



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 23, 2020 – 04:01 pm BST

PDB ID : 2YEV  
Title : Structure of caa3-type cytochrome oxidase  
Authors : Lyons, J.A.; Aragao, D.; Soulimane, T.; Caffrey, M.  
Deposited on : 2011-03-31  
Resolution : 2.36 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](mailto: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

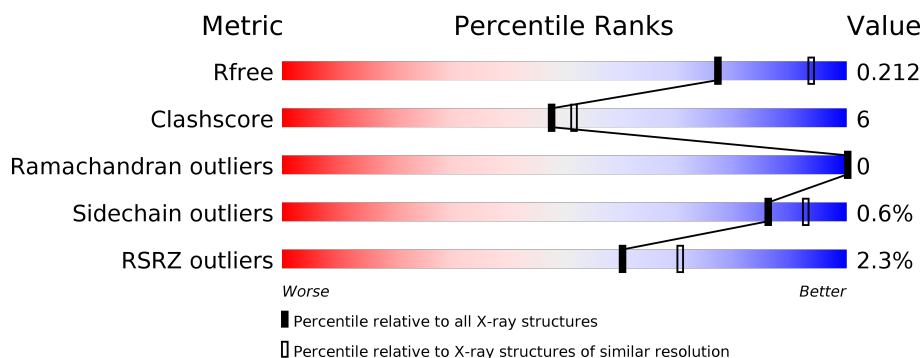
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.36 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	1164 (2.36-2.36)
Clashscore	141614	1232 (2.36-2.36)
Ramachandran outliers	138981	1211 (2.36-2.36)
Sidechain outliers	138945	1212 (2.36-2.36)
RSRZ outliers	127900	1150 (2.36-2.36)

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  $\geq 3$ , 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	791	<div> <div>2%</div> <div> <div></div> <div>85%</div> <div>13%</div> <div>•</div> </div> </div>
1	D	791	<div> <div>2%</div> <div> <div></div> <div>85%</div> <div>14%</div> <div>•</div> </div> </div>
2	B	337	<div> <div>0%</div> <div> <div></div> <div>84%</div> <div>11%</div> <div>5%</div> </div> </div>
2	E	337	<div> <div>4%</div> <div> <div></div> <div>83%</div> <div>11%</div> <div>•</div> <div>5%</div> </div> </div>
3	C	66	<div> <div>2%</div> <div> <div></div> <div>88%</div> <div>9%</div> <div>•</div> </div> </div>
3	F	66	<div> <div>5%</div> <div> <div></div> <div>85%</div> <div>11%</div> <div>5%</div> </div> </div>

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
8	7E8	A	1300	X	-	-	-

## 2 Entry composition

There are 14 unique types of molecules in this entry. The entry contains 19587 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 OXIDASE POLYPEPTIDE I+III.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	780	Total	C	N	O	S	0	1	0
			6281	4244	992	1022	23			
1	D	780	Total	C	N	O	S	0	0	0
			6276	4241	992	1020	23			

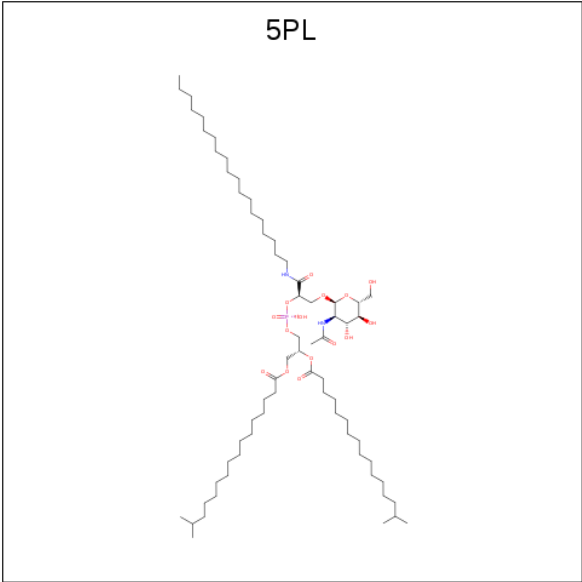
- Molecule 2 is a protein called CYTOCHROME C OXIDASE SUBUNIT 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	319	Total	C	N	O	S	0	0	0
			2507	1637	422	439	9			
2	E	319	Total	C	N	O	S	0	0	0
			2507	1637	422	439	9			

- Molecule 3 is a protein called CAA3-TYPE CYTOCHROME OXIDASE SUBUNIT IV.

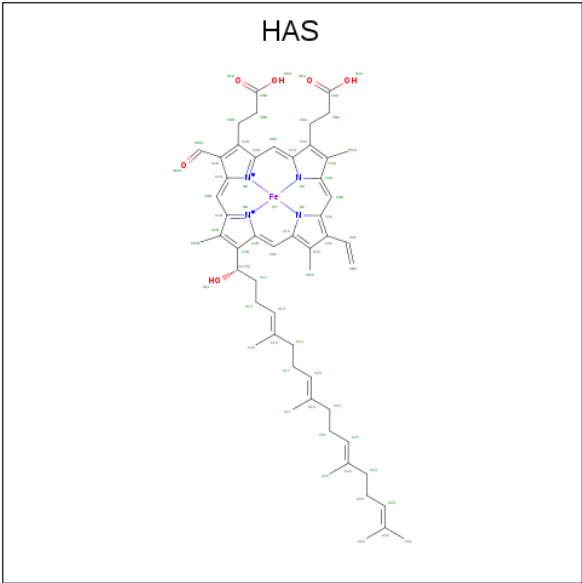
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	64	Total	C	N	O	S	0	0	0
			502	344	76	79	3			
3	F	63	Total	C	N	O	S	0	0	0
			492	338	73	78	3			

- Molecule 4 is (1R,4S,6R)-6-({[2-(ACETYLAMINO)-2-DEOXY-ALPHA-D-GLUCOPYRANOSYL]OXY}METHYL)-4-HYDROXY-1-{{[(15-METHYLHEXADECANOYL)OXY]METHYL}-4-OXIDO-7-OXO-3,5-DIOXA-8-AZA-4-PHOSPHAHEPTACOS-1-YL 15-METHYLHEXADECANOATE (three-letter code: 5PL) (formula: C<sub>67</sub>H<sub>129</sub>N<sub>2</sub>O<sub>15</sub>P).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	N	O	P	0	0
			85	67	2	15	1		
4	D	1	Total	C	N	O	P	0	0
			85	67	2	15	1		

- Molecule 5 is HEME-AS (three-letter code: HAS) (formula: C<sub>54</sub>H<sub>64</sub>FeN<sub>4</sub>O<sub>6</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	A	1	Total	C	Fe	N	O	0	0
			65	54	1	4	6		
5	A	1	Total	C	Fe	N	O	0	0
			65	54	1	4	6		

Continued on next page...

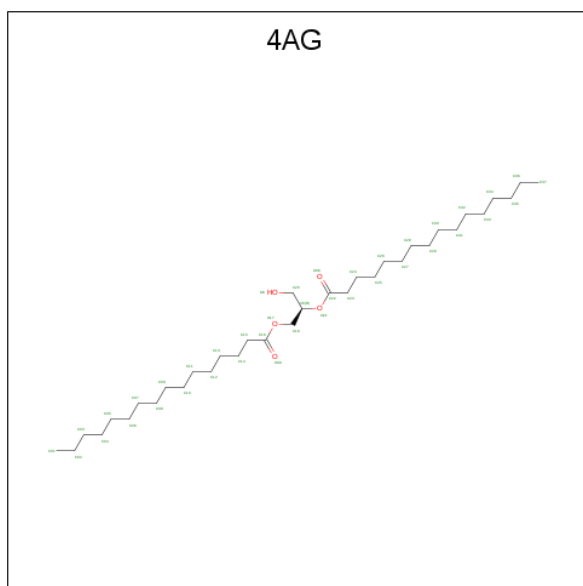
*Continued from previous page...*

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	D	1	Total	C	Fe	N	O	
			65	54	1	4	6	
5	D	1	Total	C	Fe	N	O	
			65	54	1	4	6	

- Molecule 6 is COPPER (II) ION (three-letter code: CU) (formula: Cu).

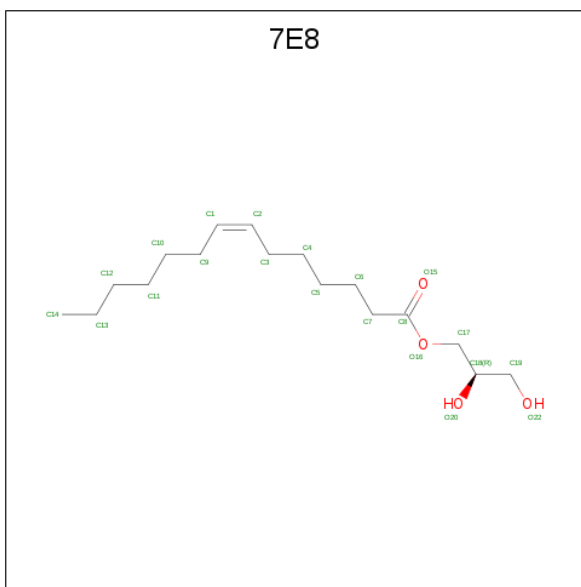
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	1	Total	Cu		
			1	1	0	0
6	D	1	Total	Cu		
			1	1	0	0

- Molecule 7 is (2R)-3-HYDROXYPROPANE-1,2-DIYL DIHEXADECANOATE (three-letter code: 4AG) (formula: C<sub>35</sub>H<sub>68</sub>O<sub>5</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total	C	O		
			40	35	5	0	0

- Molecule 8 is (2R)-2,3-DIHYDROXYPROPYL (7Z)-TETRADEC-7-ENOATE (three-letter code: 7E8) (formula: C<sub>17</sub>H<sub>32</sub>O<sub>4</sub>).

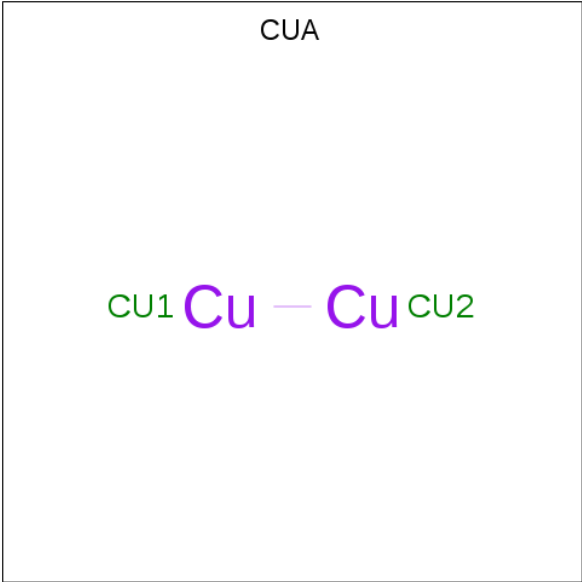


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	A	1	Total	C	O	0	0
			21	17	4		
8	A	1	Total	C	O	0	0
			21	17	4		

- Molecule 9 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

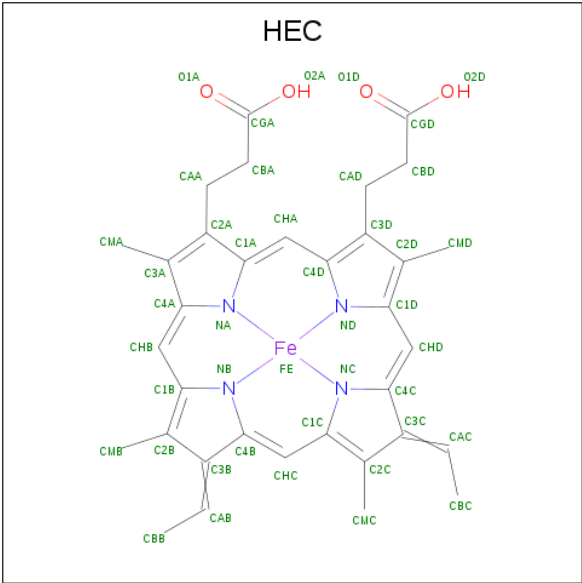
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	1	Total	Mg	0	0
			1	1		
9	D	1	Total	Mg	0	0
			1	1		

- Molecule 10 is DINUCLEAR COPPER ION (three-letter code: CUA) (formula: Cu<sub>2</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
10	B	1	Total	Cu			0	0
			2	2				
10	E	1	Total	Cu			0	0
			2	2				

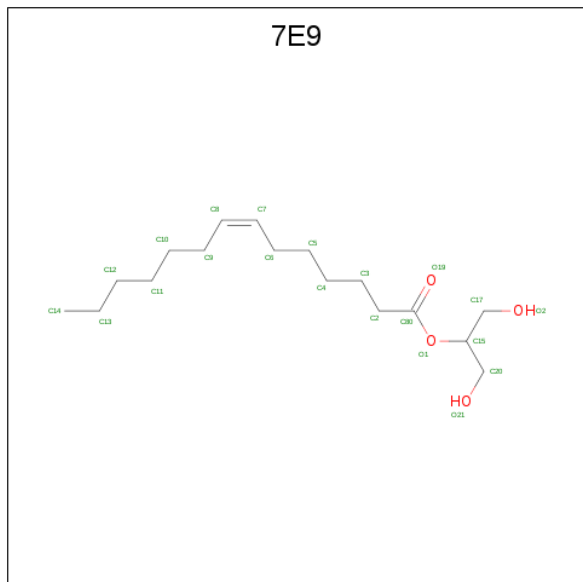
- Molecule 11 is HEME C (three-letter code: HEC) (formula: C<sub>34</sub>H<sub>34</sub>FeN<sub>4</sub>O<sub>4</sub>).



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



- Molecule 12 is 1,3-DIHYDROXYPROPAN-2-YL (Z)-TETRADEC-7-ENOATE (three-letter code: 7E9) (formula: C<sub>17</sub>H<sub>32</sub>O<sub>4</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
12	B	1	Total	C	O	0	0
			21	17	4		

- Molecule 13 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
13	D	1	Total	Cl	0	0
			1	1		
13	C	1	Total	Cl	0	0
			1	1		
13	F	1	Total	Cl	0	0
			1	1		

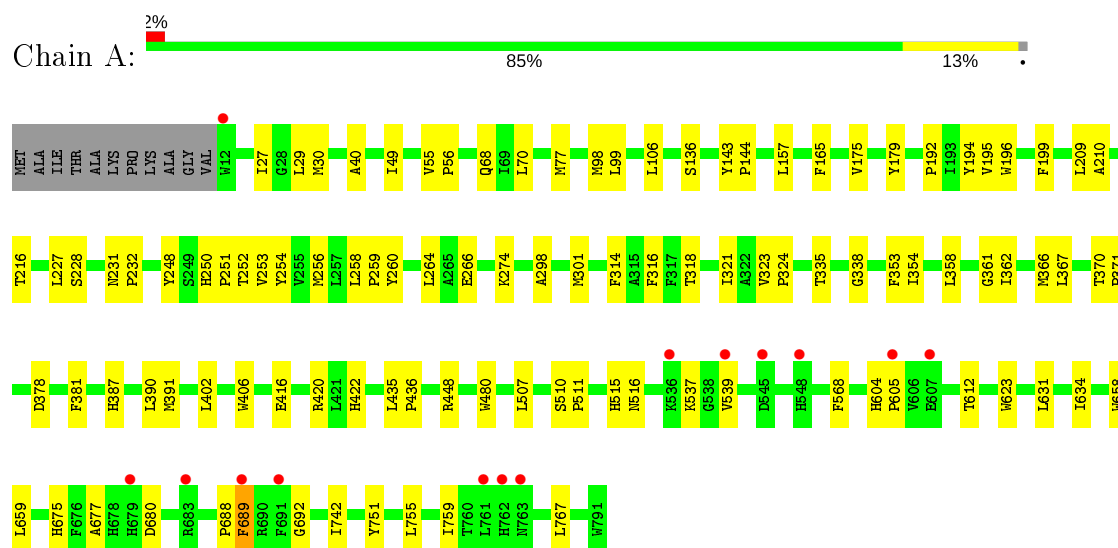
- Molecule 14 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
14	A	157	Total	O	0	0
			157	157		
14	B	90	Total	O	0	0
			90	90		
14	D	111	Total	O	0	0
			111	111		
14	E	34	Total	O	0	0
			34	34		

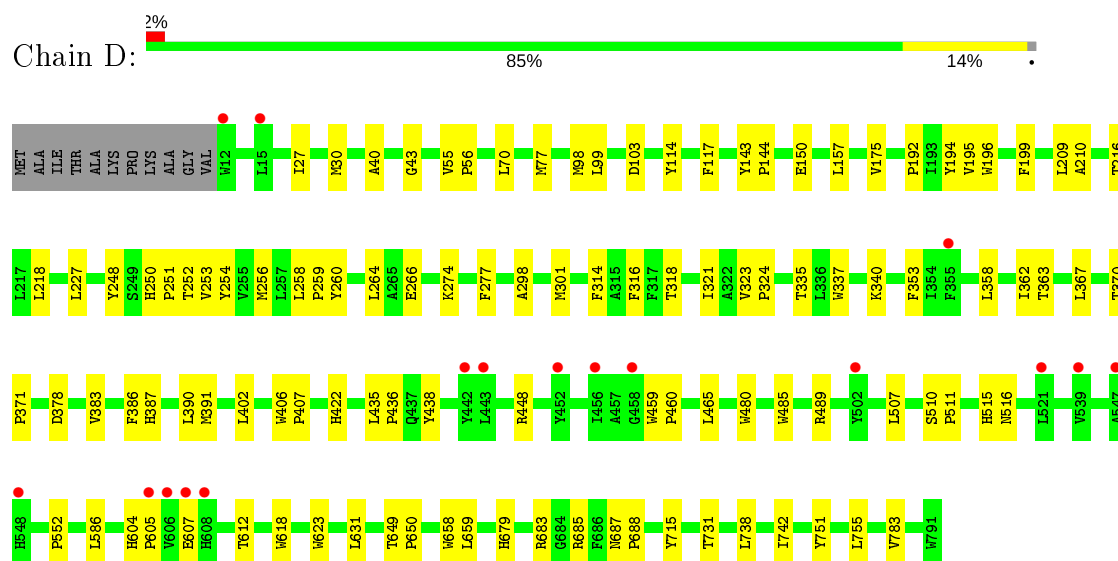
### 3 Residue-property plots [i](#)

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.

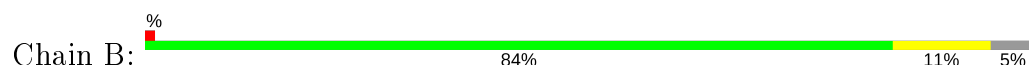
#### • Molecule 1: CYTOCHROME C OXIDASE POLYPEPTIDE I+III

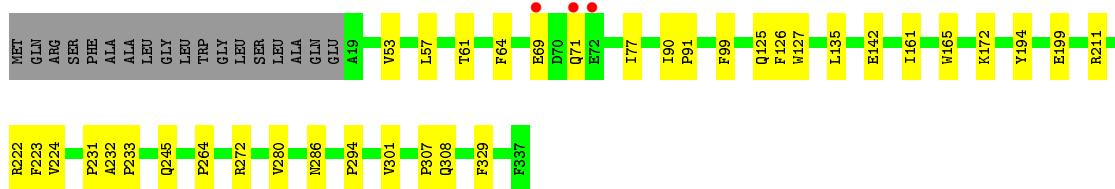


#### • Molecule 1: CYTOCHROME C OXIDASE POLYPEPTIDE I+III

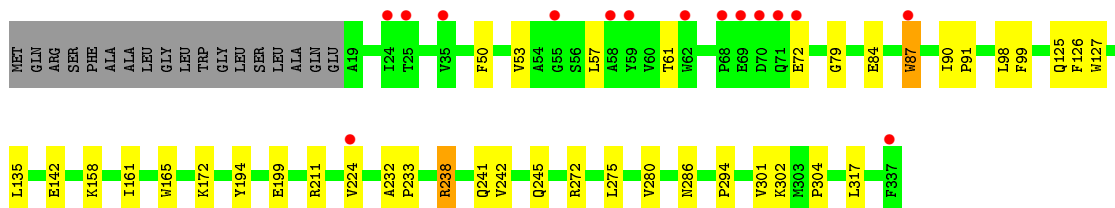
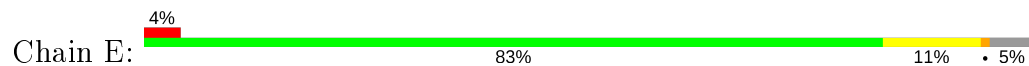


#### • Molecule 2: CYTOCHROME C OXIDASE SUBUNIT 2

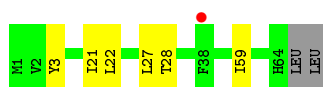
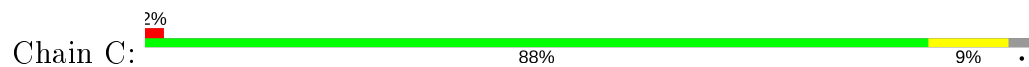




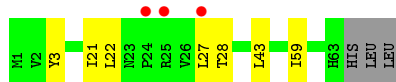
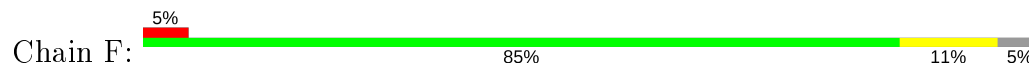
• Molecule 2: CYTOCHROME C OXIDASE SUBUNIT 2



• Molecule 3: CAA3-TYPE CYTOCHROME OXIDASE SUBUNIT IV



• Molecule 3: CAA3-TYPE CYTOCHROME OXIDASE SUBUNIT IV



## 4 Data and refinement statistics

Property	Value	Source
Space group	I 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	127.25Å 76.03Å 300.27Å 90.00° 92.21° 90.00°	Depositor
Resolution (Å)	77.18 – 2.36 115.49 – 2.36	Depositor EDS
% Data completeness (in resolution range)	100.0 (77.18-2.36) 100.0 (115.49-2.36)	Depositor EDS
$R_{merge}$	0.16	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.19 (at 2.37Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE: 1.6.4_486)	Depositor
R, $R_{free}$	0.171 , 0.218 0.163 , 0.212	Depositor DCC
$R_{free}$ test set	5970 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	50.2	Xtriage
Anisotropy	0.263	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 56.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.017 for h,-k,-l	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	19587	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	58.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.82% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: 5PL, 7E8, CL, MG, HAS, 4AG, CUA, HEC, 7E9, FME, CU

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	0.32	0/6523	0.44	0/8921
1	D	0.31	0/6515	0.44	0/8910
2	B	0.32	0/2586	0.46	0/3523
2	E	0.31	0/2586	0.46	0/3523
3	C	0.26	0/502	0.38	0/683
3	F	0.25	0/491	0.39	0/668
All	All	0.31	0/19203	0.44	0/26228

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 [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	A	6281	0	6210	82	0
1	D	6276	0	6206	97	0
2	B	2507	0	2473	26	0
2	E	2507	0	2473	27	0
3	C	502	0	545	4	0
3	F	492	0	538	6	0
4	A	85	0	128	5	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	D	85	0	128	4	0
5	A	130	0	124	18	0
5	D	130	0	124	23	0
6	A	1	0	0	0	0
6	D	1	0	0	0	0
7	A	40	0	68	1	0
8	A	42	0	64	2	0
9	A	1	0	0	0	0
9	D	1	0	0	0	0
10	B	2	0	0	0	0
10	E	2	0	0	0	0
11	B	43	0	30	3	0
11	E	43	0	30	3	0
12	B	21	0	32	0	0
13	C	1	0	0	0	0
13	D	1	0	0	0	0
13	F	1	0	0	0	0
14	A	157	0	0	4	0
14	B	90	0	0	1	0
14	D	111	0	0	0	0
14	E	34	0	0	0	0
All	All	19587	0	19173	232	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.

The worst 5 of 232 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:77:MET:HB3	5:A:1015:HAS:CAC	1.99	0.93
1:D:77:MET:HB3	5:D:1015:HAS:CAC	2.00	0.91
1:D:367:LEU:CB	5:D:1016:HAS:HMD	2.15	0.77
1:D:70:LEU:HD13	5:D:1015:HAS:HBD1	1.66	0.77
1:A:77:MET:HB3	5:A:1015:HAS:HAC	1.68	0.73

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	779/791 (98%)	762 (98%)	17 (2%)	0	100	100
1	D	778/791 (98%)	759 (98%)	19 (2%)	0	100	100
2	B	317/337 (94%)	308 (97%)	9 (3%)	0	100	100
2	E	317/337 (94%)	307 (97%)	10 (3%)	0	100	100
3	C	62/66 (94%)	62 (100%)	0	0	100	100
3	F	61/66 (92%)	61 (100%)	0	0	100	100
All	All	2314/2388 (97%)	2259 (98%)	55 (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	A	640/646 (99%)	637 (100%)	3 (0%)	88	94
1	D	639/646 (99%)	638 (100%)	1 (0%)	93	97
2	B	261/274 (95%)	259 (99%)	2 (1%)	81	89
2	E	261/274 (95%)	256 (98%)	5 (2%)	57	68
3	C	51/53 (96%)	51 (100%)	0	100	100
3	F	50/53 (94%)	50 (100%)	0	100	100
All	All	1902/1946 (98%)	1891 (99%)	11 (1%)	86	93

5 of 11 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	272	ARG
1	D	378	ASP
2	E	194	TYR
2	B	194	TYR
2	E	87	TRP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
3	C	63	HIS
3	F	63	HIS
1	D	205	ASN
2	B	245	GLN
1	D	171	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

2 non-standard protein/DNA/RNA residues 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	FME	F	1	3	8,9,10	0.38	0	7,9,11	0.69	0
3	FME	C	1	3	8,9,10	0.36	0	7,9,11	1.31	1 (14%)

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	FME	F	1	3	-	0/7/9/11	-
3	FME	C	1	3	-	0/7/9/11	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
3	C	1	FME	CA-N-CN	-2.92	118.33	122.82

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 21 ligands modelled in this entry, 7 are monoatomic - leaving 14 for Mogul analysis.

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	$\# Z  > 2$
8	7E8	A	1301	-	20,20,20	1.01	1 (5%)	21,21,21	1.04	2 (9%)
12	7E9	B	701	-	20,20,20	1.06	1 (5%)	21,21,21	1.20	1 (4%)
5	HAS	D	1016	1,14	56,72,72	4.99	22 (39%)	50,109,109	4.10	25 (50%)
4	5PL	A	900	-	84,85,85	0.81	2 (2%)	96,101,101	1.02	4 (4%)
5	HAS	A	1015	1	56,72,72	4.94	21 (37%)	50,109,109	4.02	25 (50%)
8	7E8	A	1300	-	20,20,20	1.05	1 (5%)	21,21,21	0.96	1 (4%)
11	HEC	B	587	2	26,50,50	2.43	3 (11%)	18,82,82	2.47	4 (22%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	5PL	D	900	-	84,85,85	0.81	2 (2%)	96,101,101	1.09	9 (9%)
10	CUA	B	585	2	0,1,1	0.00	-	-		
10	CUA	E	585	2	0,1,1	0.00	-	-		
7	4AG	A	1200	-	39,39,39	1.06	2 (5%)	41,41,41	1.21	3 (7%)
5	HAS	D	1015	1	56,72,72	4.95	22 (39%)	50,109,109	3.99	25 (50%)
11	HEC	E	587	2	26,50,50	2.45	3 (11%)	18,82,82	2.49	6 (33%)
5	HAS	A	1016	1,14	56,72,72	4.94	22 (39%)	50,109,109	4.05	24 (48%)

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
8	7E8	A	1301	-	-	6/20/20/20	-
12	7E9	B	701	-	-	7/21/21/21	-
5	HAS	D	1016	1,14	-	4/35/122/122	-
4	5PL	A	900	-	-	30/87/107/107	0/1/1/1
5	HAS	A	1015	1	-	9/35/122/122	-
5	HAS	D	1015	1	-	8/35/122/122	-
11	HEC	B	587	2	-	0/6/54/54	-
4	5PL	D	900	-	-	26/87/107/107	0/1/1/1
8	7E8	A	1300	-	1/1/2/4	6/20/20/20	-
7	4AG	A	1200	-	-	15/41/41/41	-
11	HEC	E	587	2	-	0/6/54/54	-
5	HAS	A	1016	1,14	-	3/35/122/122	-

The worst 5 of 102 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	D	1015	HAS	C2B-C3B	13.55	1.45	1.34
5	D	1016	HAS	C2B-C3B	13.20	1.45	1.34
5	A	1016	HAS	C2B-C3B	12.98	1.45	1.34
5	A	1015	HAS	CHD-C4A	-12.58	1.38	1.51
5	A	1015	HAS	C2B-C3B	12.42	1.44	1.34

The worst 5 of 129 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	D	1016	HAS	CHB-C1D-ND	14.95	129.52	110.94
5	A	1015	HAS	CHB-C1D-ND	13.62	127.88	110.94
5	A	1016	HAS	CHB-C1D-ND	13.32	127.50	110.94
5	D	1015	HAS	CHB-C1D-ND	13.05	127.17	110.94
5	D	1015	HAS	CHB-C1B-NB	12.29	126.22	110.94

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
8	A	1300	7E8	C18

5 of 114 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
12	B	701	7E9	C17-C15-C20-O21
12	B	701	7E9	O1-C15-C20-O21
5	D	1016	HAS	C1D-C2D-CMD-OMD
5	D	1016	HAS	C3D-C2D-CMD-OMD
5	A	1015	HAS	C1D-C2D-CMD-OMD

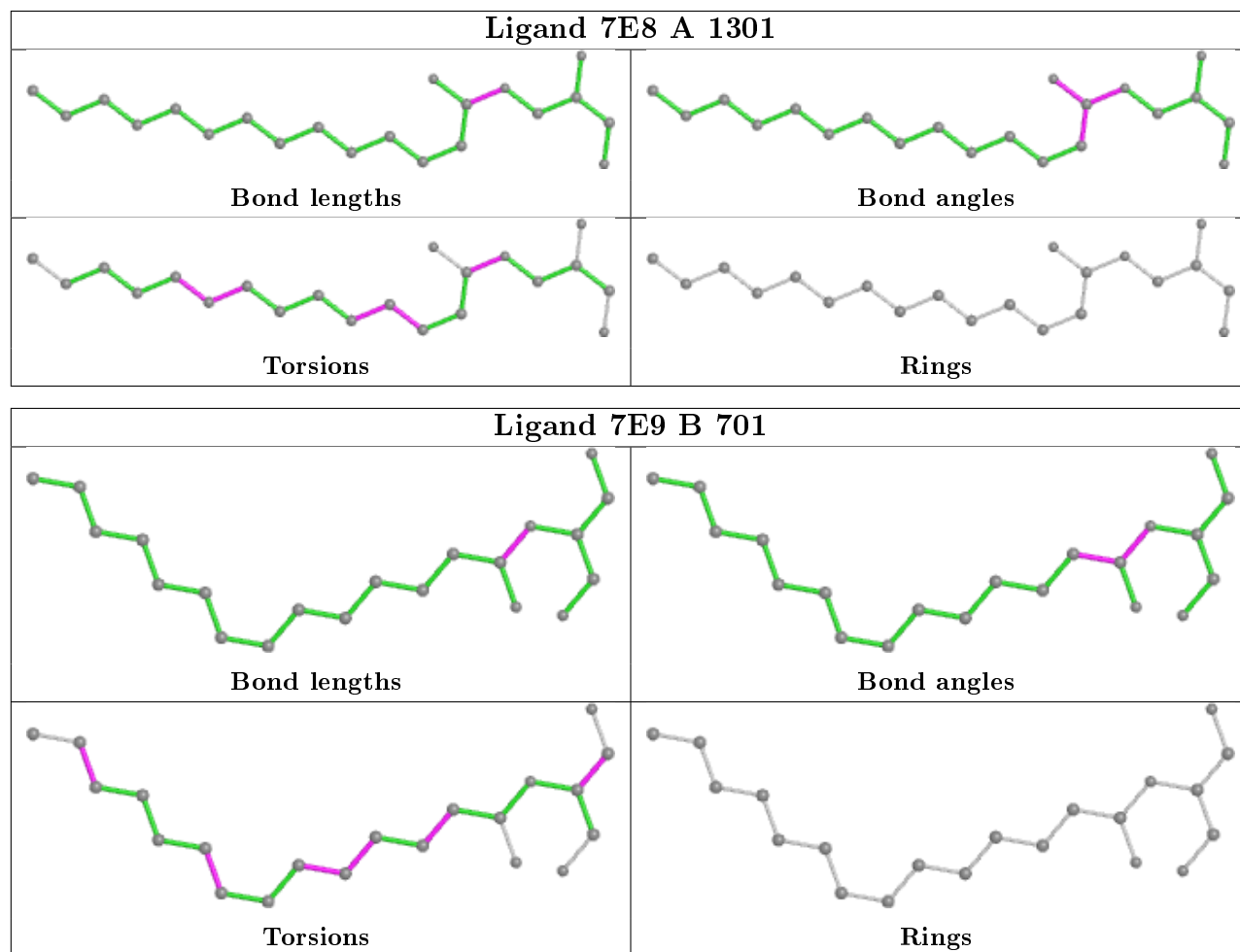
There are no ring outliers.

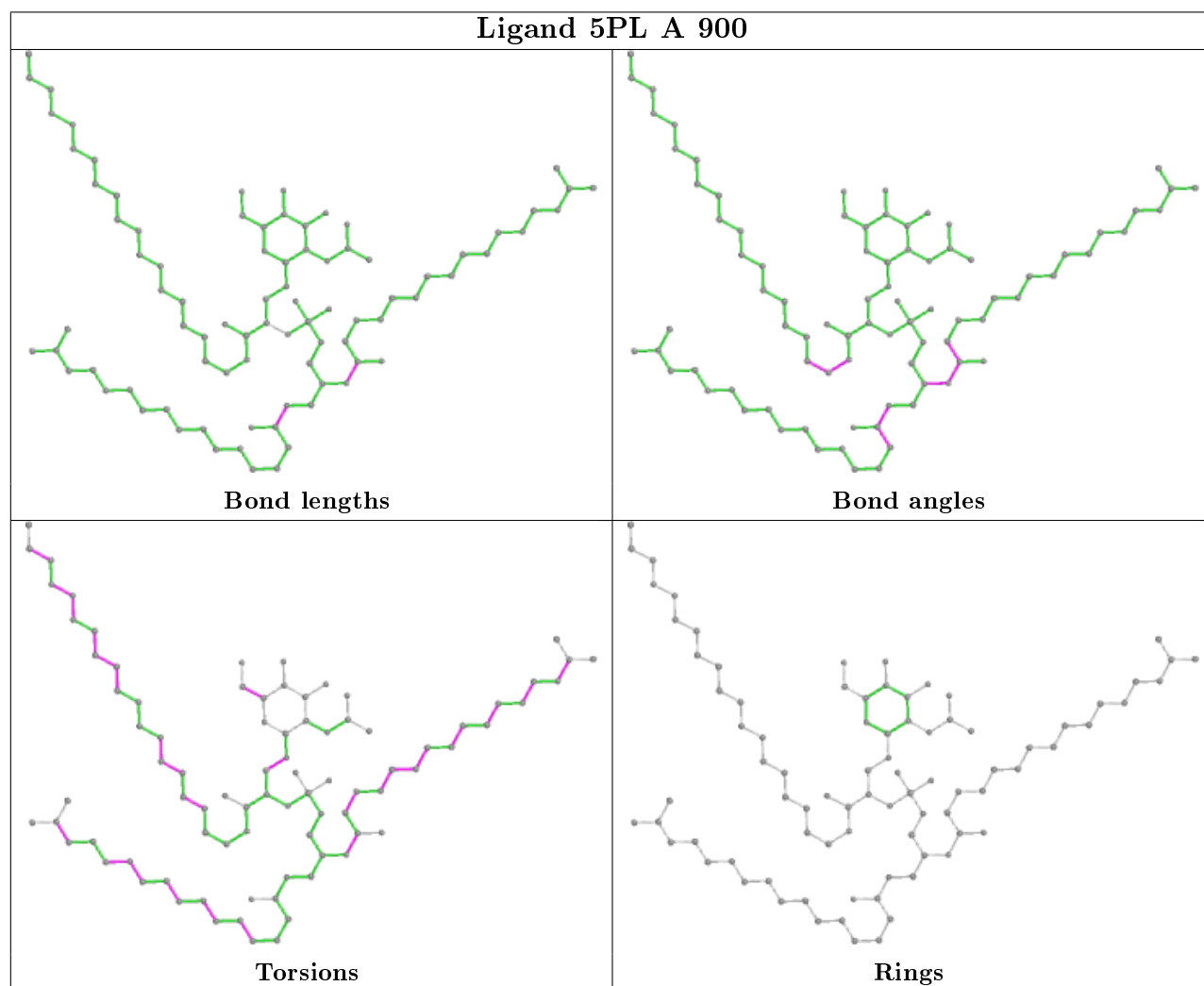
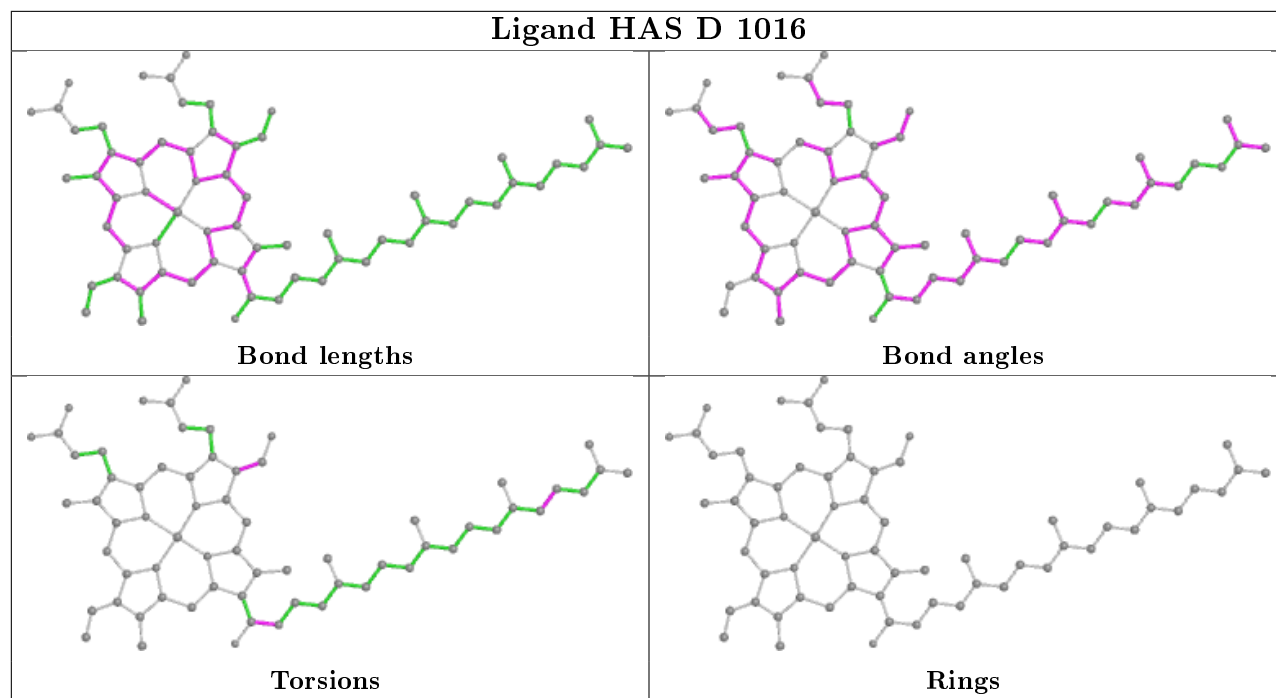
11 monomers are involved in 59 short contacts:

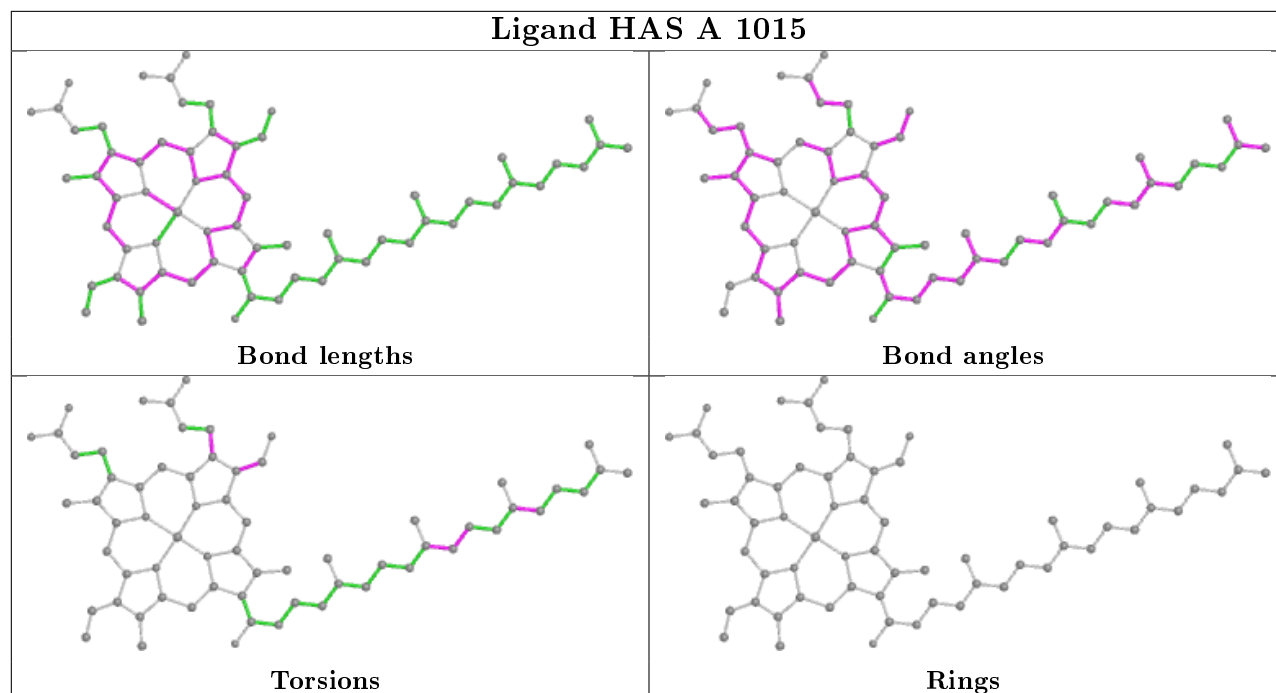
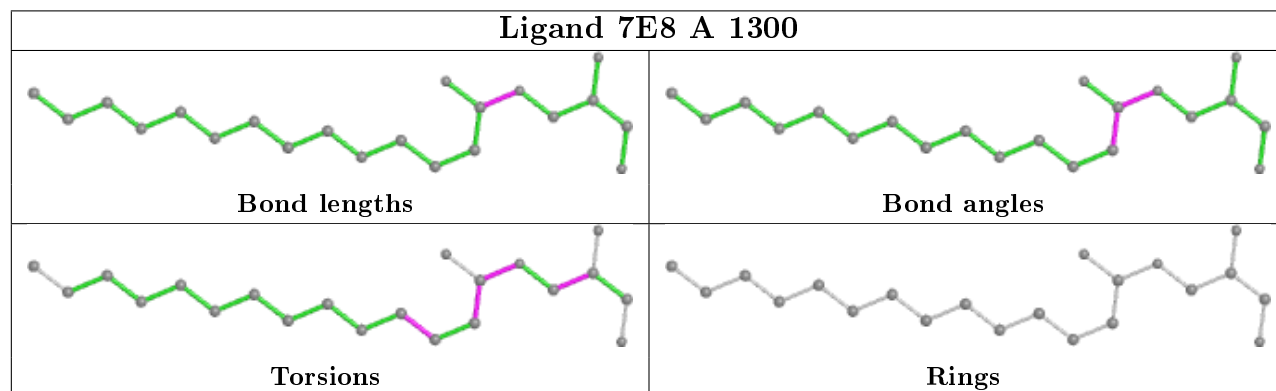
Mol	Chain	Res	Type	Clashes	Symm-Clashes
8	A	1301	7E8	1	0
5	D	1016	HAS	16	0
4	A	900	5PL	5	0
5	A	1015	HAS	9	0
8	A	1300	7E8	1	0
11	B	587	HEC	3	0
4	D	900	5PL	4	0
7	A	1200	4AG	1	0
5	D	1015	HAS	7	0
11	E	587	HEC	3	0
5	A	1016	HAS	9	0

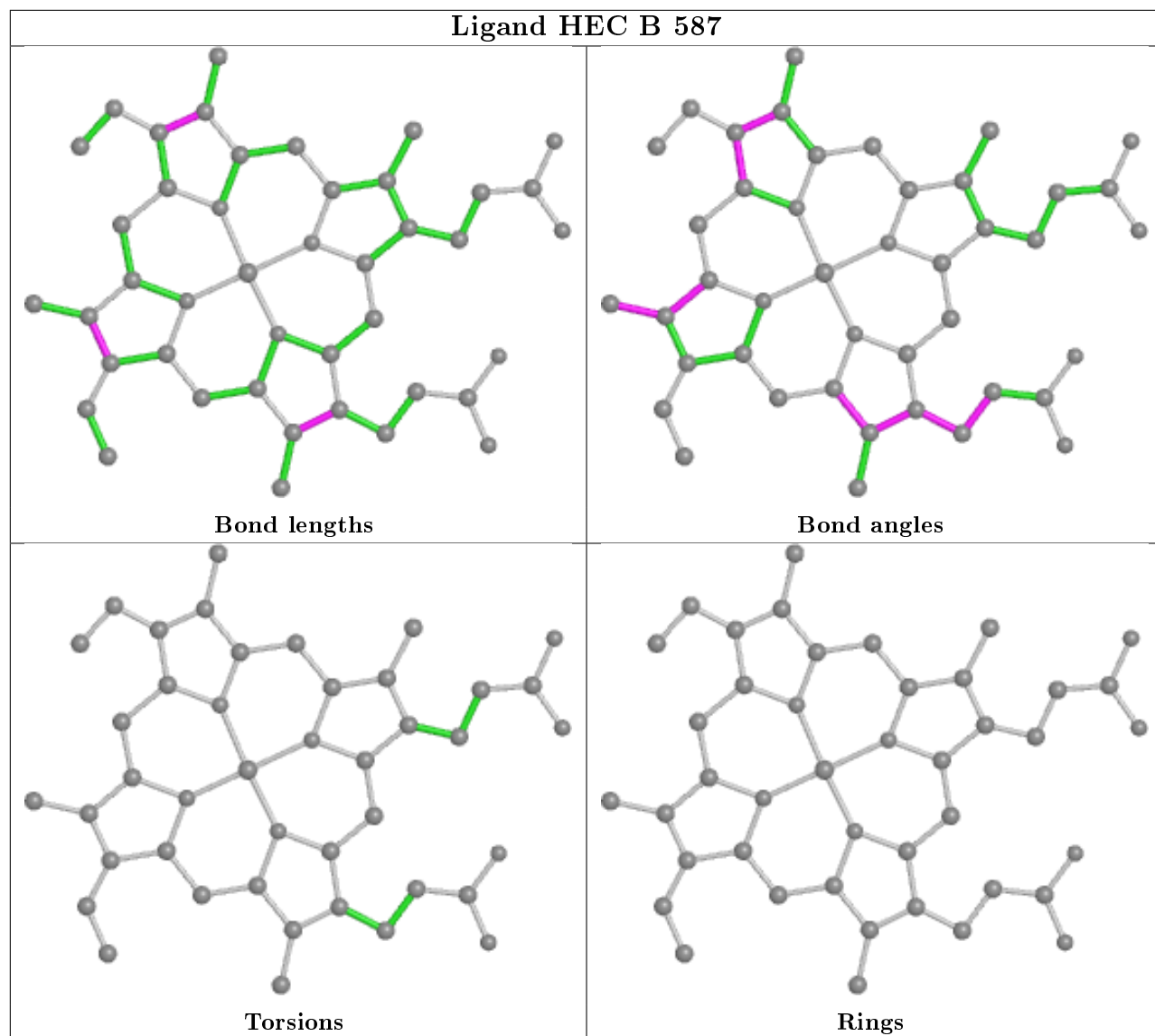
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included.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

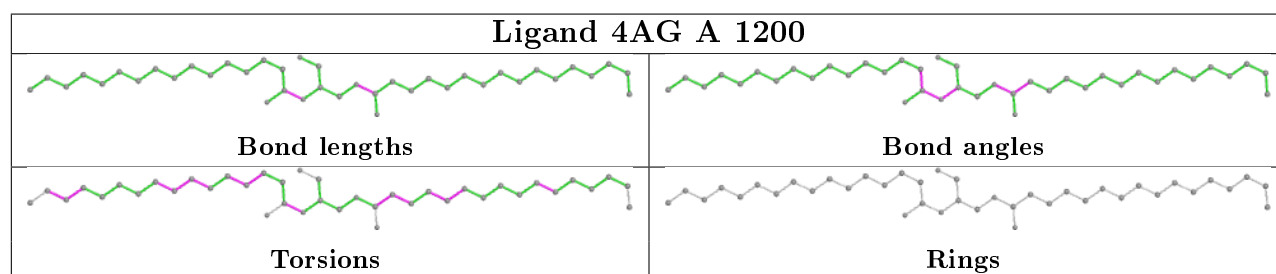
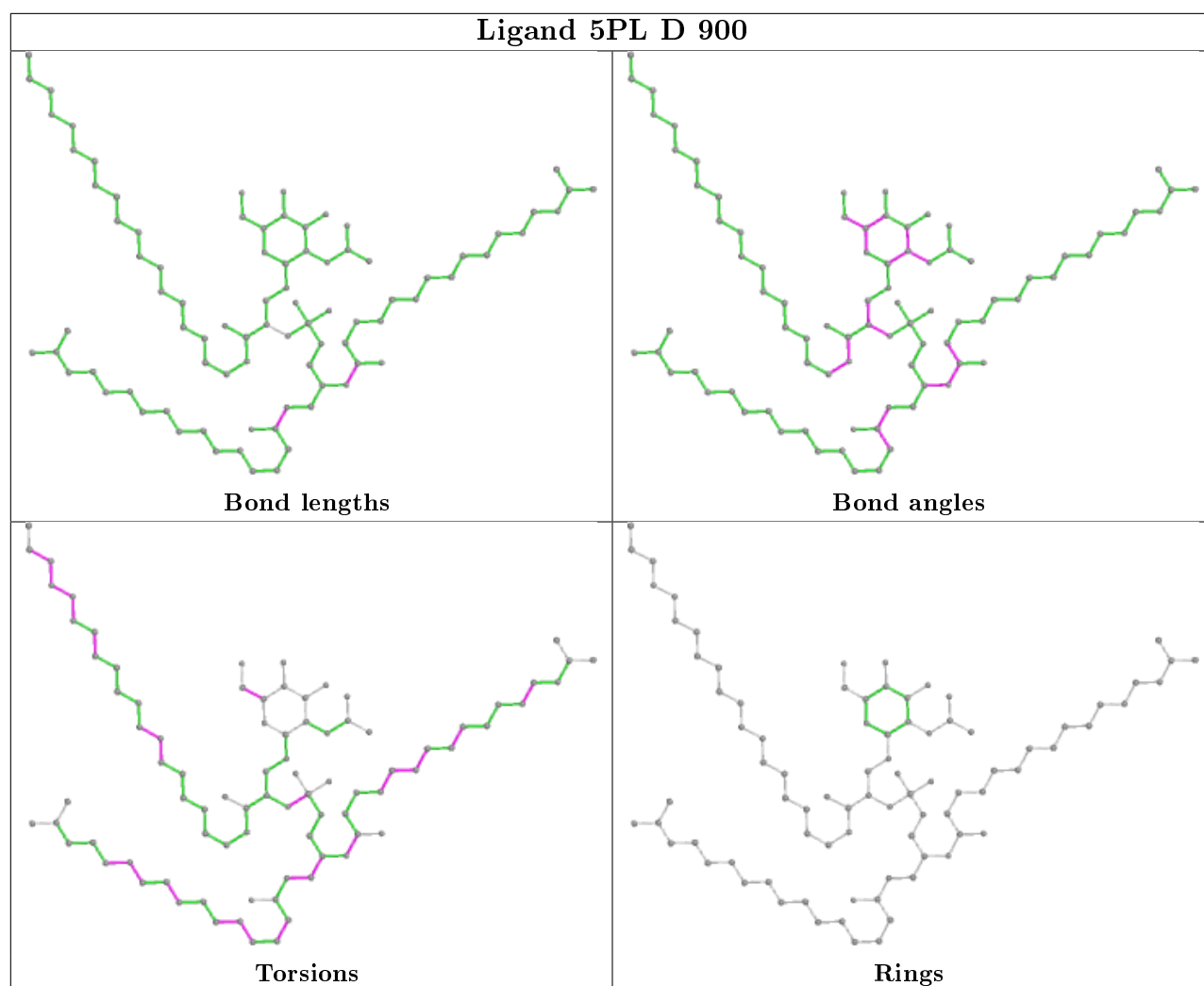
in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.



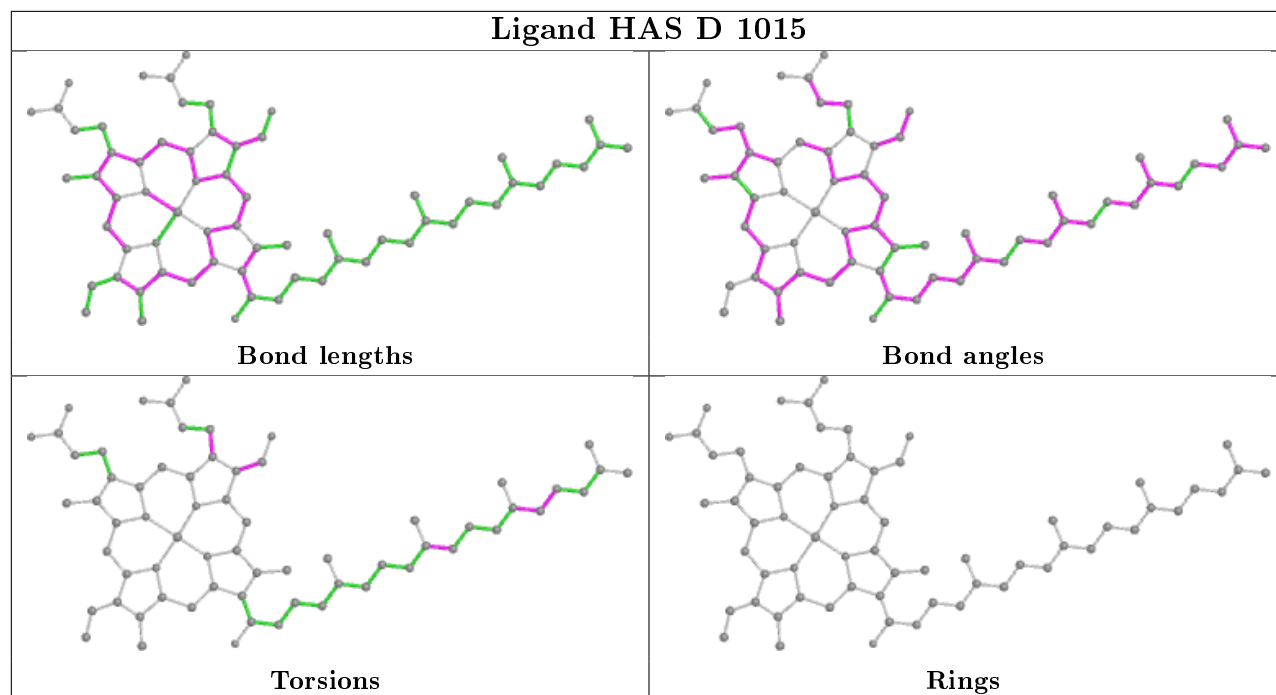


**Ligand HAS A 1015****Ligand 7E8 A 1300**

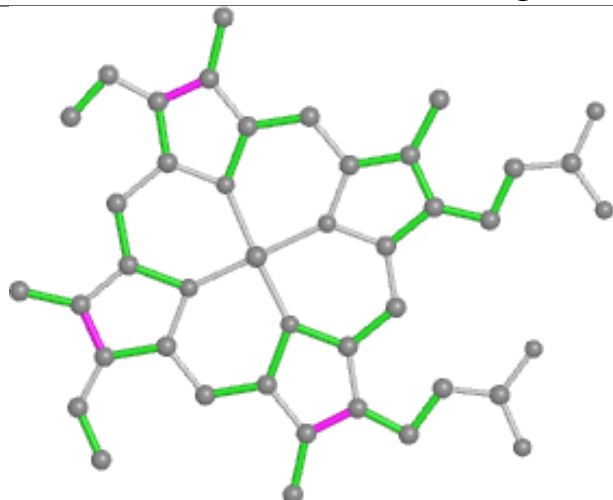




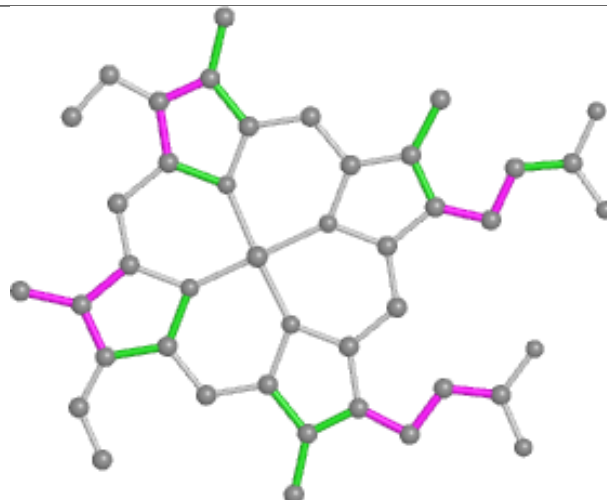




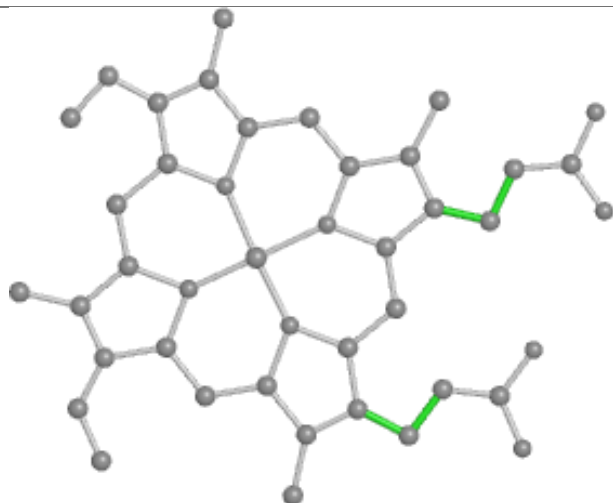
## Ligand HEC E 587



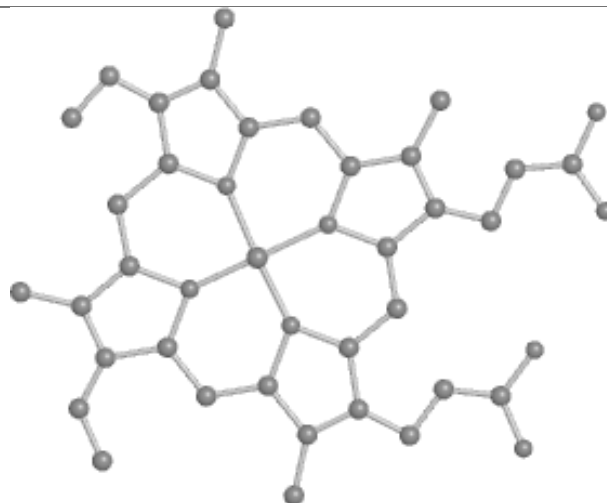
Bond lengths



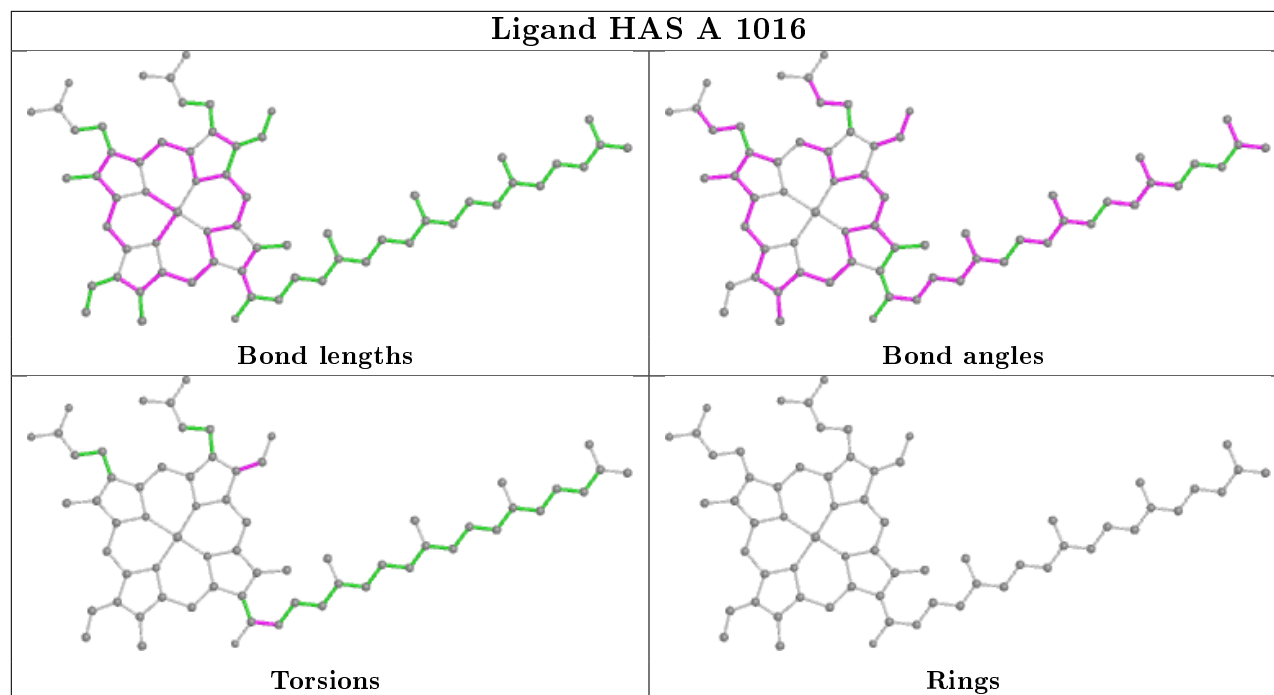
Bond angles



Torsions



Rings



## 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	780/791 (98%)	0.18	14 (1%) 68 77	32, 48, 91, 174	0
1	D	780/791 (98%)	0.19	17 (2%) 62 72	33, 53, 89, 153	0
2	B	319/337 (94%)	0.23	3 (0%) 84 90	33, 50, 85, 163	0
2	E	319/337 (94%)	0.34	15 (4%) 31 44	38, 67, 105, 189	0
3	C	63/66 (95%)	0.13	1 (1%) 72 80	44, 62, 112, 152	0
3	F	62/66 (93%)	0.13	3 (4%) 30 43	45, 65, 104, 153	0
All	All	2323/2388 (97%)	0.21	53 (2%) 60 70	32, 53, 94, 189	0

The worst 5 of 53 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	539	VAL	5.6
2	E	68	PRO	5.4
2	E	69	GLU	5.1
2	E	72	GLU	4.5
1	A	548	HIS	4.2

### 6.2 Non-standard residues in protein, DNA, RNA chains [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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
3	FME	F	1	10/11	0.92	0.14	72,91,112,160	0
3	FME	C	1	10/11	0.93	0.13	69,92,107,171	0

## 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 6.4 Ligands ⓘ

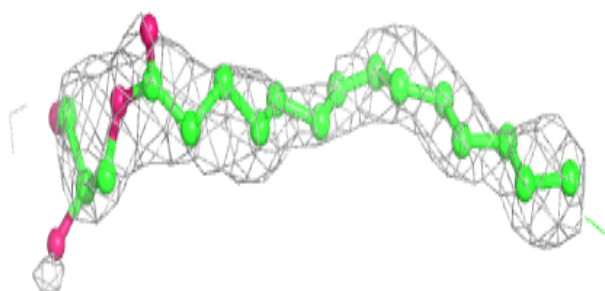
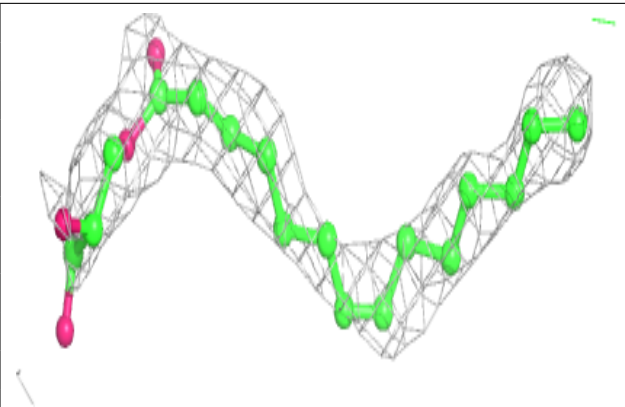
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
8	7E8	A	1300	21/21	0.76	0.31	48,79,102,115	0
12	7E9	B	701	21/21	0.84	0.26	58,90,108,109	0
8	7E8	A	1301	21/21	0.85	0.25	44,80,96,118	0
7	4AG	A	1200	40/40	0.89	0.27	33,68,115,123	0
4	5PL	D	900	85/85	0.90	0.19	37,62,107,128	0
4	5PL	A	900	85/85	0.92	0.18	35,65,105,134	0
13	CL	F	101	1/1	0.94	0.26	70,70,70,70	0
5	HAS	D	1015	65/65	0.95	0.17	30,55,128,138	0
13	CL	D	1701	1/1	0.96	0.17	86,86,86,86	0
5	HAS	A	1015	65/65	0.96	0.18	24,35,102,151	0
11	HEC	E	587	43/43	0.96	0.17	21,54,80,91	0
5	HAS	D	1016	65/65	0.97	0.16	24,47,86,103	0
5	HAS	A	1016	65/65	0.97	0.15	22,38,59,62	0
13	CL	C	101	1/1	0.98	0.28	60,60,60,60	0
9	MG	A	1801	1/1	0.98	0.19	39,39,39,39	0
9	MG	D	1801	1/1	0.98	0.12	50,50,50,50	0
11	HEC	B	587	43/43	0.98	0.15	22,34,48,52	0
10	CUA	E	585	2/2	0.99	0.16	43,43,43,44	0
10	CUA	B	585	2/2	0.99	0.17	35,35,35,36	0
6	CU	D	1017	1/1	1.00	0.16	42,42,42,42	0
6	CU	A	1017	1/1	1.00	0.20	37,37,37,37	0

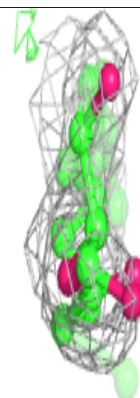
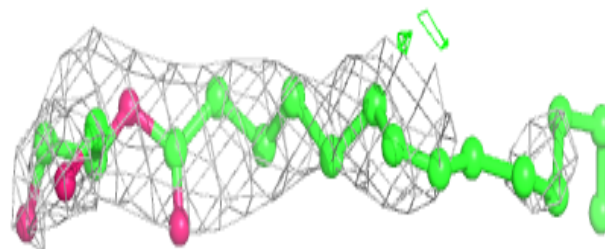
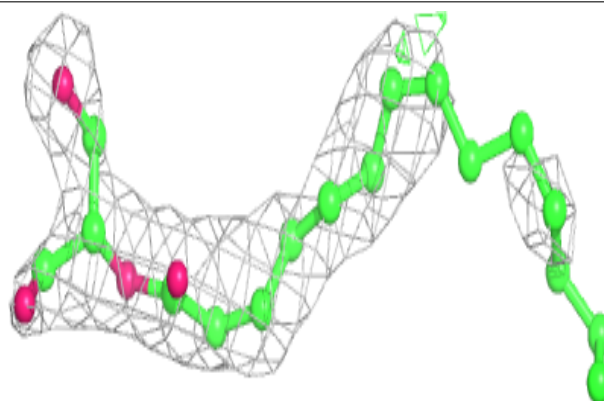
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

**Electron density around 7E8 A 1300:**

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

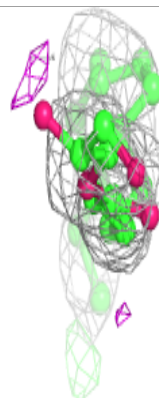
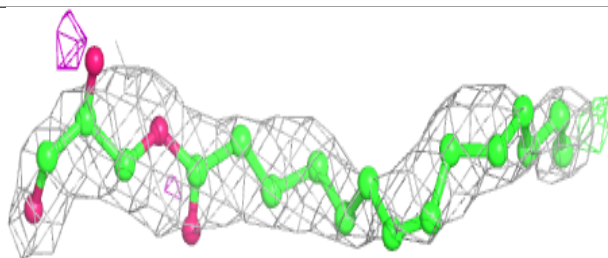
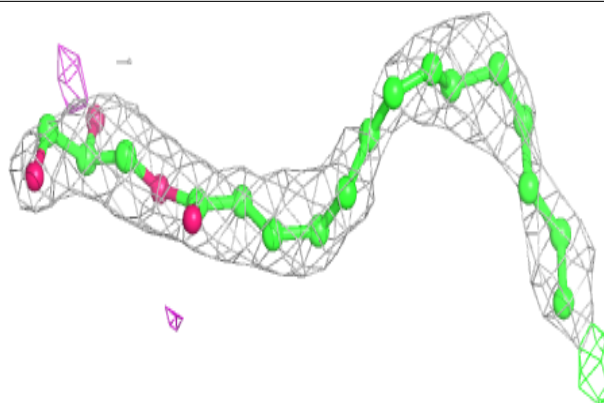
**Electron density around 7E9 B 701:**

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



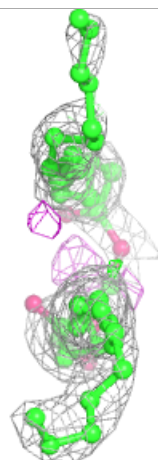
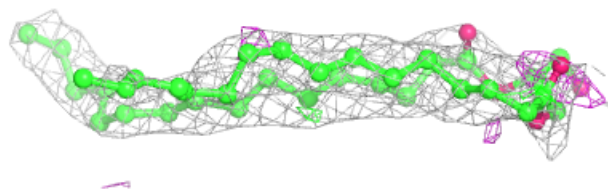
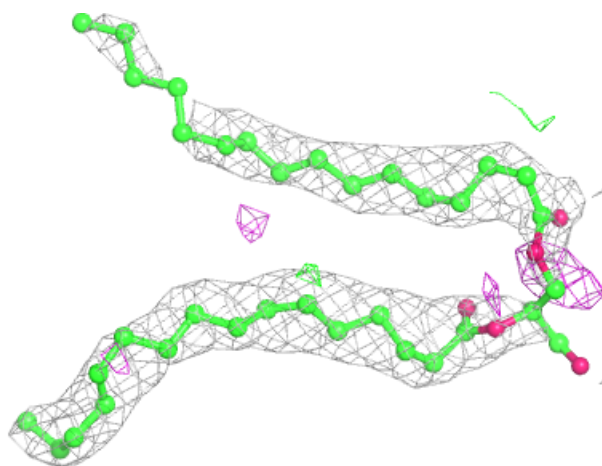
**Electron density around 7E8 A 1301:**

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



**Electron density around 4AG A 1200:**

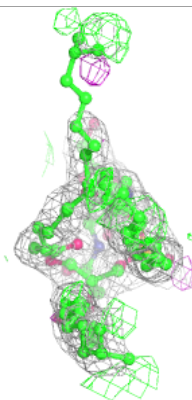
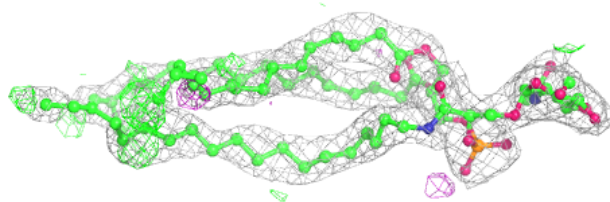
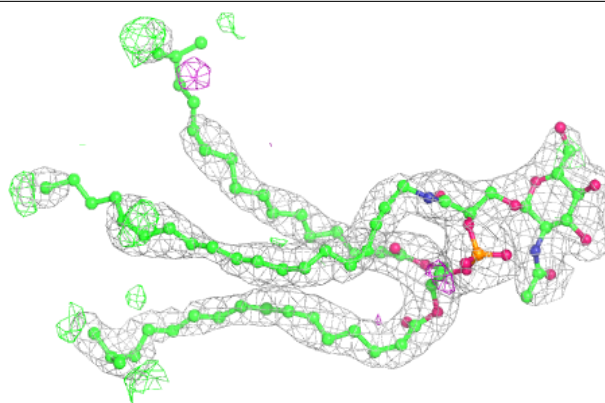
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



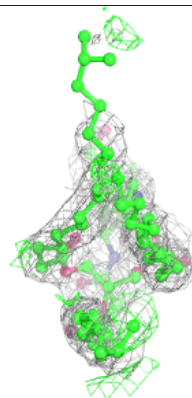
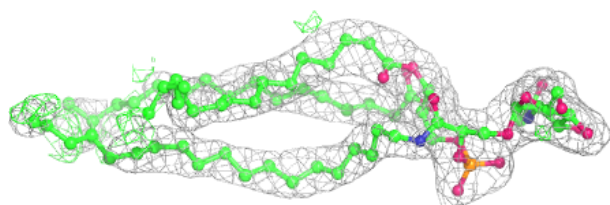
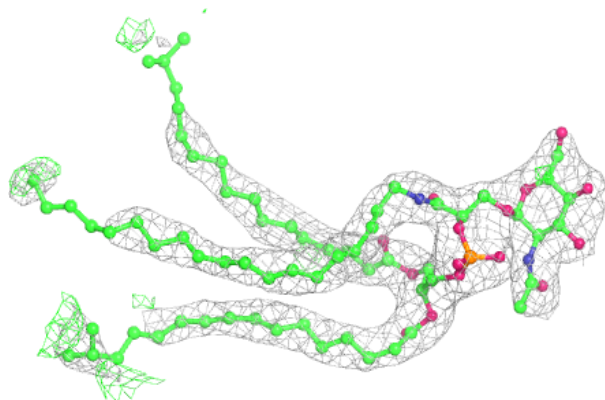


**Electron density around 5PL D 900:**

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

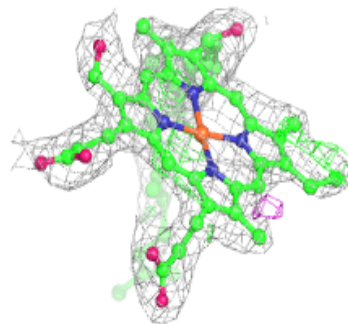
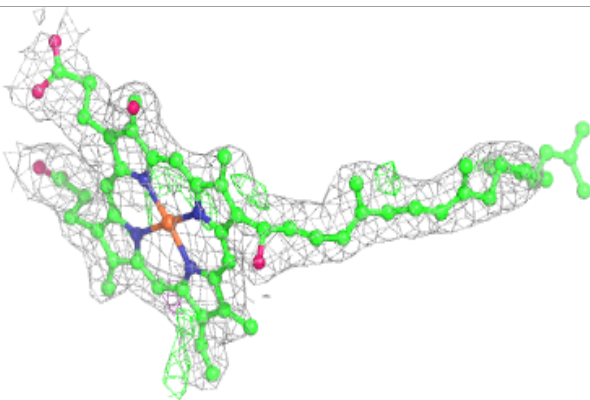
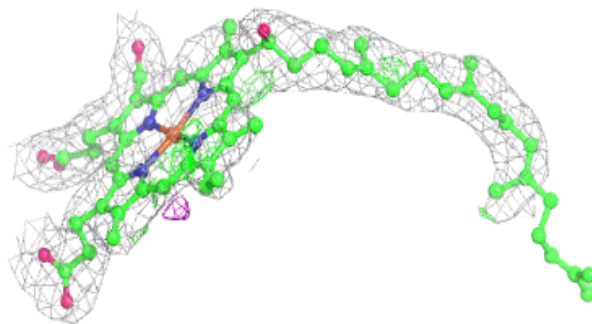
**Electron density around 5PL A 900:**

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

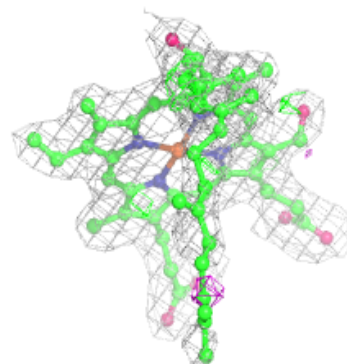
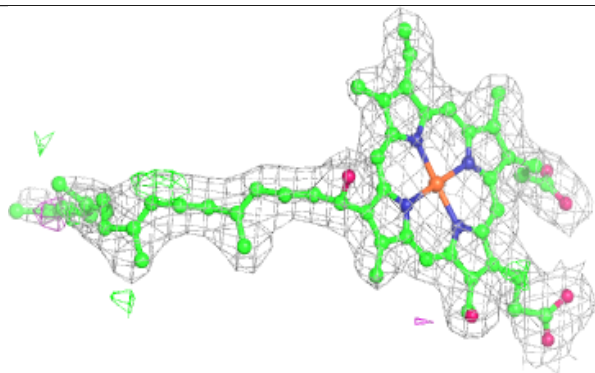
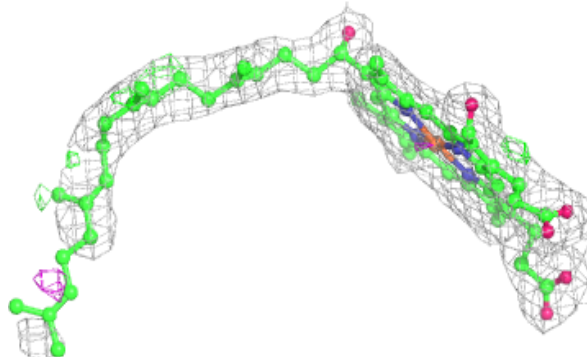


**Electron density around HAS D 1015:**

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

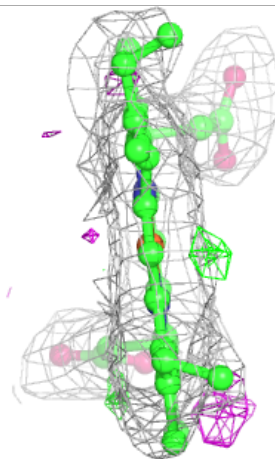
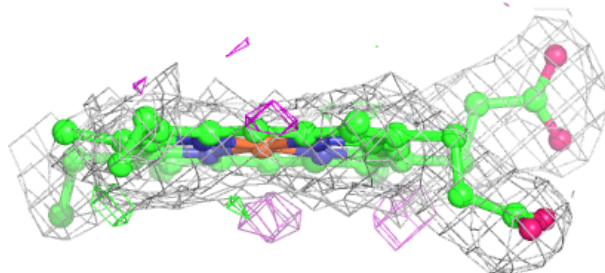
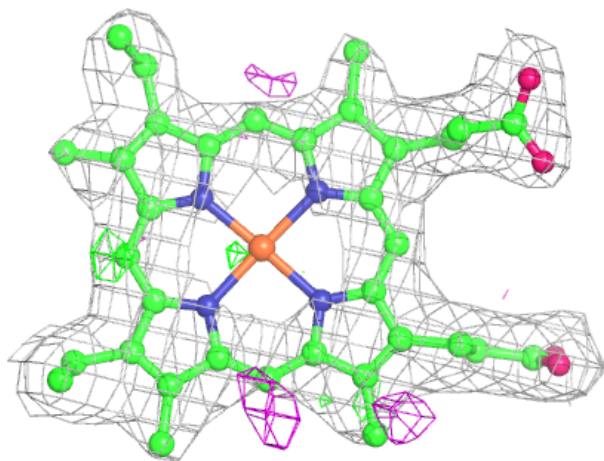
**Electron density around HAS A 1015:**

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



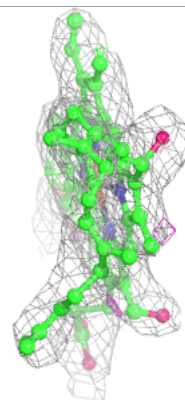
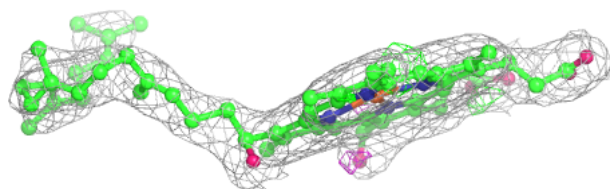
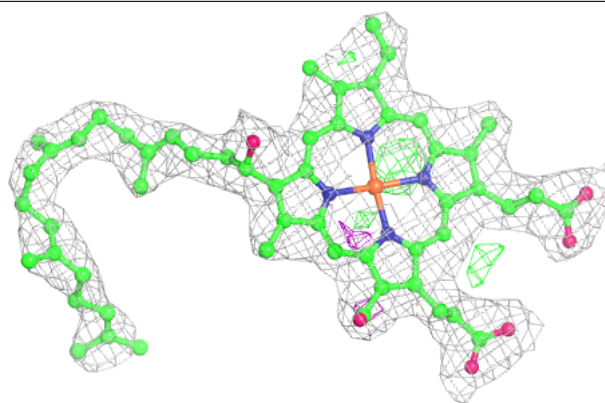
**Electron density around HEC E 587:**

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

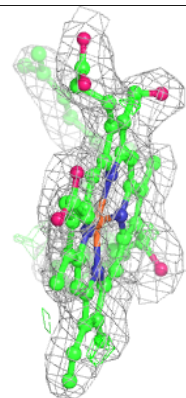
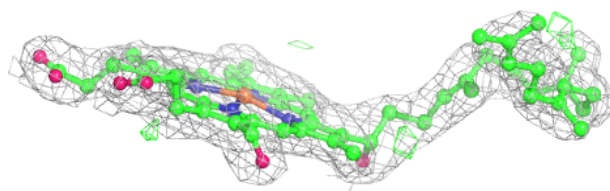
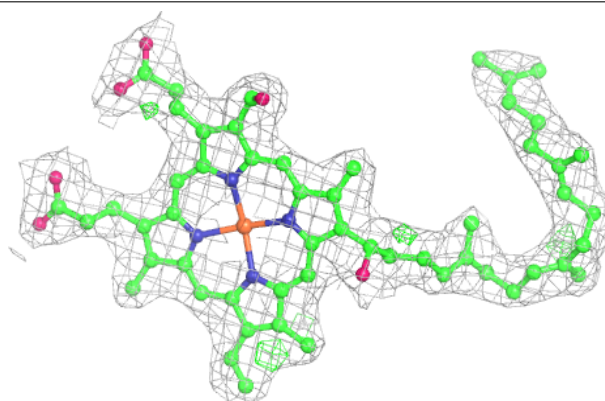


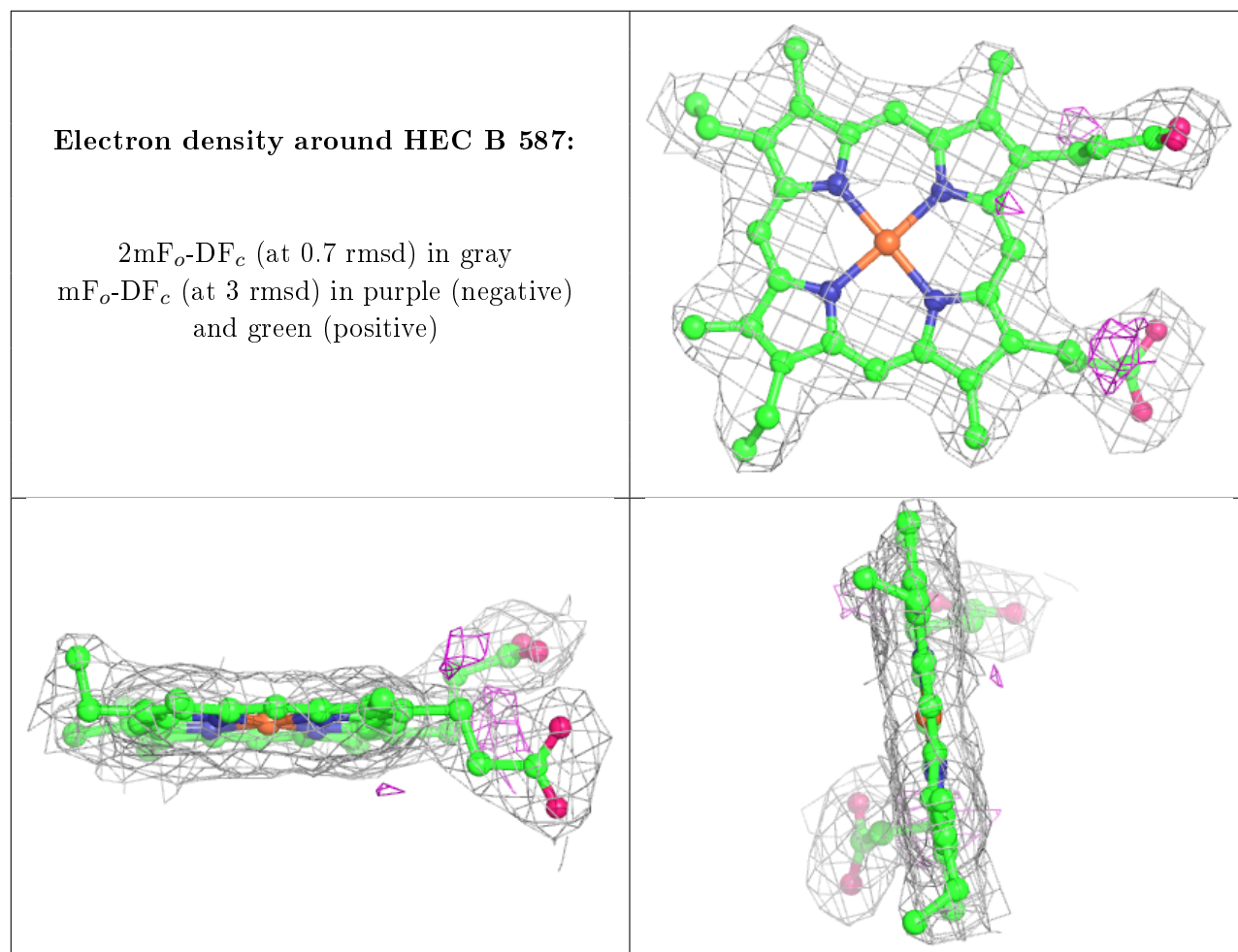
**Electron density around HAS D 1016:**

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

**Electron density around HAS A 1016:**

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





## 6.5 Other polymers [i](#)

There are no such residues in this entry.