

Full wwPDB X-ray Structure Validation Report (i)

Mar 19, 2024 - 12:55 pm GMT

:	80E7
:	Crystal structure of the Z-DNA duplex d(CGCGCG) containing ordered cop-
	per(II) and soaked in hydrogen peroxide for 5 minutes
:	Lambert, M.C.; Hall, J.P.
	2023-03-10
:	1.25 Å(reported)
	: : :

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

We welcome your comments at validation@mail.wwpdb.org A user guide is available at https://www.wwpdb.org/validation/2017/XrayValidationReportHelp with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (1)) were used in the production of this report:

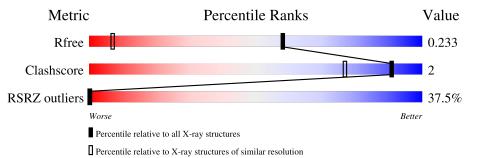
MolProbity Mogul	:	4.02b-467 1.8.4, CSD as541be (2020)
Xtriage (Phenix)		1.13
EDS	:	2.36
buster-report	:	1.1.7(2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure: $X\text{-}RAY\;DIFFRACTION$

The reported resolution of this entry is 1.25 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	$egin{array}{c} { m Whole \ archive} \ (\#{ m Entries}) \end{array}$	${f Similar\ resolution}\ (\#{ m Entries,\ resolution\ range}({ m \AA}))$
R_{free}	130704	1023 (1.28-1.24)
Clashscore	141614	1060 (1.28-1.24)
RSRZ outliers	127900	1004 (1.28-1.24)

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					
			50%					
1	А	6	83%	17%				
			33%					
1	В	6	100%					
			17%					
1	С	6	83%	17%				
			67%					
1	D	6	100%					
			17%					
1	Ε	6	83%	17%				
			67%					
1	F	6	100%					
			17%					
1	G	6	100%					



Continued from previous page...

Mol	Chain	Length		Quality of chain
			33%	
1	Н	6		100%

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	CU	С	103	-	-	-	Х



2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 1911 atoms, of which 621 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.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	٨	6	Total	С	Η	Ν	0	Р	0	1	0
	А	0	218	66	79	27	40	6	0	1	0
1	В	6	Total	С	Η	Ν	Ο	Р	0	1	0
	D	0	218	66	79	27	40	6	0	1	0
1	С	6	Total	С	Η	Ν	Ο	Р	0	1	0
	U	0	218	66	79	27	40	6	0	I	0
1	D	6	Total	С	Η	Ν	Ο	Р	0	1	0
	D	0	218	66	79	27	40	6	0	1	0
1	Е	6	Total	С	Η	Ν	Ο	Р	0	1	0
	Ľ	0	218	66	79	27	40	6	0		U
1	F	6	Total	С	Η	Ν	Ο	Р	0	1	0
	Ľ	0	218	66	79	27	40	6	0	T	0
1	G	6	Total	С	Η	Ν	Ο	Р	0	0	0
	G	0	188	57	68	24	34	5	0	0	0
1	Н	6	Total	С	Η	Ν	Ο	Р	0	1	0
	11	0	218	66	79	27	40	6	0		0

• Molecule 1 is a DNA chain called DNA (5'-D(*CP*GP*CP*GP*CP*G)-3').

• Molecule 2 is COPPER (II) ION (three-letter code: CU) (formula: Cu) (labeled as "Ligand of Interest" by depositor).

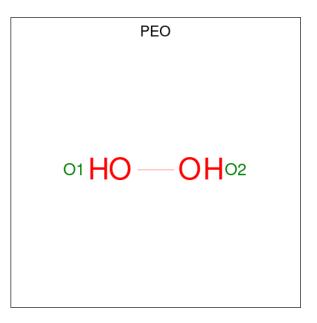
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	А	3	Total Cu 3 3	0	0
2	В	2	Total Cu 2 2	0	0
2	С	3	Total Cu 3 3	0	0
2	D	2	Total Cu 2 2	0	0
2	Е	3	Total Cu 3 3	0	0
2	F	2	Total Cu 2 2	0	0



Continued	from	previous	page
-----------	------	----------	------

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	G	3	Total Cu 3 3	0	0
2	Н	2	Total Cu 2 2	0	0

• Molecule 3 is HYDROGEN PEROXIDE (three-letter code: PEO) (formula: H_2O_2) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	А	1	Total O 2 2	0	0
3	А	1	Total O 2 2	0	0
3	С	1	Total O 2 2	0	0
3	С	1	Total O 2 2	0	0
3	С	1	Total O 2 2	0	0
3	G	1	Total O 2 2	0	0
3	G	1	Total O 2 2	0	0

• Molecule 4 is water.



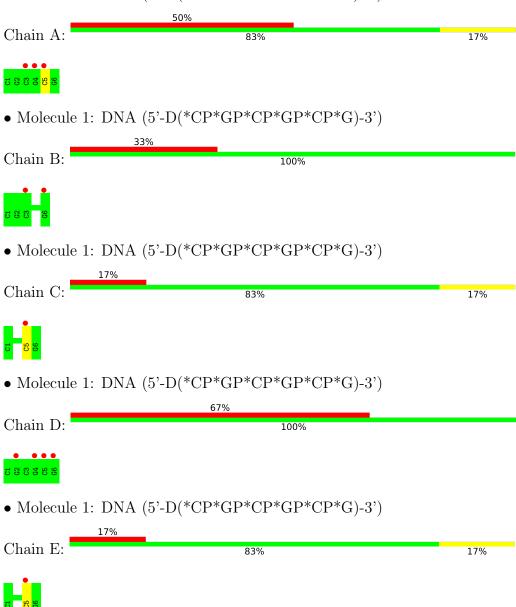
80E1	80E7
------	------

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	А	29	Total O 29 29	0	0
4	В	14	Total O 16 16	0	2
4	С	18	Total O 18 18	0	0
4	D	15	Total O 15 15	0	0
4	Е	29	Total O 29 29	0	0
4	F	18	Total O 18 18	0	0
4	G	23	TotalO2323	0	0
4	Н	14	Total O 15 15	0	1



3 Residue-property plots (i)

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: DNA (5'-D(*CP*GP*CP*GP*CP*G)-3')

• Molecule 1: DNA (5'-D(*CP*GP*CP*GP*CP*G)-3')



	67%	
Chain F:	100%	
C C C C C C C C C C C C C C C C C C C		
	NA $(5'-D(*CP*GP*CP*GP*CP*G)-3')$	
Chain G:	100%	
88		
• Molecule 1: D	NA $(5'-D(*CP*GP*CP*GP*CP*G)-3')$	
Chain H:	33%	
	100%	
8 8 3 3 5		



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants	18.28Å 30.74Å 42.56Å	Depositor
a, b, c, α , β , γ	90.01° 89.95° 90.01°	Depositor
Resolution (Å)	24.92 - 1.25	Depositor
Resolution (A)	30.74 - 1.25	EDS
% Data completeness	93.9 (24.92-1.25)	Depositor
(in resolution range)	94.9 (30.74-1.25)	EDS
R _{merge}	(Not available)	Depositor
R _{sym}	(Not available)	Depositor
$< I/\sigma(I) > 1$	$3.78 (at 1.25 \text{\AA})$	Xtriage
Refinement program	PHENIX 1.20.1_4487	Depositor
D D	0.199 , 0.233	Depositor
R, R_{free}	0.197 , 0.233	DCC
R_{free} test set	1292 reflections (5.30%)	wwPDB-VP
Wilson B-factor $(Å^2)$	11.2	Xtriage
Anisotropy	0.549	Xtriage
Bulk solvent $k_{sol}(e/Å^3), B_{sol}(Å^2)$	0.33 , 43.0	EDS
L-test for twinning ²	$< L > = 0.44, < L^2 > = 0.27$	Xtriage
	0.418 for h,-k,-l	
Estimated twinning fraction	0.448 for -h,k,-l	Xtriage
	0.426 for -h,-k,l	
F_o, F_c correlation	0.97	EDS
Total number of atoms	1911	wwPDB-VP
Average B, all atoms $(Å^2)$	13.0	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 35.46 % 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 5.8169e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.



¹Intensities estimated from amplitudes.

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: PEO, 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	
	Unam	RMSZ	# Z > 5	RMSZ	# Z > 5
1	А	0.69	0/155	0.88	0/237
1	В	0.80	0/155	0.99	0/237
1	С	0.70	0/155	0.89	0/237
1	D	0.76	0/155	0.98	0/237
1	Е	0.70	0/155	0.88	0/237
1	F	0.75	0/155	0.94	0/237
1	G	0.72	0/134	0.90	0/205
1	Н	0.86	0/155	1.00	0/237
All	All	0.75	0/1219	0.93	0/1864

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	А	139	79	79	1	0
1	В	139	79	79	0	0
1	С	139	79	79	1	0
1	D	139	79	79	0	0
1	Е	139	79	79	1	0



Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	F	139	79	79	0	0
1	G	120	68	68	0	0
1	Н	139	79	79	0	0
2	А	3	0	0	0	0
2	В	2	0	0	0	0
2	С	3	0	0	0	0
2	D	2	0	0	0	0
2	Ε	3	0	0	0	0
2	F	2	0	0	0	0
2	G	3	0	0	0	0
2	Η	2	0	0	0	0
3	А	4	0	0	0	0
3	С	6	0	0	0	0
3	G	4	0	0	0	0
4	А	29	0	0	1	0
4	В	16	0	0	0	0
4	С	18	0	0	1	0
4	D	15	0	0	0	0
4	Е	29	0	0	1	0
4	F	18	0	0	0	0
4	G	23	0	0	0	0
4	Н	15	0	0	0	0
All	All	1290	621	621	3	0

Continued from previous page...

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 2.

All (3) 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:C:5[B]:DC:OP1	4:C:201:HOH:O	1.90	0.90
1:E:5[A]:DC:OP2	4:E:201:HOH:O	2.07	0.72
1:A:5[A]:DC:OP2	4:A:202:HOH:O	2.22	0.41

There are no symmetry-related clashes.



5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

There are no protein molecules in this entry.

5.3.2 Protein sidechains (i)

There are no protein molecules in this entry.

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 27 ligands modelled in this entry, 20 are monoatomic - leaving 7 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	Turne	Chain	Dec	Link	B	ond leng	$_{ m gths}$	B	ond ang	gles
IVIOI	Mol Type	Chain	nes	LIIIK	Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z >2
3	PEO	С	104	2	1,1,1	0.05	0	-		
3	PEO	G	104	2	1,1,1	0.03	0	-		
3	PEO	С	106	2	1,1,1	0.04	0	-		
3	PEO	С	105	2	1,1,1	0.00	0	-		
3	PEO	А	104	2	1,1,1	0.05	0	-		
3	PEO	А	105	2	1,1,1	0.03	0	-		
3	PEO	G	105	2	1,1,1	0.05	0	-		



There are no bond length outliers.

There are no bond angle outliers.

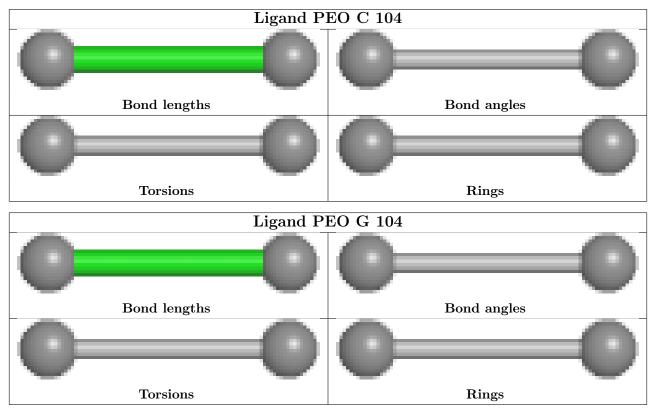
There are no chirality outliers.

There are no torsion outliers.

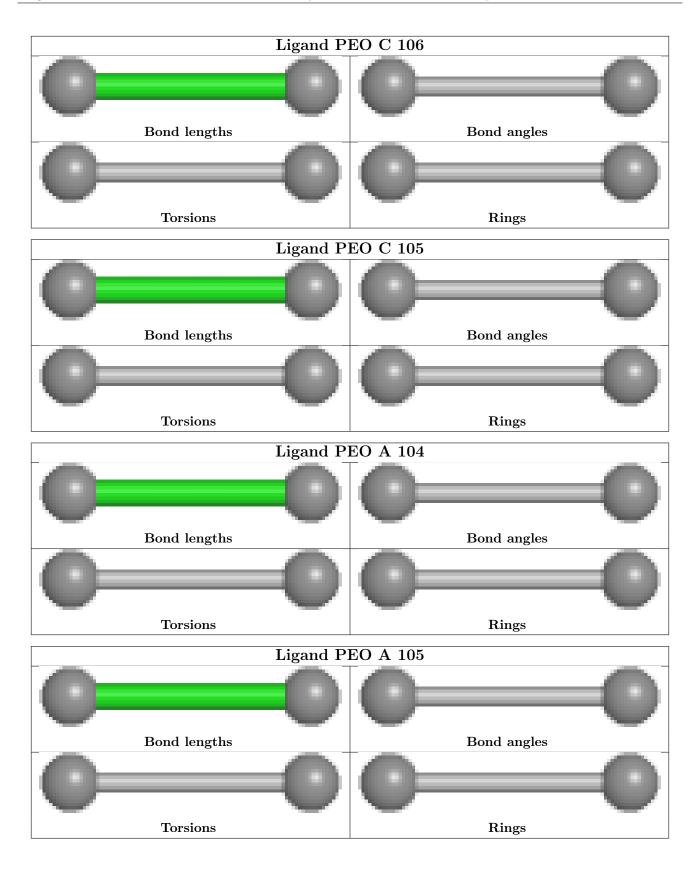
There are no ring outliers.

No monomer is involved in short contacts.

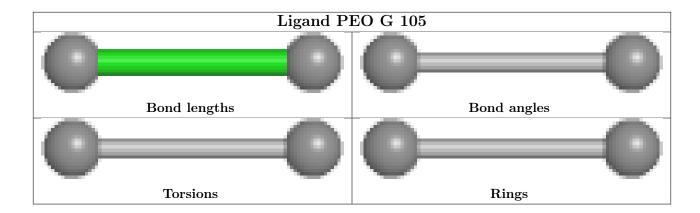
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. 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 sufficient must be highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.











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^{th} 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	$\langle RSRZ \rangle$	#RSR2	$\mathbf{Z} \!>\! \mathbf{\hat{z}}$	2	$OWAB(A^2)$	Q < 0.9
1	А	6/6~(100%)	1.89	3~(50%)	0	0	11, 11, 12, 12	0
1	В	6/6~(100%)	1.85	2 (33%)	0	0	11, 11, 12, 14	0
1	С	6/6~(100%)	1.65	1 (16%)	1	1	11, 12, 12, 12	0
1	D	6/6~(100%)	2.29	4 (66%)	0	0	11, 11, 12, 13	0
1	Ε	6/6~(100%)	1.77	1 (16%)	1	1	11, 12, 12, 12	0
1	F	6/6~(100%)	2.12	4 (66%)	0	0	11, 12, 12, 12	0
1	G	6/6~(100%)	1.45	1 (16%)	1	1	10, 11, 12, 13	0
1	Н	6/6~(100%)	1.68	2 (33%)	0	0	11, 11, 12, 14	0
All	All	48/48~(100%)	1.84	18~(37%)	0	0	10, 12, 13, 14	0

All (18) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	4	DG	3.1
1	F	4	DG	3.1
1	В	3[A]	DC	2.9
1	Н	6	DG	2.9
1	D	5[A]	DC	2.9
1	F	3[A]	DC	2.6
1	С	5[A]	DC	2.5
1	D	2	DG	2.4
1	Ε	5[A]	DC	2.4
1	А	5[A]	DC	2.3
1	А	4	DG	2.3
1	F	5	DC	2.2
1	D	6	DG	2.2
1	F	6	DG	2.2
1	А	3	DC	2.1
1	Н	3[A]	DC	2.1



Continued from previous page...

Mol	Chain	\mathbf{Res}	Type	RSRZ	
1	G	5	DC	2.1	
1	В	6	DG	2.0	

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, 95^{th} 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	CU	Е	103	1/1	0.35	0.30	75,75,75,75	0
2	CU	С	103	1/1	0.41	0.73	$67,\!67,\!67,\!67$	1
2	CU	Е	102	1/1	0.43	0.24	70,70,70,70	0
2	CU	В	102	1/1	0.57	0.36	57,57,57,57	0
2	CU	G	101	1/1	0.59	0.12	68,68,68,68	0
2	CU	С	101	1/1	0.65	0.12	64,64,64,64	0
2	CU	В	101	1/1	0.67	0.16	64,64,64,64	0
2	CU	Н	101	1/1	0.69	0.33	61,61,61,61	0
3	PEO	С	106	2/2	0.73	0.23	17,17,17,20	2
2	CU	Е	101	1/1	0.74	0.12	46,46,46,46	0
2	CU	С	102	1/1	0.80	0.07	45,45,45,45	0
3	PEO	С	105	2/2	0.84	0.26	18,18,18,18	0
3	PEO	С	104	2/2	0.84	0.10	23,23,23,25	0
2	CU	D	101	1/1	0.86	0.08	59,59,59,59	0
3	PEO	G	104	2/2	0.88	0.09	24,24,24,25	0
2	CU	D	102	1/1	0.90	0.07	51,51,51,51	1
2	CU	А	101	1/1	0.90	0.15	68,68,68,68	0
2	CU	А	102	1/1	0.90	0.15	44,44,44,44	0
2	CU	Н	102	1/1	0.90	0.26	61,61,61,61	1
2	CU	G	103	1/1	0.91	0.09	56, 56, 56, 56	0
2	CU	G	102	1/1	0.92	0.12	48,48,48,48	0

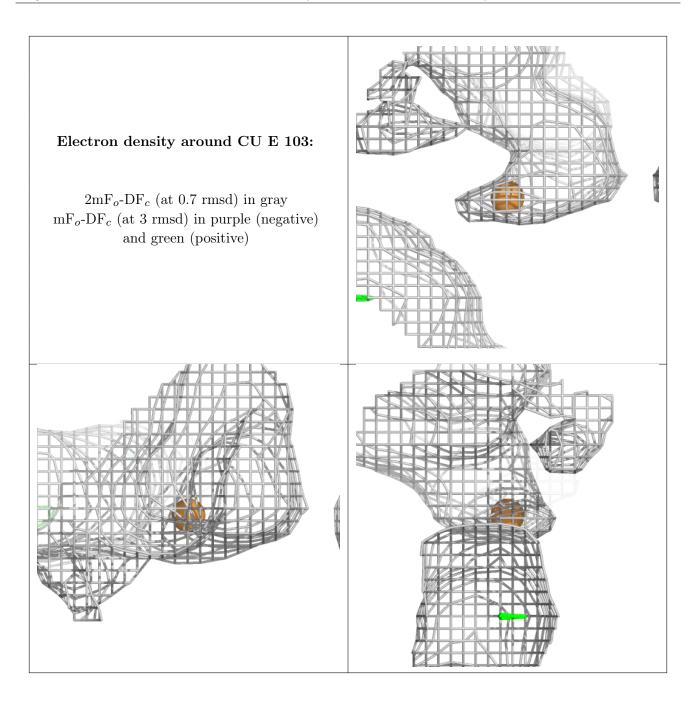


Mol	Type	Chain	Res	Atoms	RSCC	\mathbf{RSR}	$\mathbf{B} ext{-factors}(\mathbf{A}^2)$	Q < 0.9
3	PEO	А	104	2/2	0.92	0.16	$23,\!23,\!23,\!23$	2
3	PEO	А	105	2/2	0.93	0.21	31,31,31,32	0
2	CU	F	102	1/1	0.94	0.04	55,55,55,55	0
2	CU	А	103	1/1	0.95	0.12	62,62,62,62	0
3	PEO	G	105	2/2	0.96	0.12	19,19,19,21	0
2	CU	F	101	1/1	0.97	0.03	57,57,57,57	0

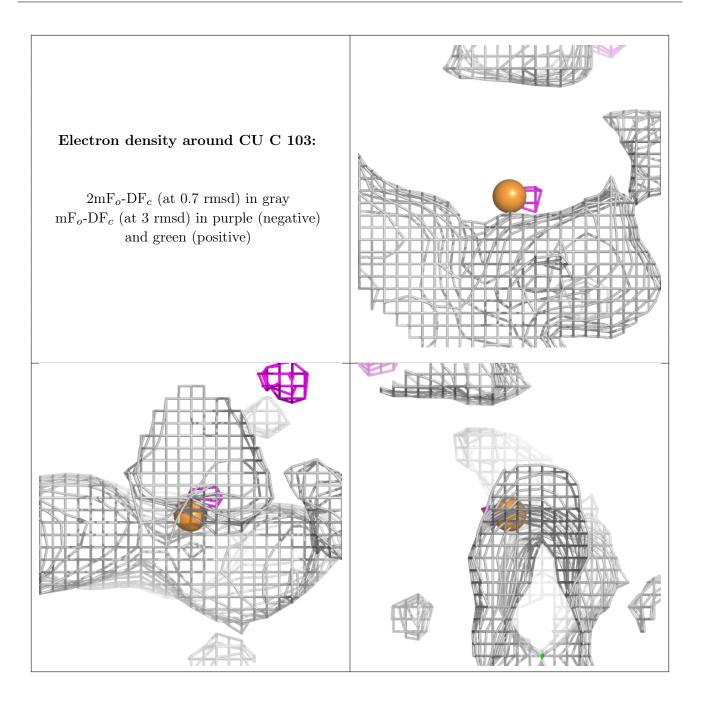
Continued from previous page...

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.

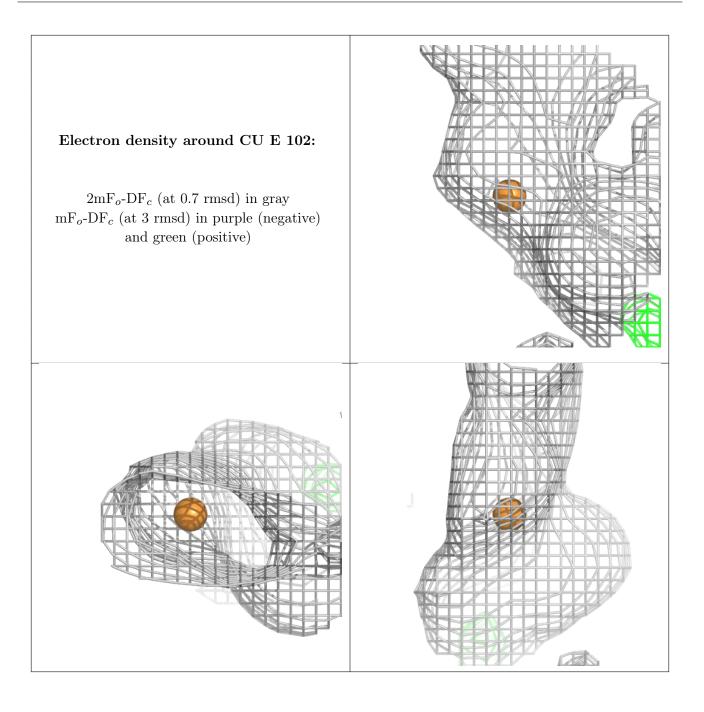




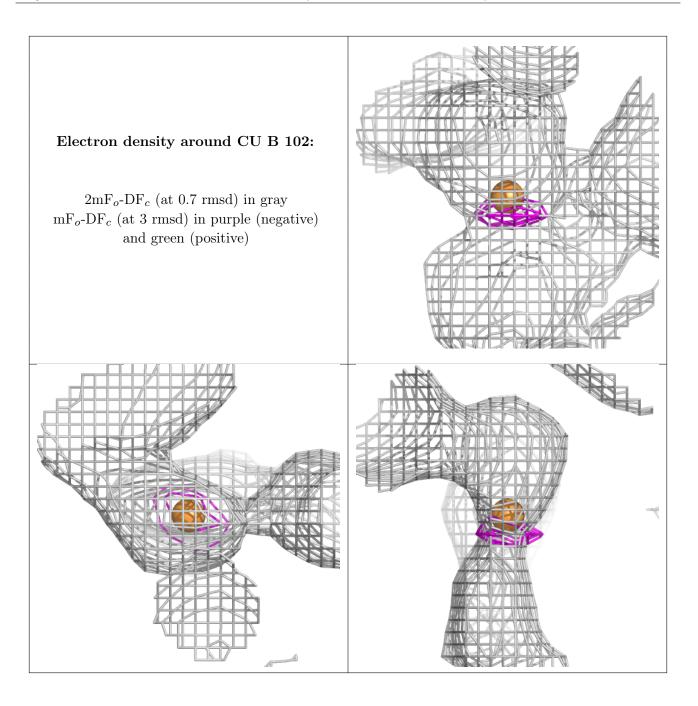




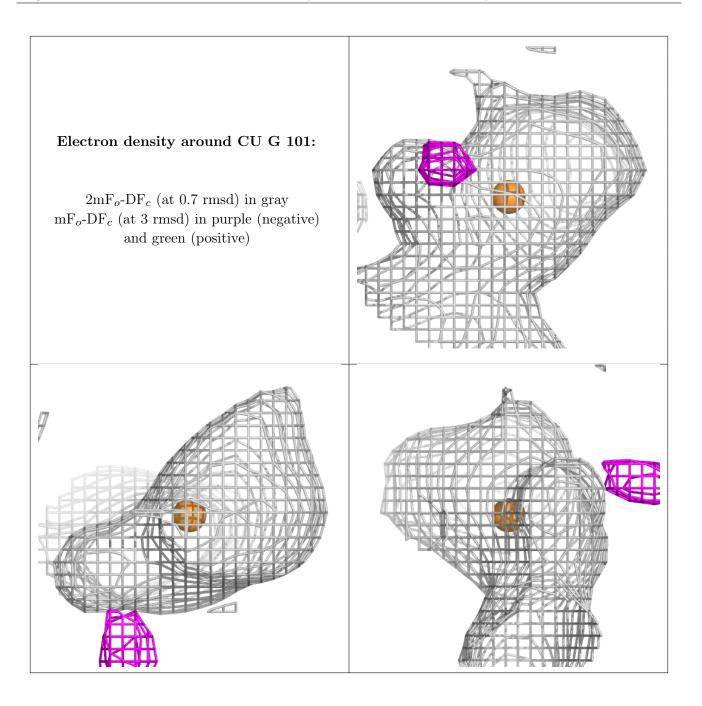




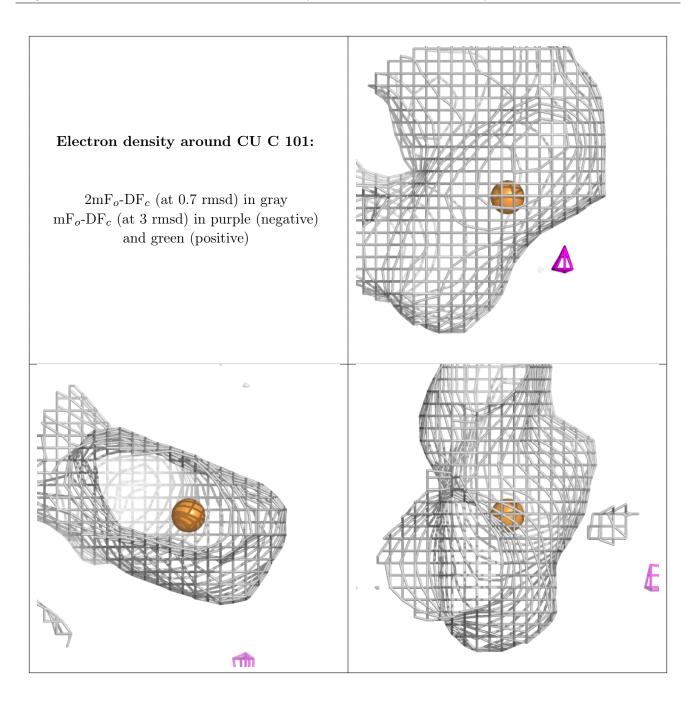




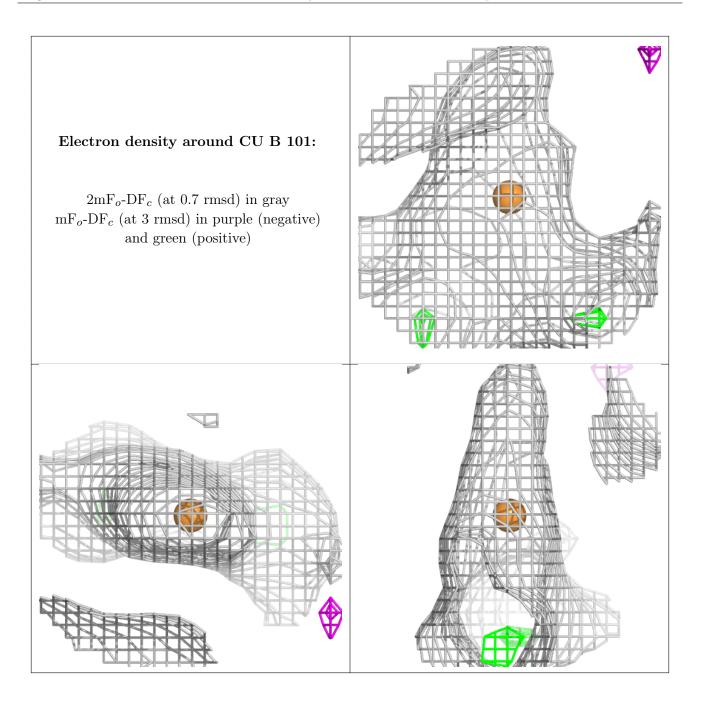




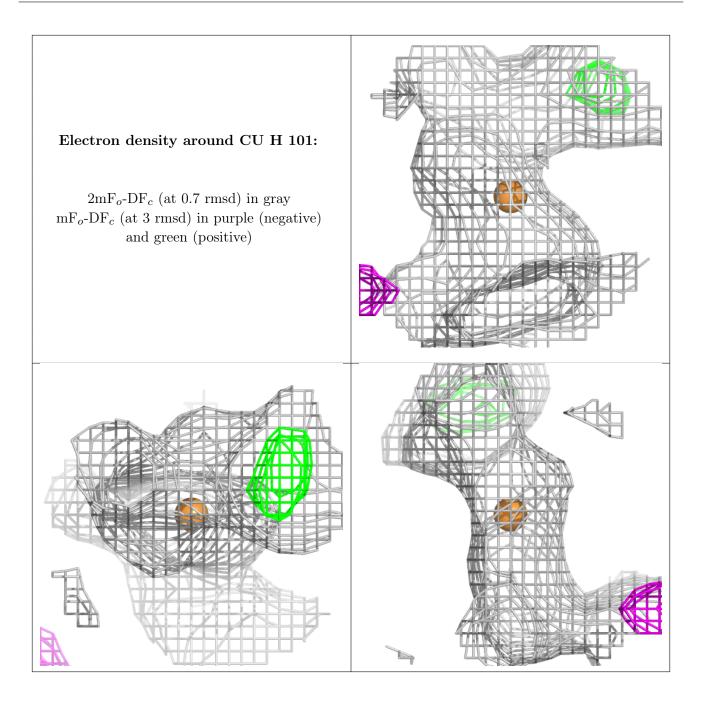




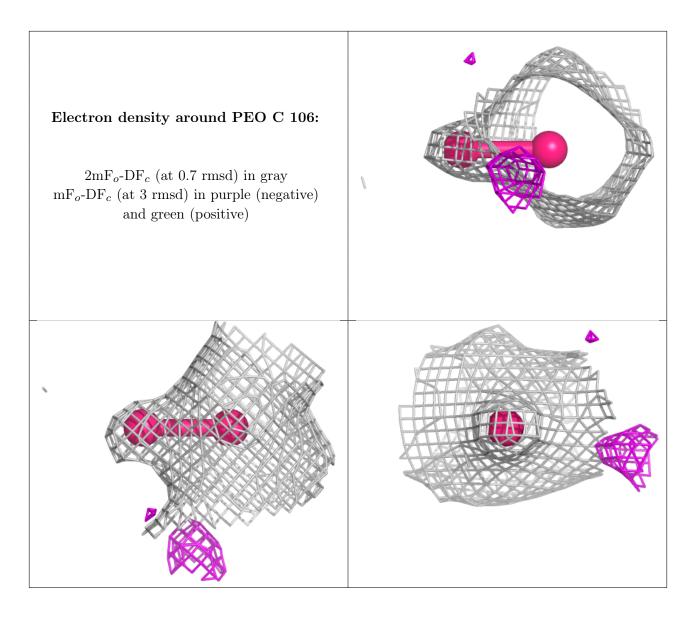




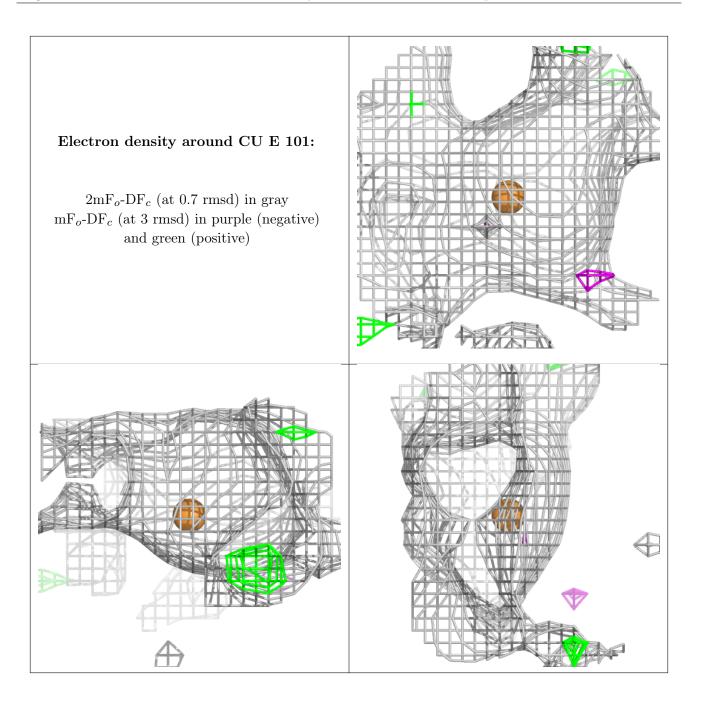




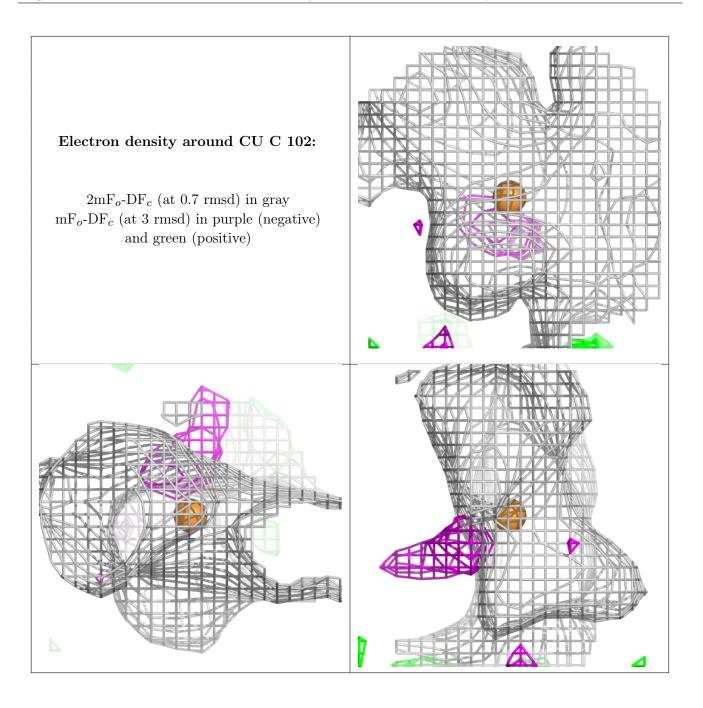




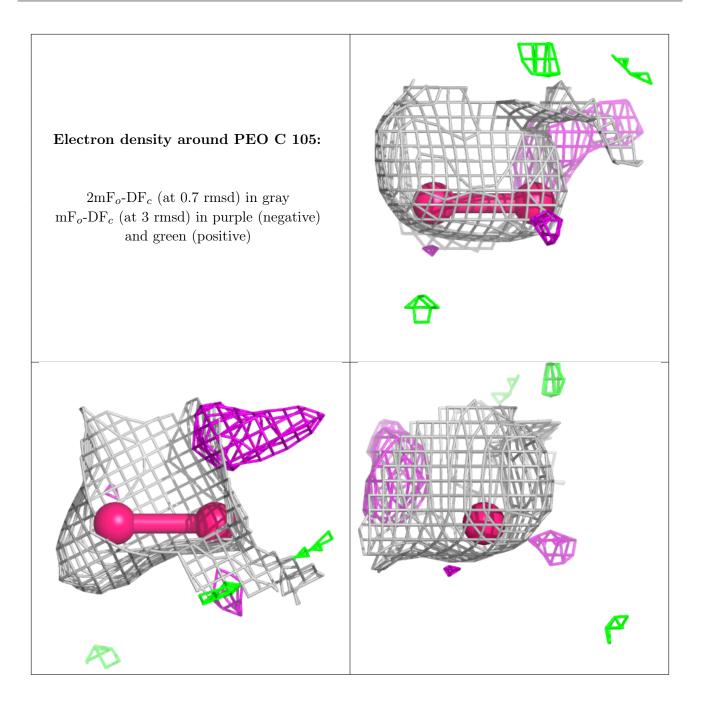




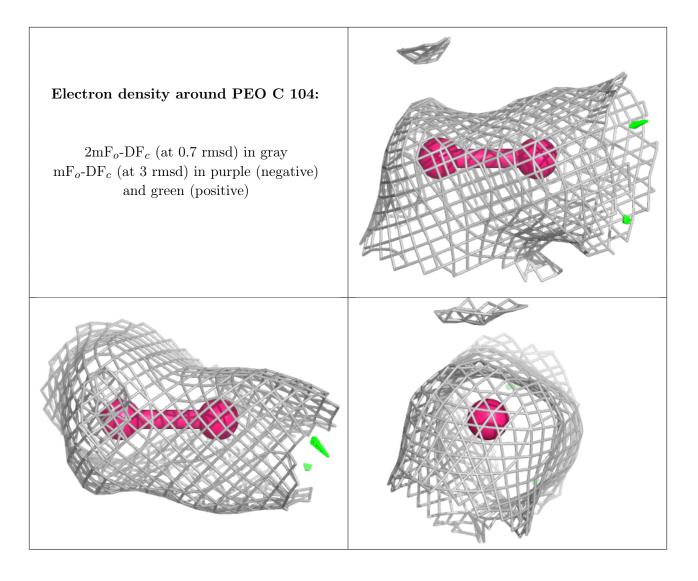




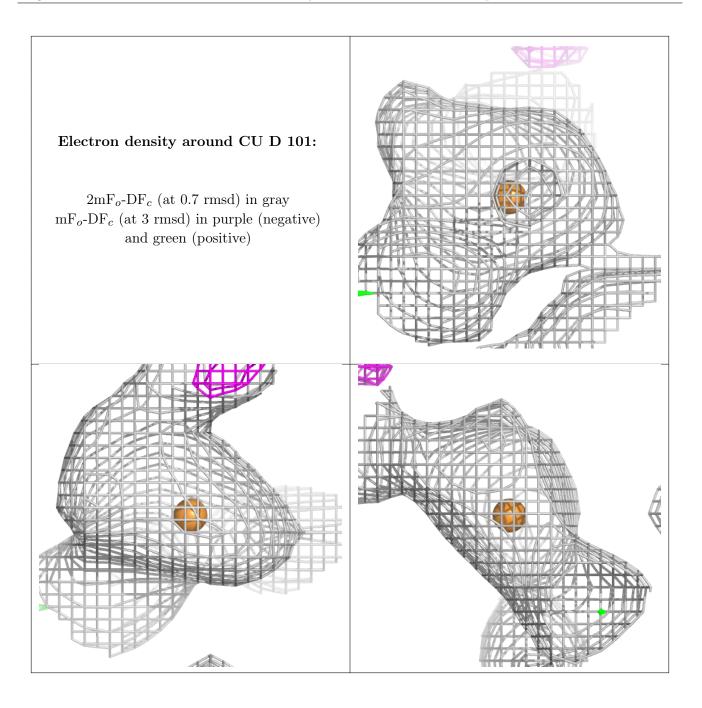




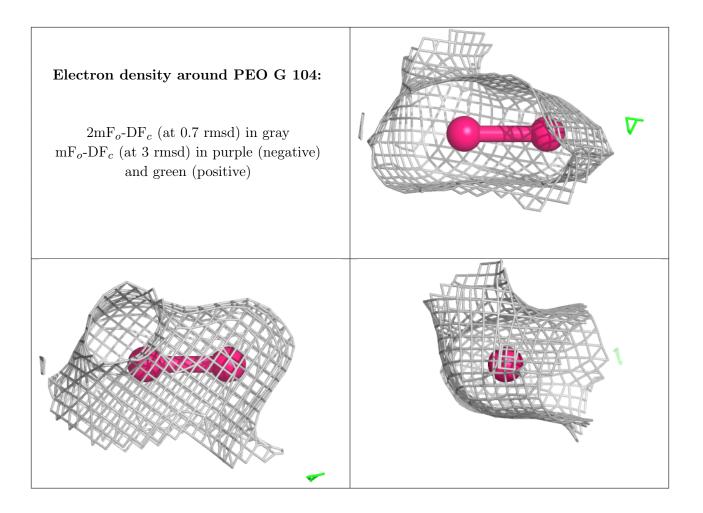




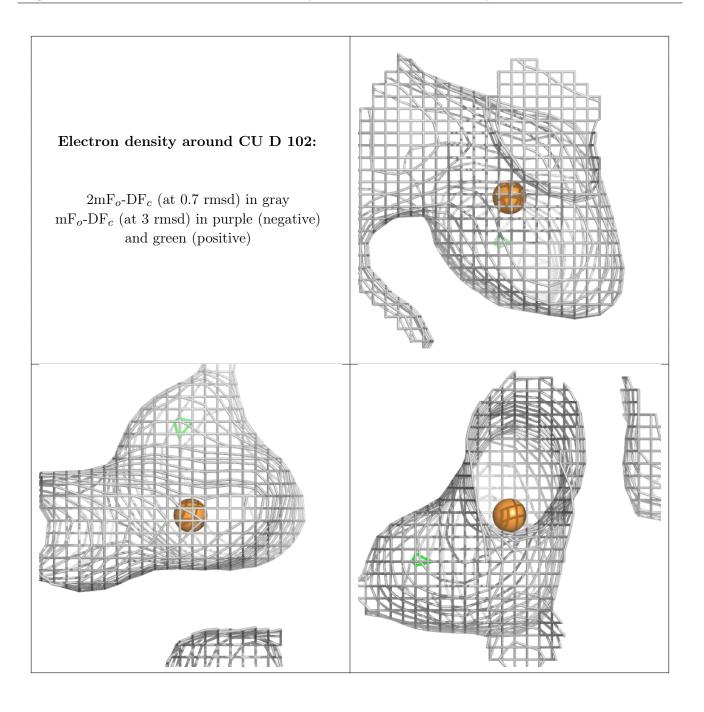




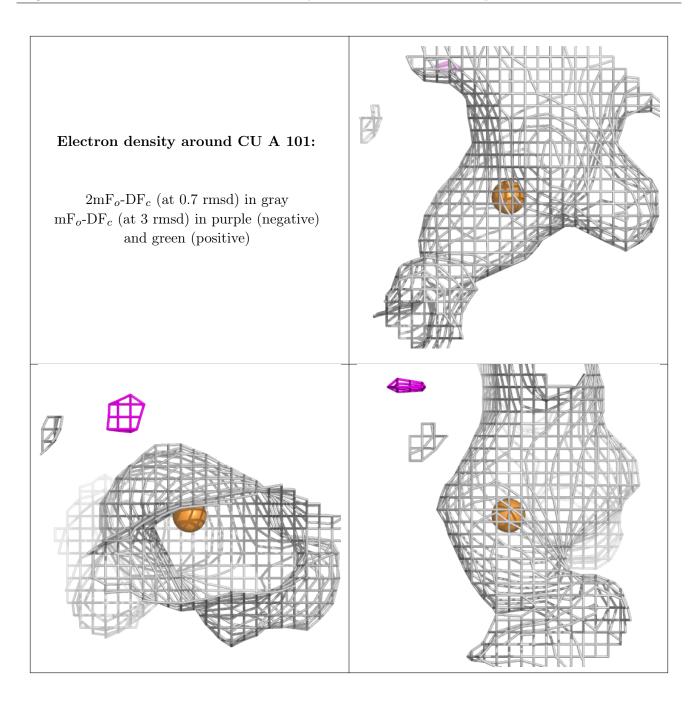




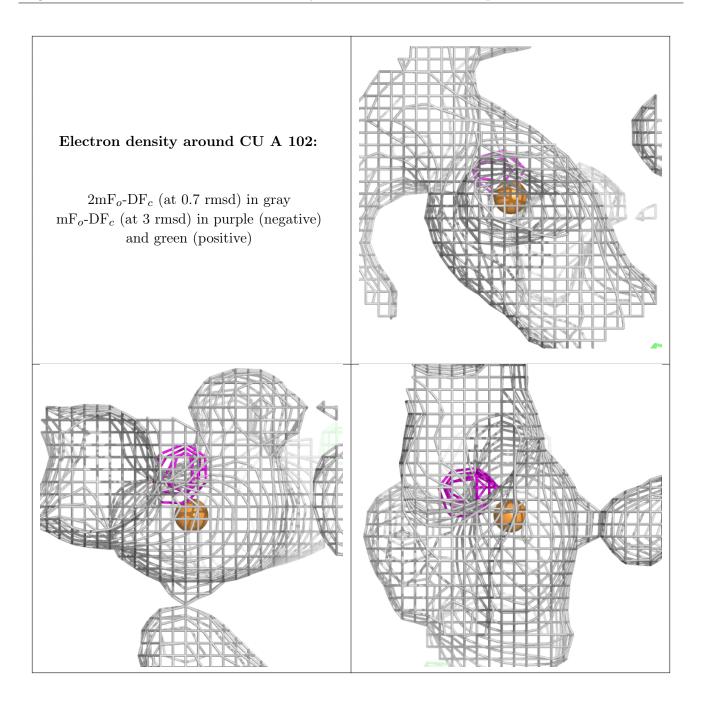




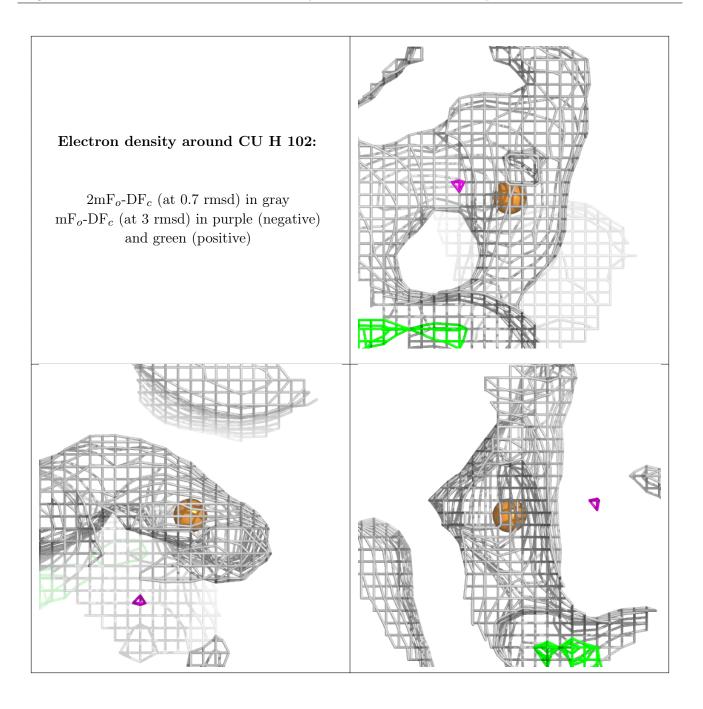




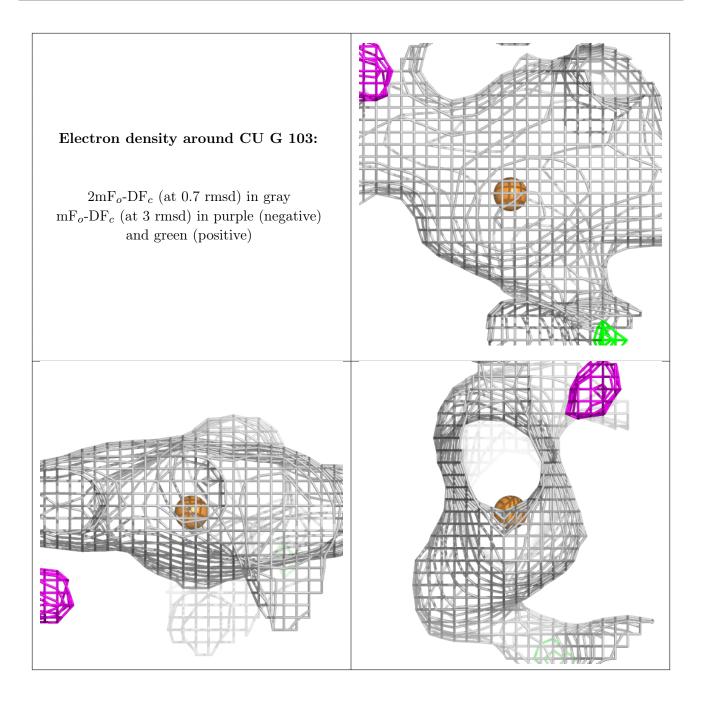




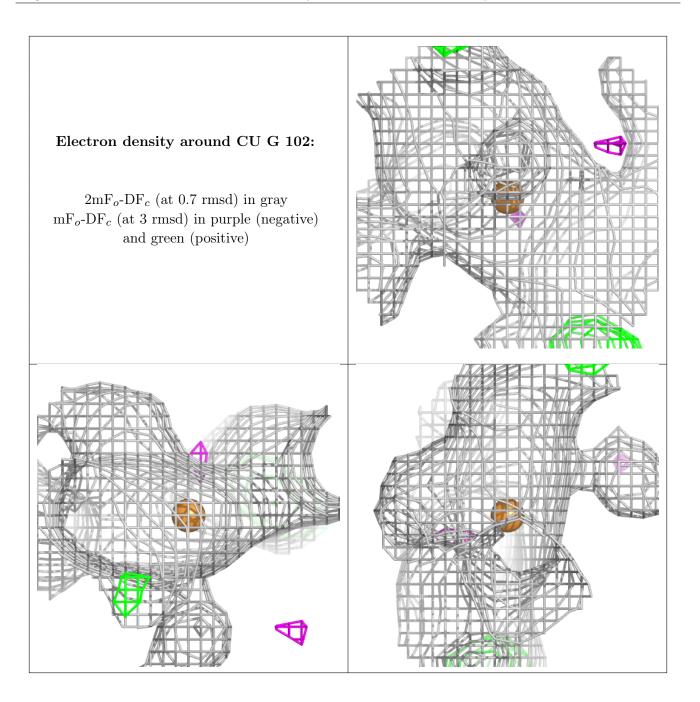




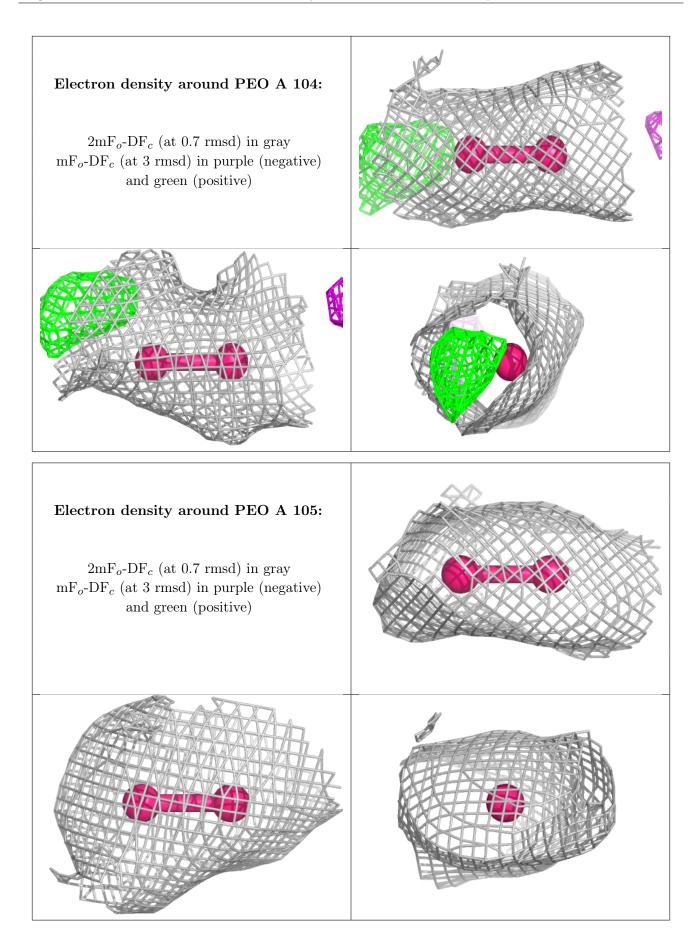




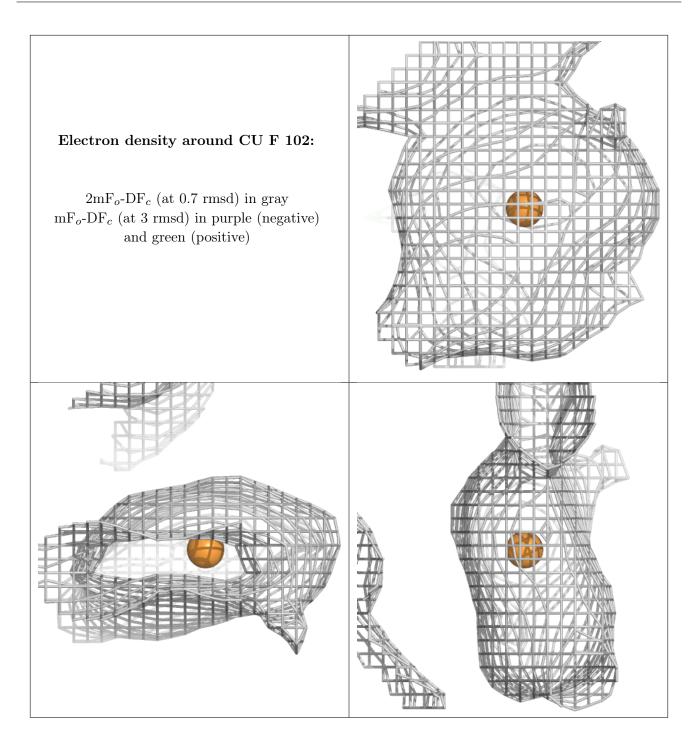




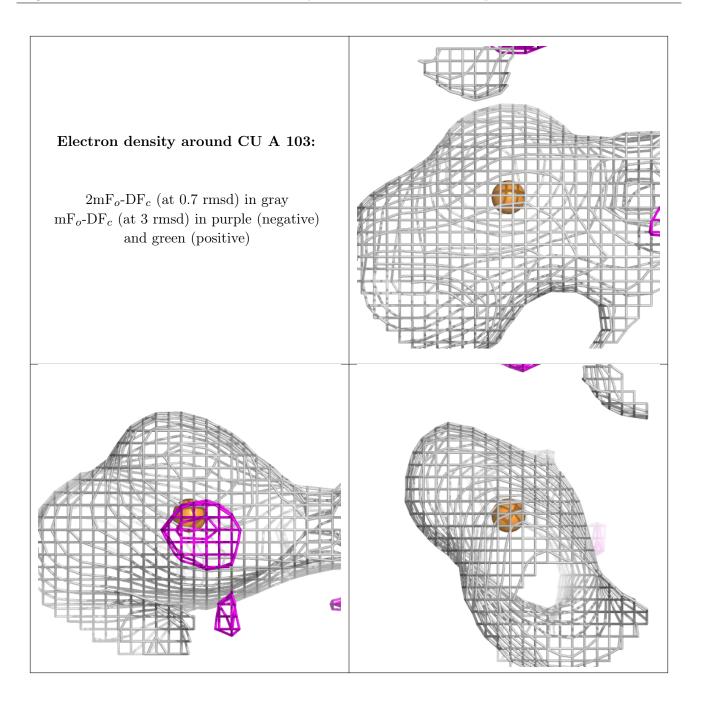




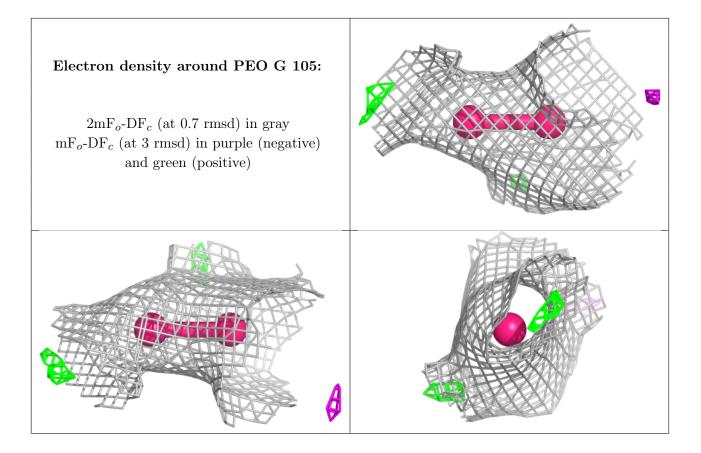




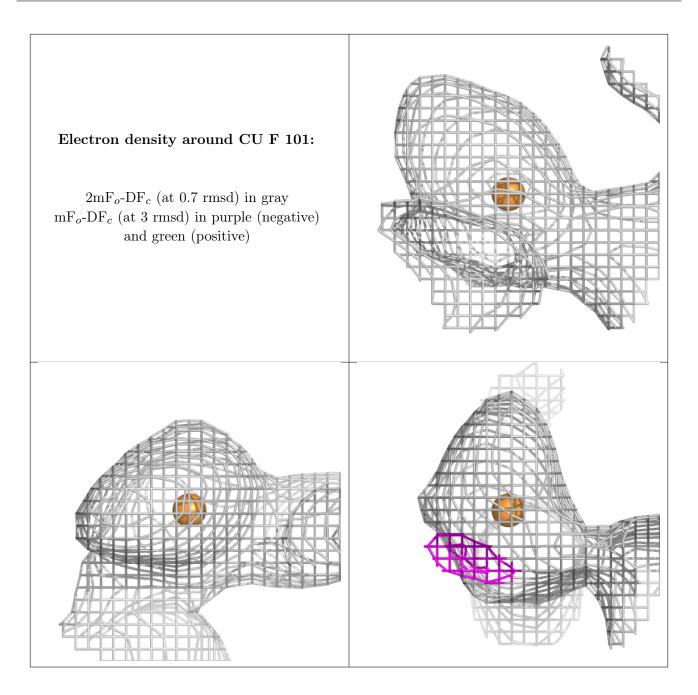












6.5 Other polymers (i)

There are no such residues in this entry.

