

# wwPDB X-ray Structure Validation Summary Report (i)

#### Sep 25, 2023 – 08:42 AM EDT

calprotectin

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 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:

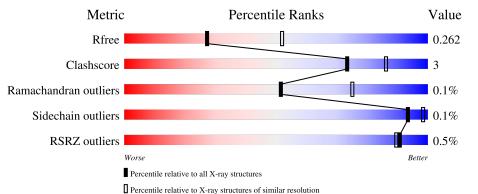
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
$\mathrm{EDS}$	:	2.35.1
buster-report	:	1.1.7(2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

# 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 2.60 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	$\begin{array}{c} \textbf{Whole archive} \\ (\#\textbf{Entries}) \end{array}$	${f Similar\ resolution}\ (\#{ m Entries,\ resolution\ range}({ m \AA}))$
$R_{free}$	130704	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)
RSRZ outliers	127900	3104 (2.60-2.60)

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	А	93	.%	9% •	6%
1	С	93	83%	12%	5%
1	Е	93	87%	8%	5%
1	G	93	83%	12%	5%
2	В	114	91%	•	6%

Continued on next page...



Continued from previous page...

Mol	Chain	Length	Quality of chain		
2	D	114	89%	5%	5%
2	F	114	85%	8%	7%
2	Н	114	81%	13%	6%



# 2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 6610 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.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	Λ	87	Total	С	Ν	0	$\mathbf{S}$	0	0 2	0
	А	01	728	475	120	131	2	0		0
1	С	88	Total	С	Ν	0	S	0	1	0
	U	00	729	472	121	134	2	0	1	0
1	C	G 88	Total	С	Ν	0	S	0	2	0
	I G		735	476	121	136	2	0		0
1	1 E	<u> </u>	Total	С	Ν	0	S	0	2	0
		88	737	480	121	134	2			U

• Molecule 1 is a protein called Protein S100-A8.

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
А	42	SER	CYS	engineered mutation	UNP P05109
С	42	SER	CYS	engineered mutation	UNP P05109
G	42	SER	CYS	engineered mutation	UNP P05109
Е	42	SER	CYS	engineered mutation	UNP P05109

• Molecule 2 is a protein called Protein S100-A9.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
2	В	107	Total	С	Ν	Ο	$\mathbf{S}$	0	0	0	
	D	107	879	550	155	169	5	0	0	0	
2	Л	108	Total	С	Ν	0	S	0	0	0	
	D	108	883	552	156	170	5	0	0	0	
2	Н	107	Total	С	Ν	0	S	0	0	0	
	11	107	879	550	155	169	5	0	0	0	
2	F	Б	106	Total	С	Ν	0	S	0	0	0
	Г	100	870	545	154	166	5	0	0	U	

There are 4 discrepancies between the modelled and reference sequences:



Chain	Residue	Modelled	Actual	Comment	Reference
В	3	SER	CYS	engineered mutation	UNP P06702
D	3	SER	CYS	engineered mutation	UNP P06702
Н	3	SER	CYS	engineered mutation	UNP P06702
F	3	SER	CYS	engineered mutation	UNP P06702

• Molecule 3 is SODIUM ION (three-letter code: NA) (formula: Na) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	А	1	Total Na 1 1	0	0
3	В	1	Total Na 1 1	0	0
3	С	1	Total Na 1 1	0	0
3	G	1	Total Na 1 1	0	0
3	Н	1	Total Na 1 1	0	0
3	Ε	1	Total Na 1 1	0	0

• Molecule 4 is CALCIUM ION (three-letter code: CA) (formula: Ca) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	А	1	Total Ca 1 1	0	0
4	В	1	Total Ca 1 1	0	0
4	С	1	Total Ca 1 1	0	0
4	D	2	Total Ca 2 2	0	0
4	G	1	Total Ca 1 1	0	0
4	Н	1	Total Ca 1 1	0	0
4	Е	1	Total Ca 1 1	0	0
4	F	2	Total Ca 2 2	0	0



• Molecule 5 is NICKEL (II) ION (three-letter code: NI) (formula: Ni) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	В	1	Total Ni 1 1	0	0
5	С	1	Total Ni 1 1	0	0
5	D	1	Total Ni 1 1	0	0
5	Н	1	Total Ni 1 1	0	0
5	Ε	1	Total Ni 1 1	0	0
5	F	1	Total Ni 1 1	0	0

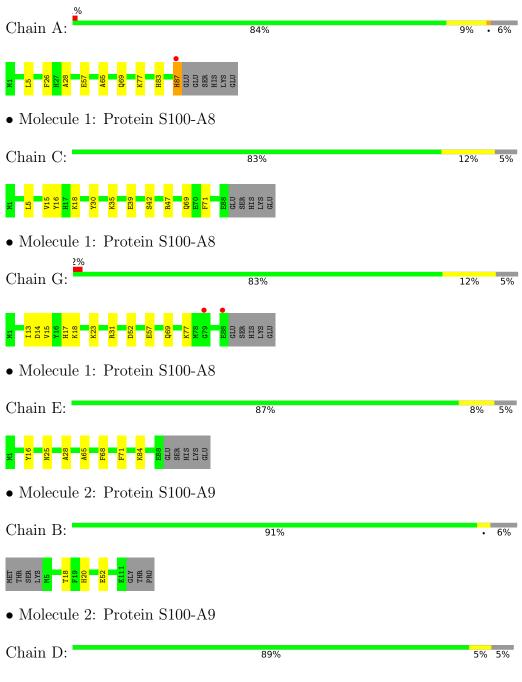
• Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	А	22	TotalO2222	0	0
6	В	20	TotalO2020	0	0
6	С	17	Total O 17 17	0	0
6	D	19	Total O 19 19	0	0
6	G	15	Total O 15 15	0	0
6	Н	16	Total O 16 16	0	0
6	Е	14	Total         O           14         14	0	0
6	F	25	Total O 25 25	0	0



# 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.

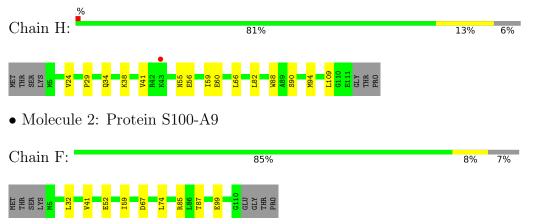


• Molecule 1: Protein S100-A8





• Molecule 2: Protein S100-A9





# 4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants	57.02Å 77.61Å 222.72Å	Depositor
a, b, c, $\alpha$ , $\beta$ , $\gamma$	$90.00^{\circ}$ $90.00^{\circ}$ $90.00^{\circ}$	Depositor
Resolution (Å)	45.24 - 2.60	Depositor
Resolution (A)	45.95 - 2.59	EDS
% Data completeness	93.7 (45.24-2.60)	Depositor
(in resolution range)	$93.7 \ (45.95 - 2.59)$	EDS
R <sub>merge</sub>	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$< I/\sigma(I) > 1$	$4.78 (at 2.58 \text{\AA})$	Xtriage
Refinement program	PHENIX (1.11.1_2575: ???)	Depositor
D D.	0.224 , $0.262$	Depositor
$R, R_{free}$	0.224 , $0.262$	DCC
$R_{free}$ test set	989 reflections $(3.35\%)$	wwPDB-VP
Wilson B-factor $(Å^2)$	32.5	Xtriage
Anisotropy	1.078	Xtriage
Bulk solvent $k_{sol}(e/Å^3), B_{sol}(Å^2)$	0.33, 27.5	EDS
L-test for twinning <sup>2</sup>	$ < L >=0.48, < L^2>=0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	6610	wwPDB-VP
Average B, all atoms $(Å^2)$	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 7.63% of the height of the origin peak. No significant pseudotranslation is detected.

<sup>&</sup>lt;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.



<sup>&</sup>lt;sup>1</sup>Intensities 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: NA, NI, CA

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
IVIOI	Unam	RMSZ	# Z  > 5	RMSZ	# Z  > 5
1	А	0.24	0/746	0.36	0/1002
1	С	0.23	0/743	0.36	0/998
1	Е	0.23	0/755	0.35	0/1014
1	G	0.23	0/752	0.36	0/1010
2	В	0.23	0/897	0.36	0/1202
2	D	0.23	0/901	0.36	0/1207
2	F	0.23	0/888	0.36	0/1190
2	Н	0.23	0/897	0.36	0/1202
All	All	0.23	0/6579	0.36	0/8825

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	728	0	744	7	0
1	С	729	0	741	8	0
1	Е	737	0	750	5	0
1	G	735	0	747	7	0
2	В	879	0	849	3	0

Continued on next page...



5W	$1\mathrm{F}$
----	---------------

Mol	Chain	<b>Non-H</b>	H(model)	H(added)	Clashes	Symm-Clashes
2	D	883	0	852	5	0
2	F	870	0	843	6	0
2	Н	879	0	849	10	0
3	А	1	0	0	0	0
3	В	1	0	0	0	0
3	С	1	0	0	0	0
3	Ε	1	0	0	0	0
3	G	1	0	0	0	0
3	Η	1	0	0	0	0
4	А	1	0	0	0	0
4	В	1	0	0	0	0
4	С	1	0	0	0	0
4	D	2	0	0	0	0
4	Ε	1	0	0	0	0
4	F	2	0	0	0	0
4	G	1	0	0	0	0
4	Н	1	0	0	0	0
5	В	1	0	0	0	0
5	С	1	0	0	0	0
5	D	1	0	0	0	0
5	Ε	1	0	0	0	0
5	F	1	0	0	0	0
5	Н	1	0	0	0	0
6	А	22	0	0	1	0
6	В	20	0	0	0	0
6	С	17	0	0	1	0
6	D	19	0	0	0	0
6	Ε	14	0	0	0	0
6	F	25	0	0	1	0
6	G	15	0	0	0	0
6	Н	16	0	0	0	0
All	All	6610	0	6375	42	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 3.

The worst 5 of 42 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)
2:F:85:ARG:NH2	6:F:301:HOH:O	2.33	0.60
1:A:87:HIS:CE1	2:B:20:HIS:HE1	2.21	0.59

Continued on next page...



Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:G:15:VAL:HA	1:G:18:LYS:HE2	1.85	0.59
1:A:57[A]:GLU:OE2	1:A:77:LYS:NZ	2.27	0.59
1:A:69:GLN:NE2	6:A:203:HOH:O	2.40	0.55

Continued from previous page...

There are no symmetry-related clashes.

#### 5.3 Torsion angles (i)

#### 5.3.1 Protein backbone (i)

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	Perce	entiles
1	А	87/93~(94%)	85~(98%)	2(2%)	0	100	100
1	С	87/93~(94%)	85~(98%)	2(2%)	0	100	100
1	Ε	88/93~(95%)	85~(97%)	3~(3%)	0	100	100
1	G	88/93~(95%)	86~(98%)	2(2%)	0	100	100
2	В	105/114~(92%)	103~(98%)	2(2%)	0	100	100
2	D	106/114~(93%)	103~(97%)	3~(3%)	0	100	100
2	$\mathbf{F}$	104/114~(91%)	101 (97%)	2(2%)	1 (1%)	15	32
2	Н	105/114~(92%)	102~(97%)	3~(3%)	0	100	100
All	All	770/828~(93%)	750~(97%)	19 (2%)	1 (0%)	51	75

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	F	67	ASP

#### 5.3.2 Protein sidechains (i)

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.



Mol	Chain	Analysed	Rotameric	Outliers	Perce	ntiles
1	А	79/83~(95%)	78~(99%)	1 (1%)	69	86
1	$\mathbf{C}$	79/83~(95%)	79~(100%)	0	100	100
1	Ε	80/83~(96%)	80 (100%)	0	100	100
1	G	80/83~(96%)	80 (100%)	0	100	100
2	В	97/103~(94%)	97~(100%)	0	100	100
2	D	97/103~(94%)	97~(100%)	0	100	100
2	F	96/103~(93%)	96 (100%)	0	100	100
2	Н	97/103~(94%)	97~(100%)	0	100	100
All	All	705/744~(95%)	704 (100%)	1 (0%)	93	98

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	А	87	HIS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
2	В	20	HIS

#### 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 22 ligands modelled in this entry, 22 are monoatomic - leaving 0 for Mogul analysis. 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.

### 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	<RSRZ $>$	# RSRZ > 2	$\mathbf{OWAB}(\mathrm{\AA}^2)$	Q<0.9
1	А	87/93~(93%)	-0.06	1 (1%) 80 78	29, 41, 60, 106	0
1	С	88/93~(94%)	-0.18	0 100 100	23, 33, 47, 56	0
1	E	88/93~(94%)	-0.26	0 100 100	26, 38, 49, 58	0
1	G	88/93~(94%)	0.11	2 (2%) 60 54	26, 38, 64, 97	0
2	В	107/114~(93%)	-0.05	0 100 100	32, 47, 62, 68	0
2	D	108/114~(94%)	-0.15	0 100 100	27, 38, 53, 69	0
2	F	106/114~(92%)	-0.23	0 100 100	23, 34, 56, 64	0
2	Н	107/114~(93%)	0.14	1 (0%) 84 82	30, 49, 62, 73	0
All	All	779/828~(94%)	-0.08	4 (0%) 91 89	23, 40, 60, 106	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	88	GLU	3.7
1	А	87	HIS	3.6
1	G	79	GLY	3.0
2	Н	43	LYS	2.7

## 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.



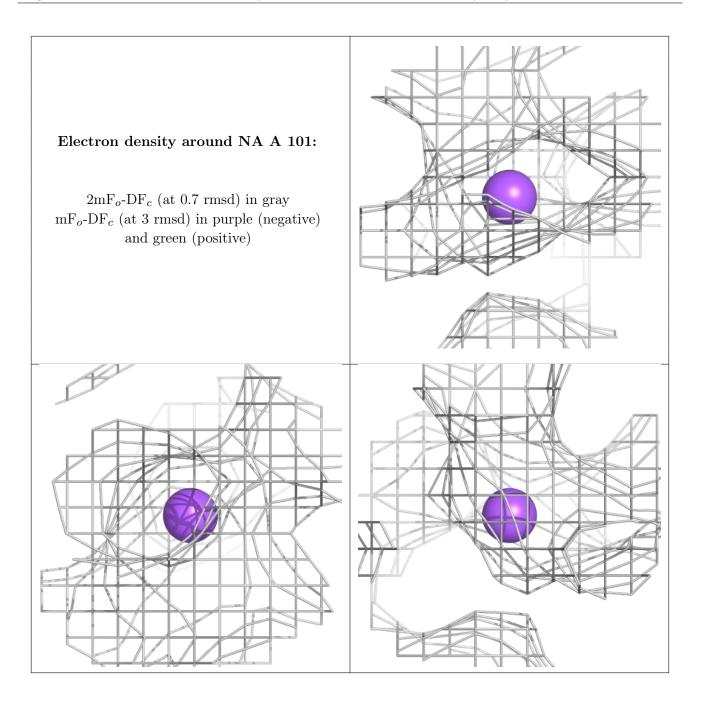
### 6.4 Ligands (i)

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.

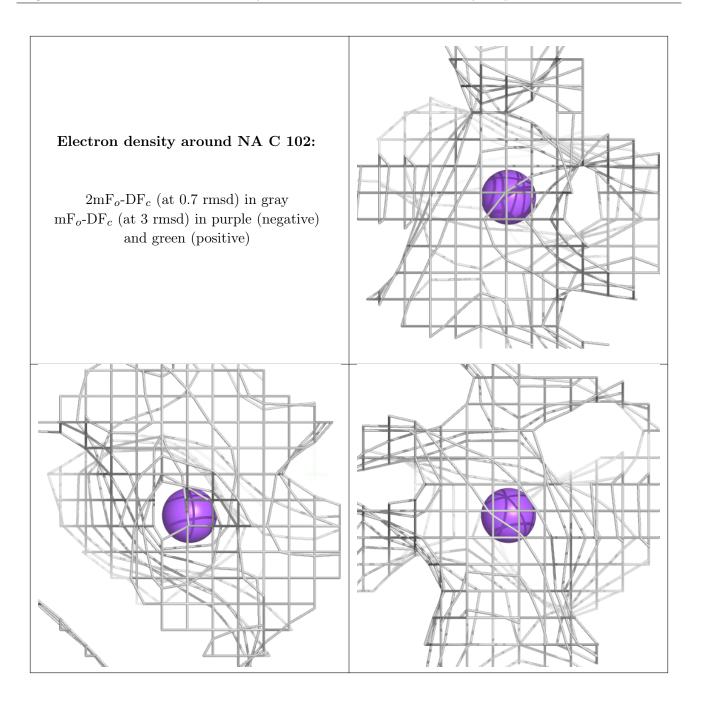
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q < 0.9
3	NA	А	101	1/1	0.77	0.13	$50,\!50,\!50,\!50$	0
3	NA	С	102	1/1	0.87	0.17	34,34,34,34	0
4	CA	Н	203	1/1	0.89	0.10	$53,\!53,\!53,\!53$	0
3	NA	В	202	1/1	0.90	0.09	31,31,31,31	0
3	NA	Н	202	1/1	0.94	0.13	46,46,46,46	0
4	CA	Е	103	1/1	0.95	0.09	42,42,42,42	0
3	NA	Е	102	1/1	0.96	0.14	33,33,33,33	0
4	CA	В	203	1/1	0.97	0.11	46,46,46,46	0
4	CA	D	202	1/1	0.97	0.07	34,34,34,34	0
4	CA	F	202	1/1	0.97	0.09	33,33,33,33	0
4	CA	F	203	1/1	0.97	0.16	$37,\!37,\!37,\!37$	0
5	NI	С	101	1/1	0.98	0.11	31,31,31,31	1
3	NA	G	101	1/1	0.99	0.14	22,22,22,22	0
4	CA	С	103	1/1	0.99	0.08	31,31,31,31	0
4	CA	А	102	1/1	0.99	0.08	26,26,26,26	0
4	CA	D	203	1/1	0.99	0.05	33,33,33,33	0
5	NI	В	201	1/1	0.99	0.08	36,36,36,36	0
4	CA	G	102	1/1	0.99	0.10	34,34,34,34	0
5	NI	D	201	1/1	0.99	0.11	30,30,30,30	0
5	NI	Н	201	1/1	0.99	0.18	32,32,32,32	0
5	NI	Е	101	1/1	0.99	0.17	30,30,30,30	1
5	NI	F	201	1/1	0.99	0.13	39,39,39,39	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.

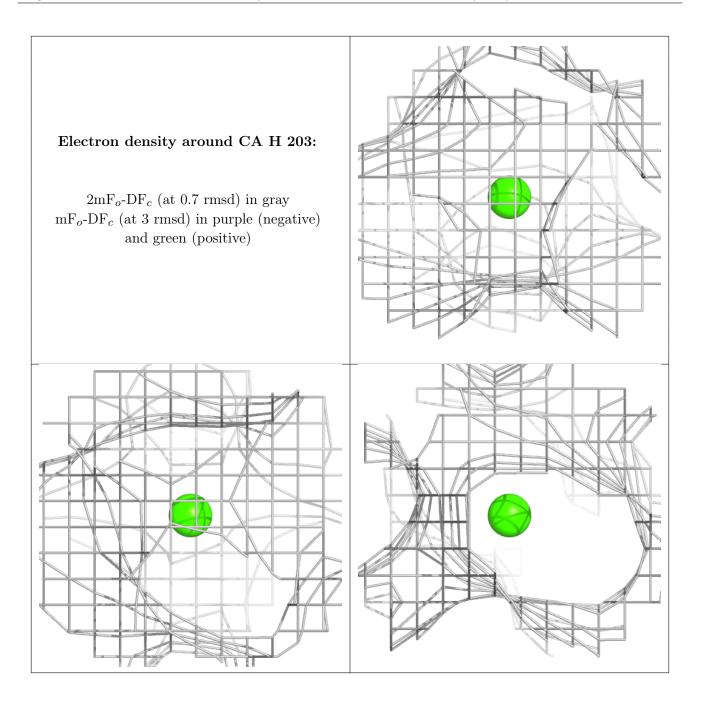




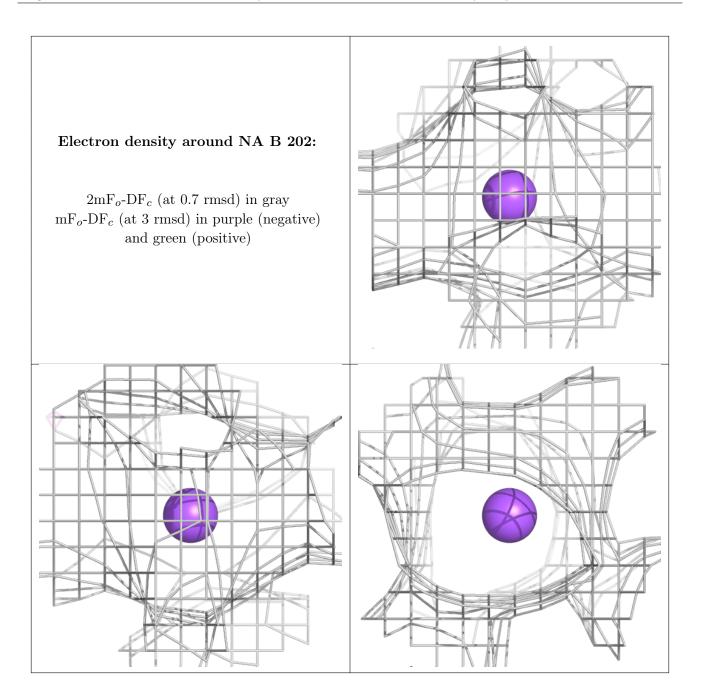




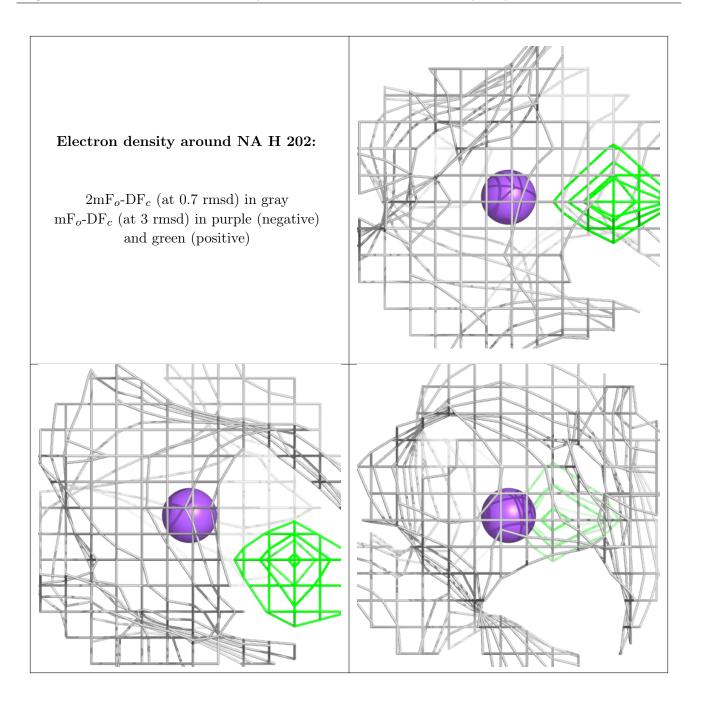




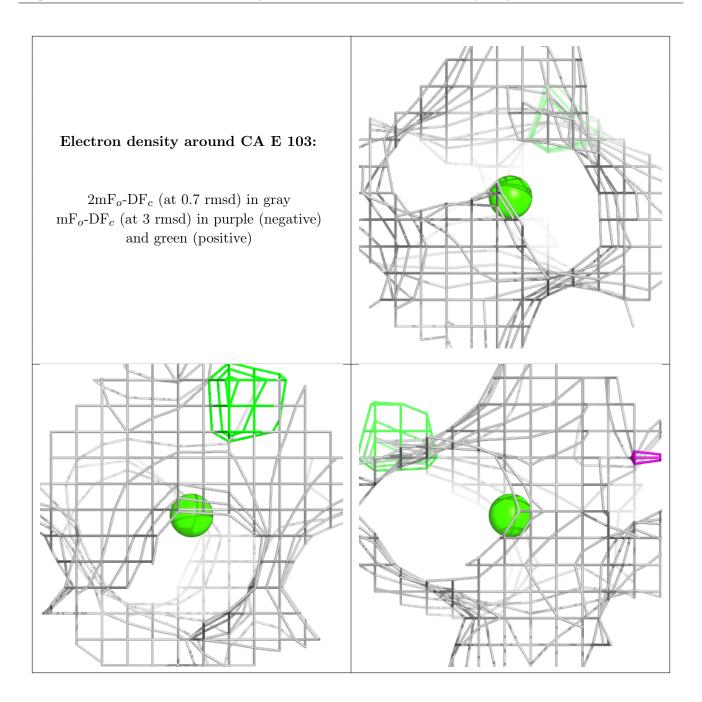




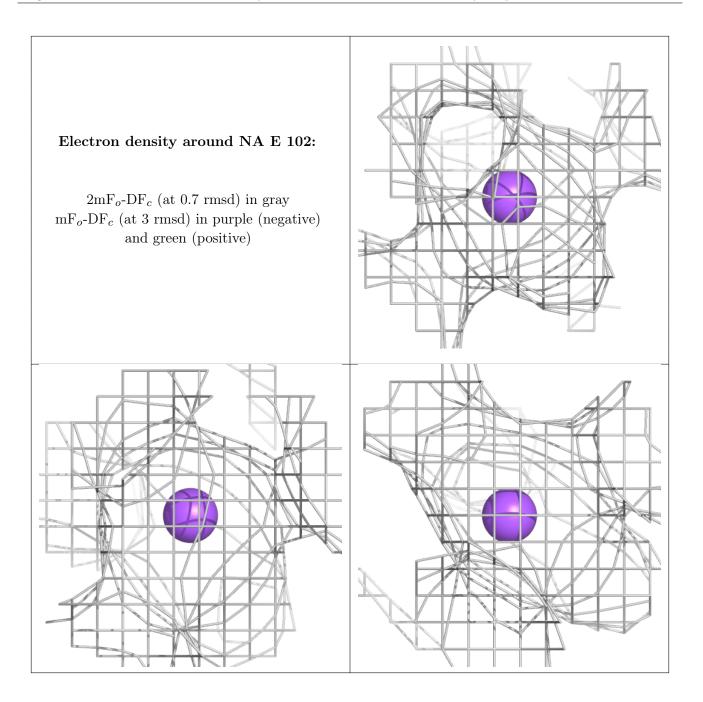




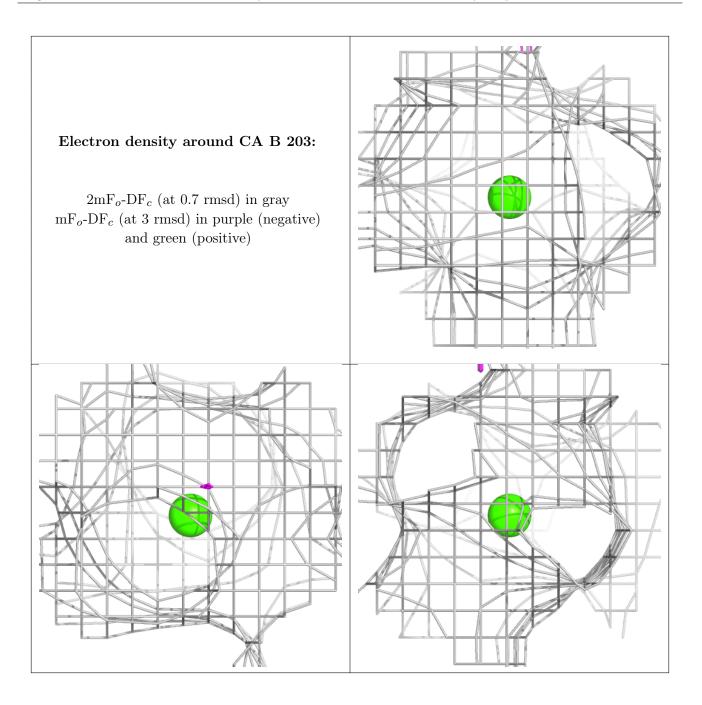




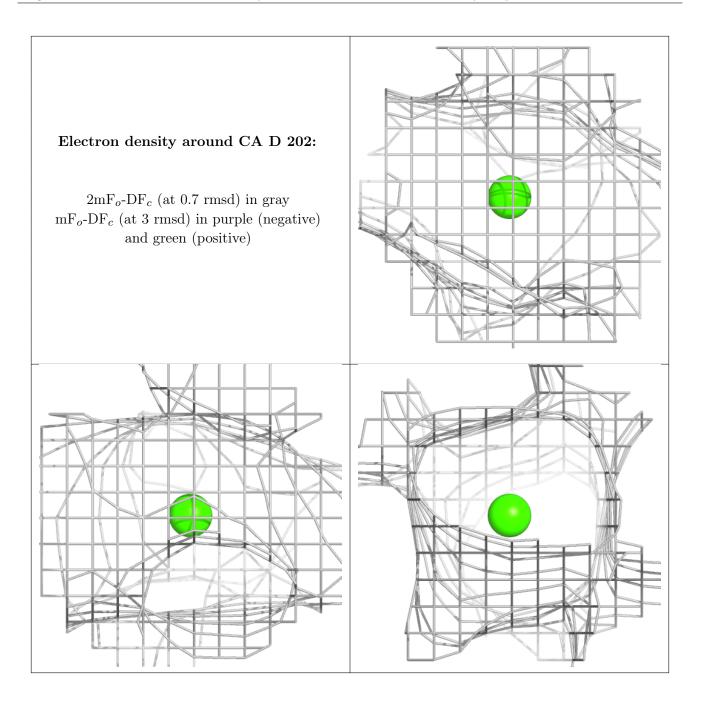




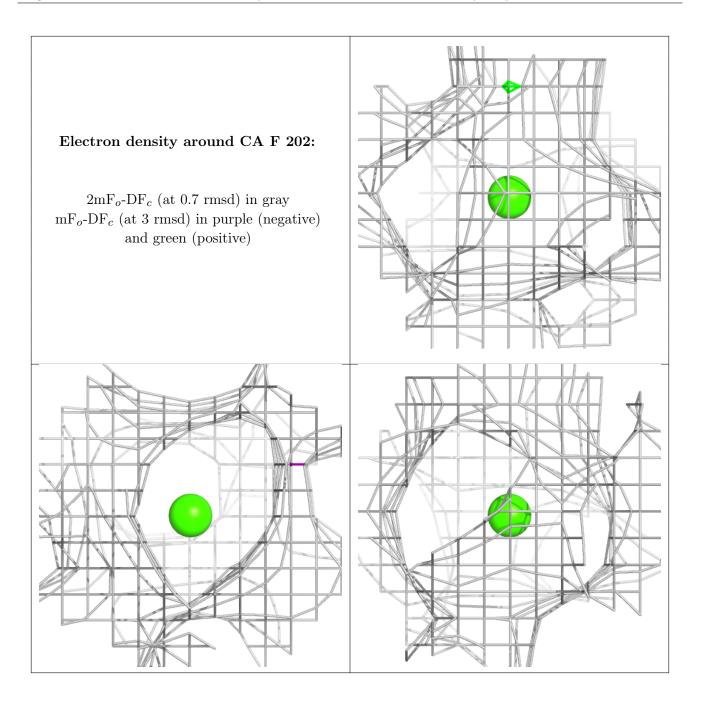




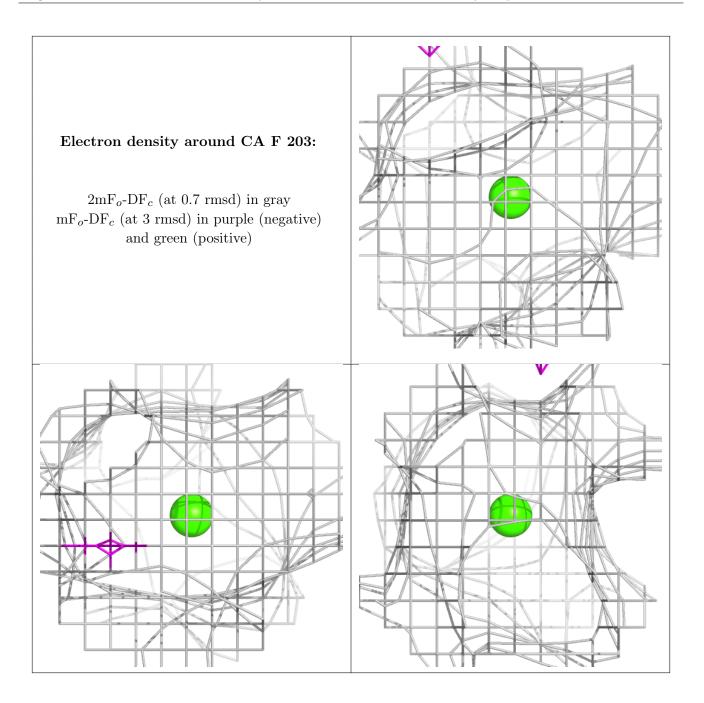




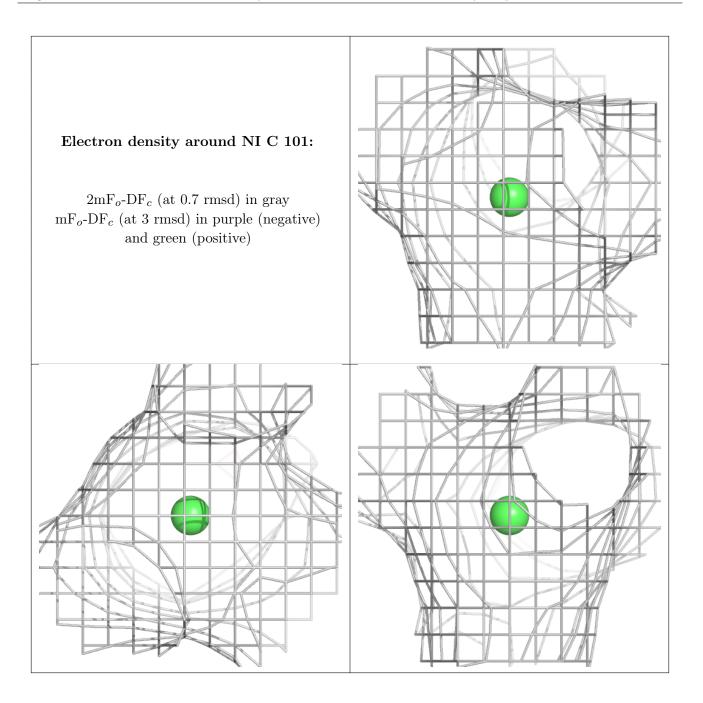




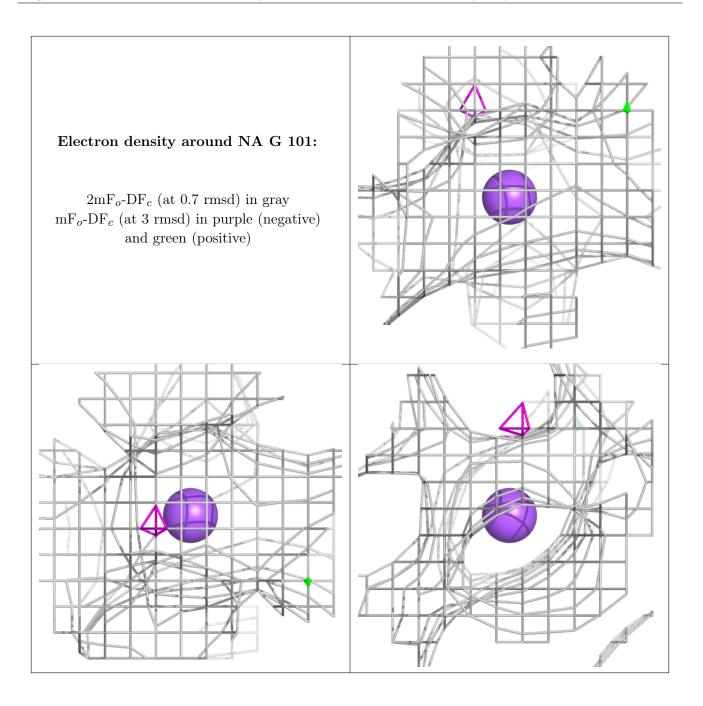




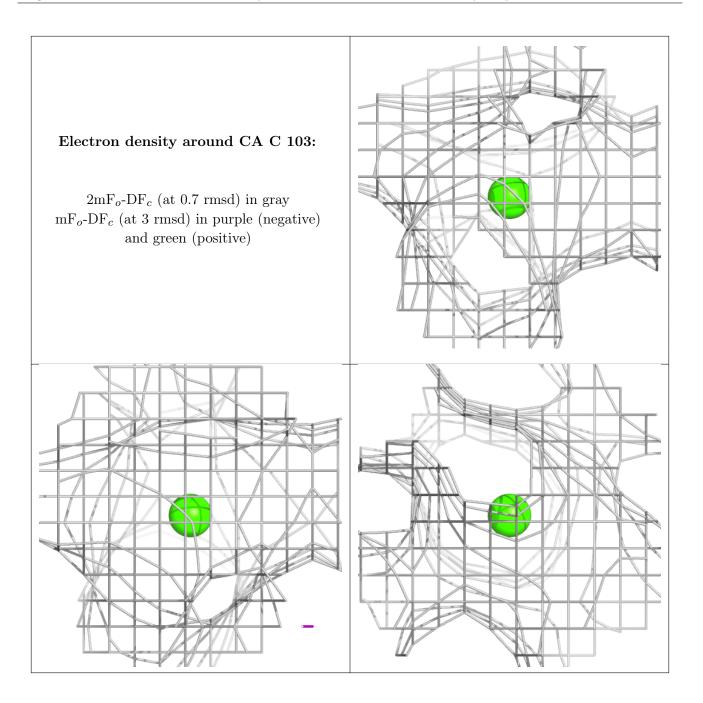




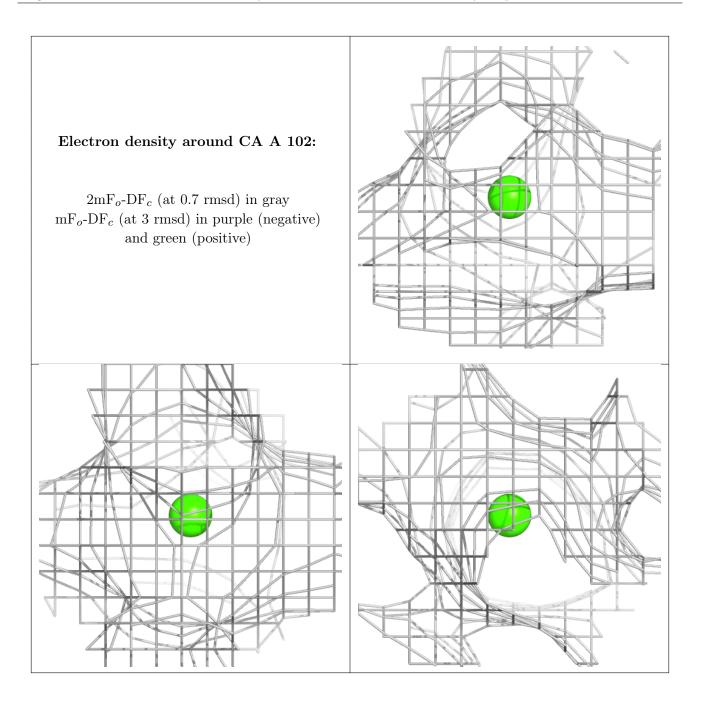




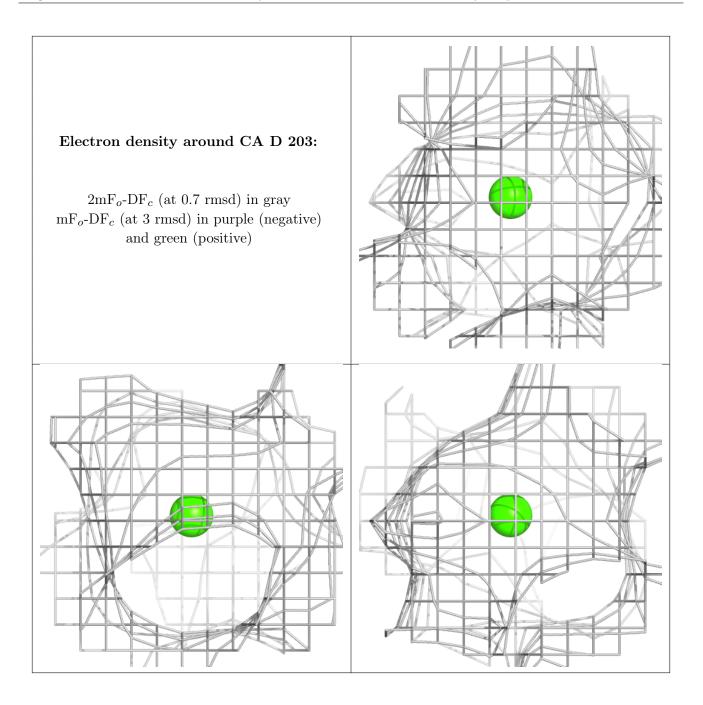




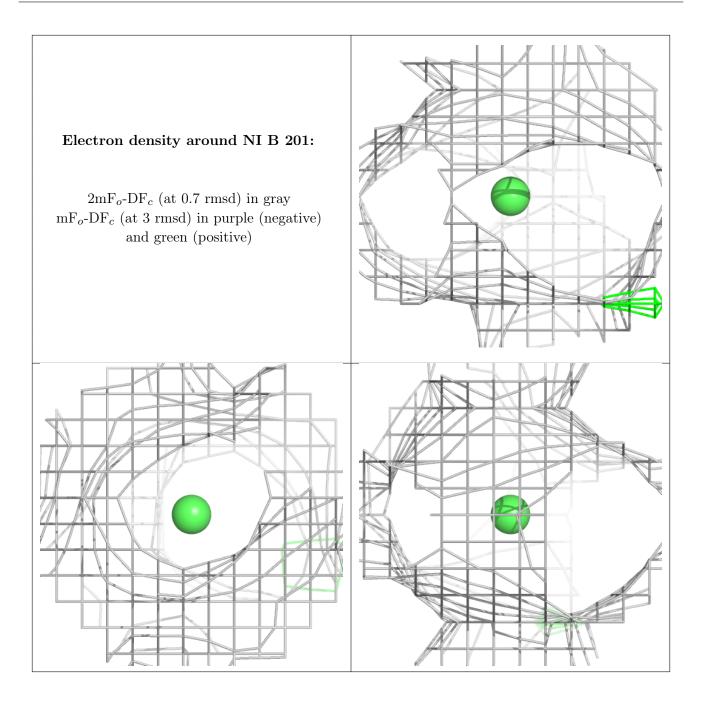




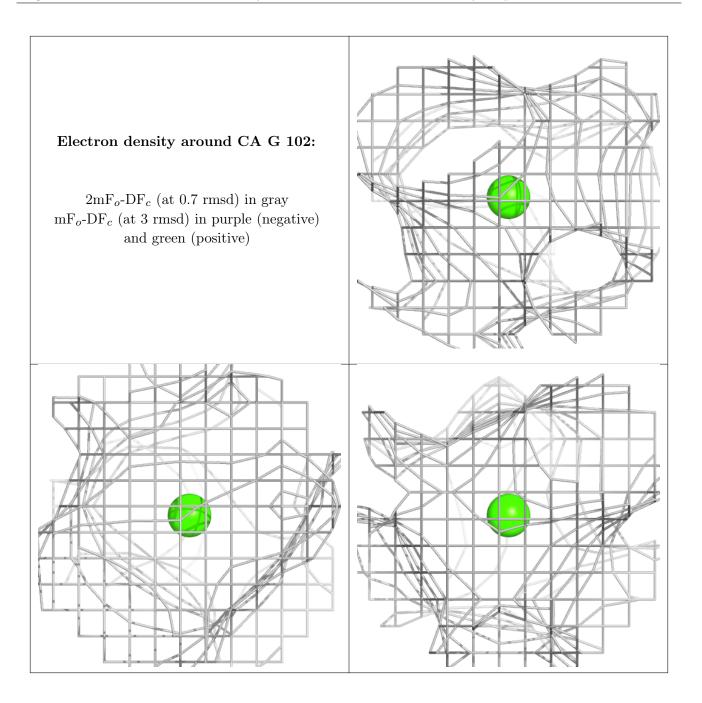




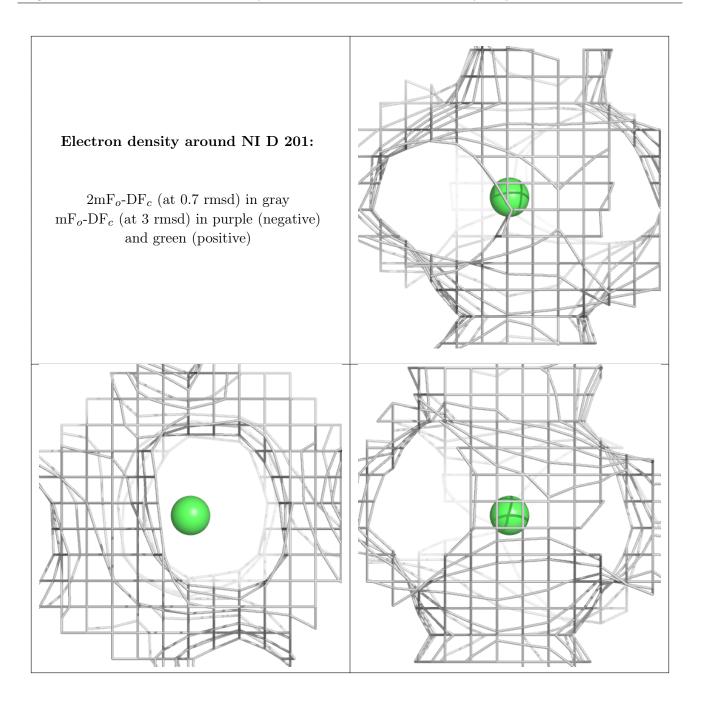




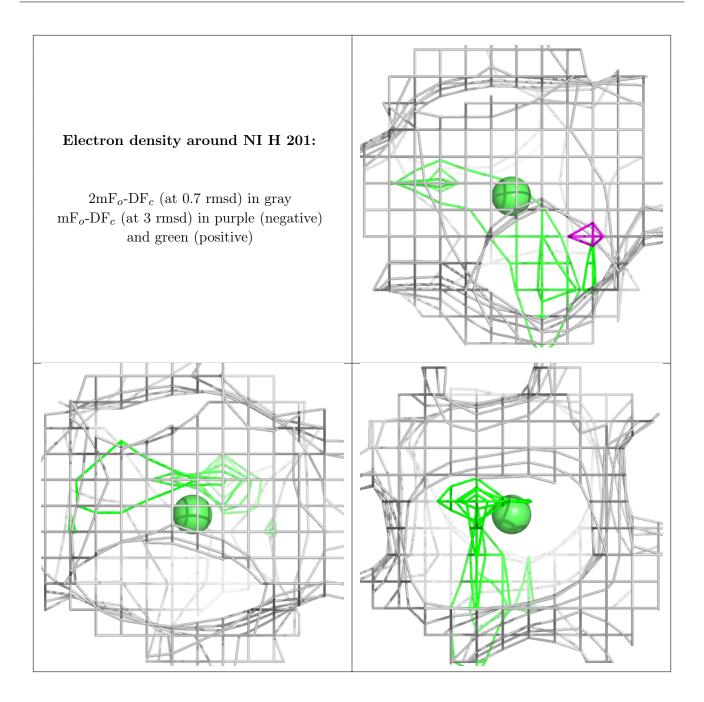




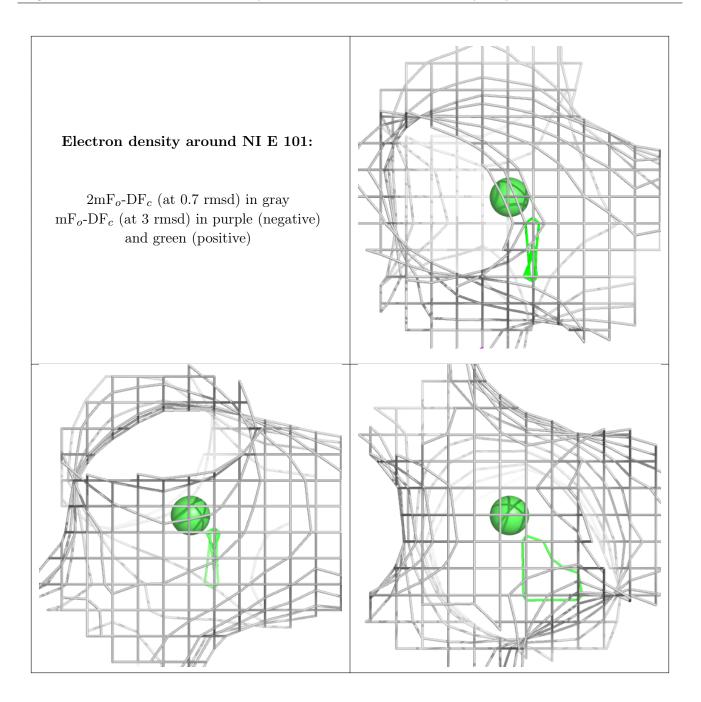




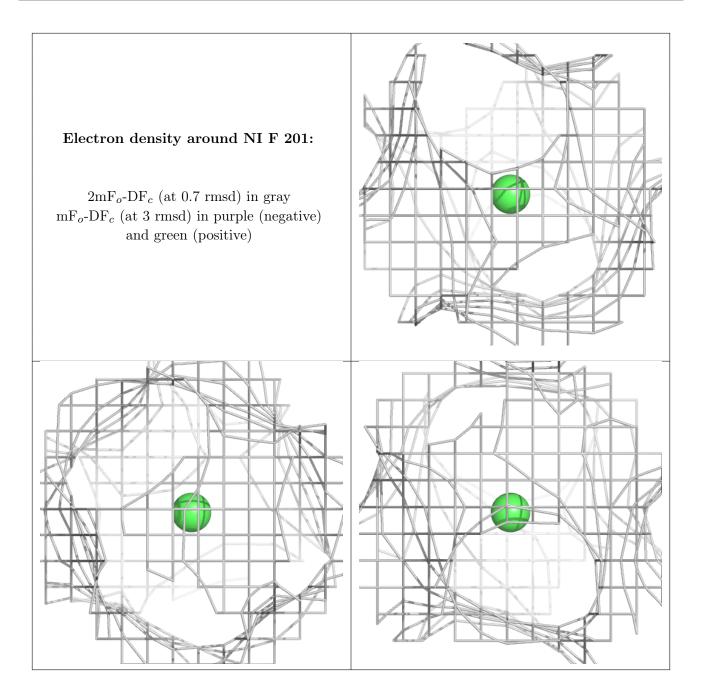












## 6.5 Other polymers (i)

There are no such residues in this entry.

