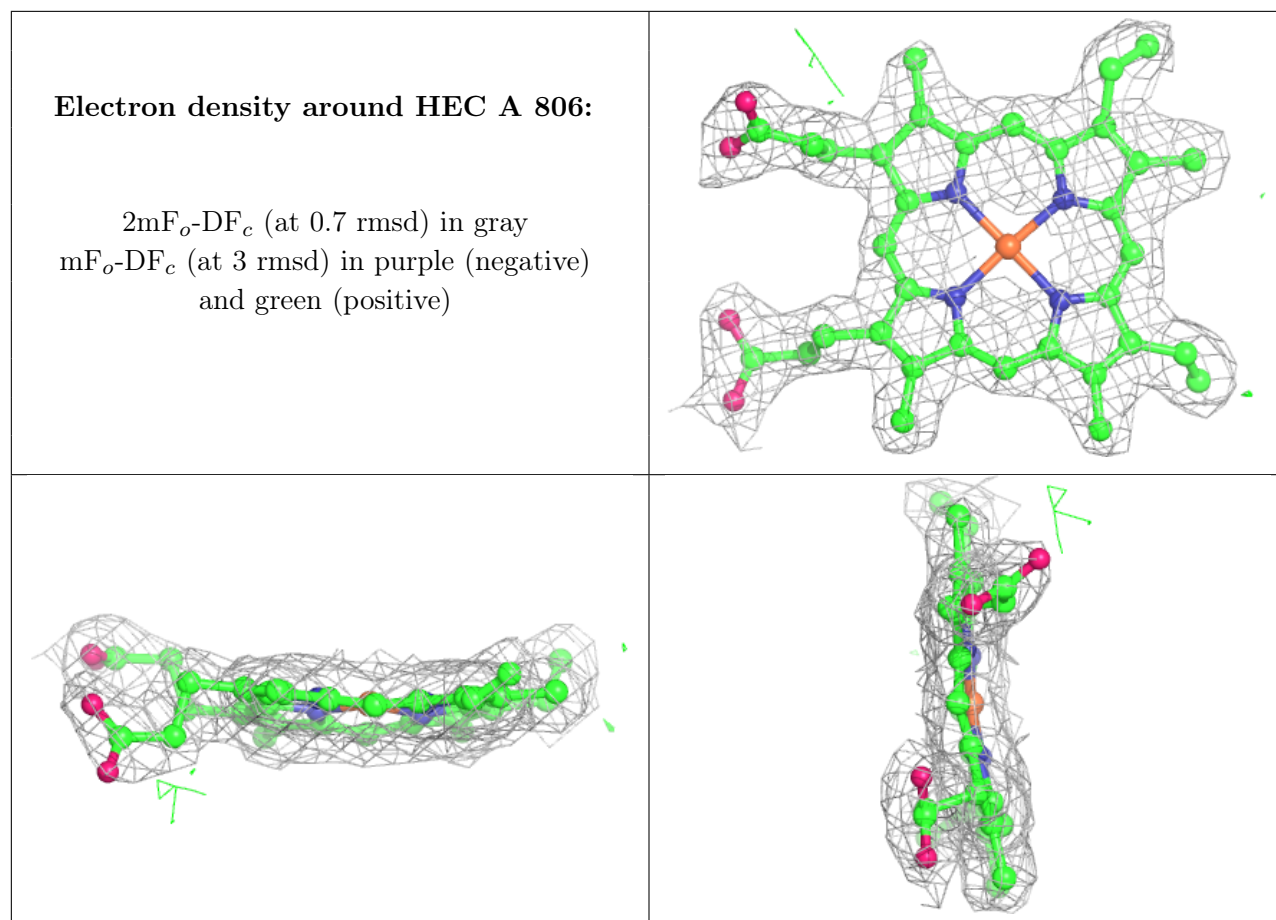
Electron density around HEC A 807:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)
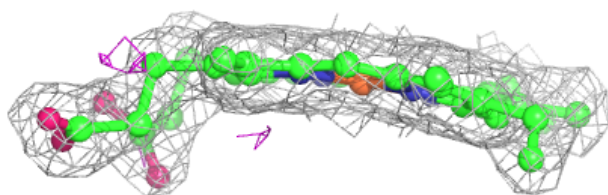
Electron density around HEC A 808:

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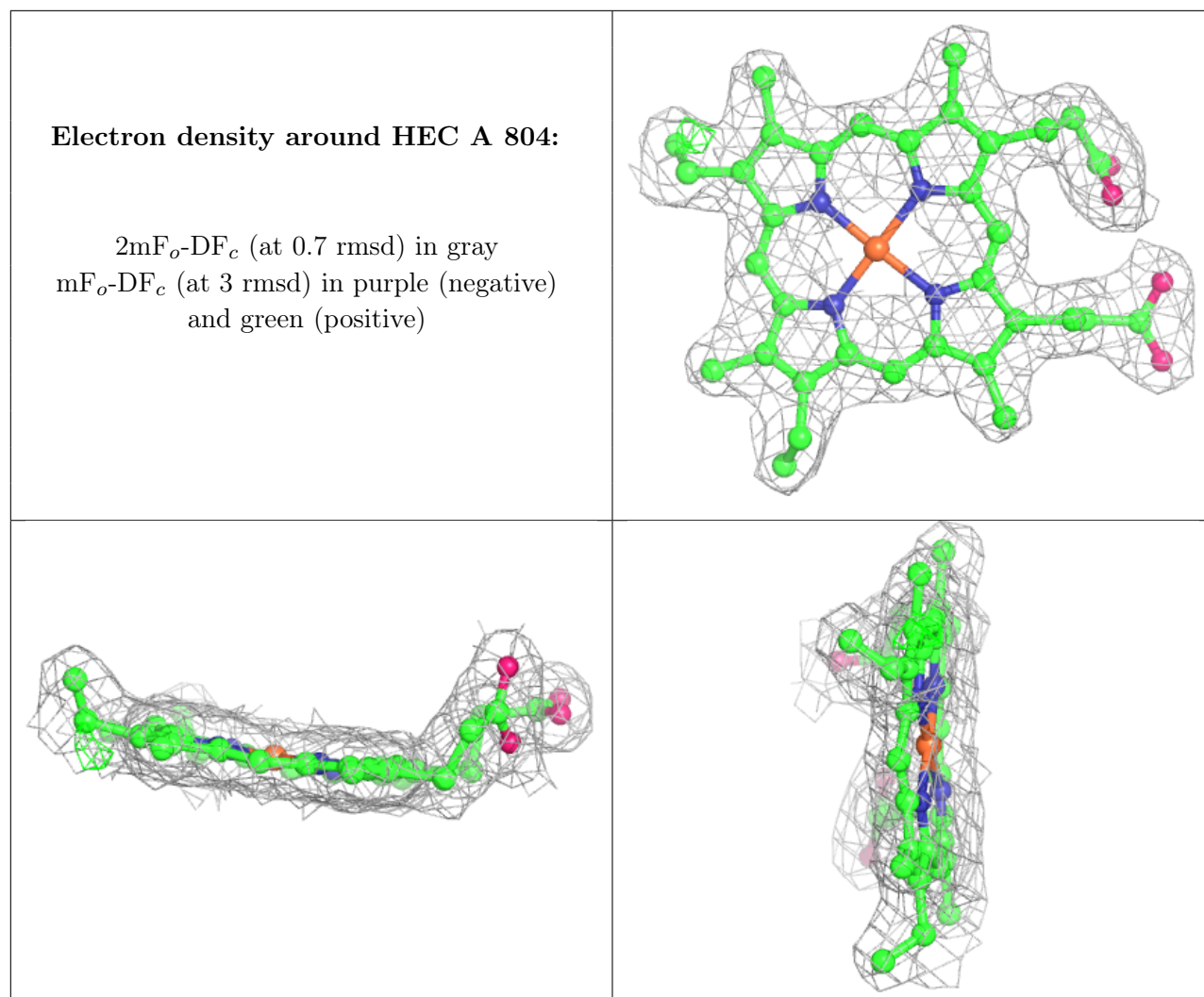and green (positive)

Electron density around HEC A 805:

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

Electron density around HEC A 804:

$2mF_o\text{-}DF_c$ (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.