

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

Jun 22, 2024 – 11:54 PM EDT

PDB ID	:	6SJ0
Title	:	Amidohydrolase, AHS
Authors	:	Naismith, J.H.; Song, H.
1		2019-08-12
Resolution	:	1.75  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 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:

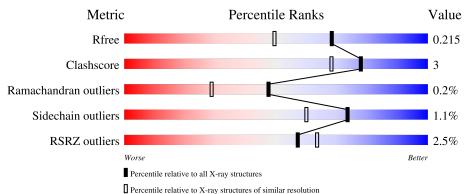
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
$\mathrm{EDS}$	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

# 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 1.75 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 <sub>free</sub>	130704	2340 (1.76-1.76)
Clashscore	141614	2466 (1.76-1.76)
Ramachandran outliers	138981	2437 (1.76-1.76)
Sidechain outliers	138945	2437 (1.76-1.76)
RSRZ outliers	127900	2298 (1.76-1.76)

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	519	89%	6%	5%			
1	В	519	88%	6%	5%			



# 2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 7911 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 Amidohydrolase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1		493	Total	С	Ν	0	S	0	4	0
	490	3758	2332	689	713	24	0	4	0	
1	1 B	B 493	Total	С	Ν	0	S	0	9	0
		490	3740	2318	689	711	22	0		0

Chain	Residue	Modelled	Actual	Comment	Reference
А	-24	MET	-	initiating methionine	UNP A0A022MQ12
А	-23	SER	-	expression tag	UNP A0A022MQ12
А	-22	TYR	-	expression tag	UNP A0A022MQ12
А	-21	TYR	-	expression tag	UNP A0A022MQ12
А	-20	HIS	-	expression tag	UNP A0A022MQ12
А	-19	HIS	-	expression tag	UNP A0A022MQ12
А	-18	HIS	-	expression tag	UNP A0A022MQ12
А	-17	HIS	-	expression tag	UNP A0A022MQ12
А	-16	HIS	-	expression tag	UNP A0A022MQ12
А	-15	HIS	-	expression tag	UNP A0A022MQ12
А	-14	ASP	-	expression tag	UNP A0A022MQ12
А	-13	TYR	-	expression tag	UNP A0A022MQ12
А	-12	ASP	-	expression tag	UNP A0A022MQ12
А	-11	ILE	-	expression tag	UNP A0A022MQ12
А	-10	PRO	-	expression tag	UNP A0A022MQ12
А	-9	THR	-	expression tag	UNP A0A022MQ12
А	-8	THR	-	expression tag	UNP A0A022MQ12
А	-7	GLU	-	expression tag	UNP A0A022MQ12
А	-6	ASN	-	expression tag	UNP A0A022MQ12
А	-5	LEU	-	expression tag	UNP A0A022MQ12
А	-4	TYR	-	expression tag	UNP A0A022MQ12
А	-3	PHE	-	expression tag	UNP A0A022MQ12
А	-2	GLN	-	expression tag	UNP A0A022MQ12
А	-1	GLY	-	expression tag	UNP A0A022MQ12
А	0	ALA	-	expression tag	UNP A0A022MQ12

There are 50 discrepancies between the modelled and reference sequences:

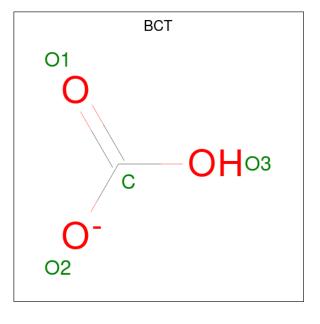
Continued on next page...



Chain	Residue	Modelled	Actual	Comment	Reference
В	-24	MET	_	initiating methionine	UNP A0A022MQ12
В	-23	SER	_	expression tag	UNP A0A022MQ12
В	-22	TYR	-	expression tag	UNP A0A022MQ12
В	-21	TYR	-	expression tag	UNP A0A022MQ12
В	-20	HIS	-	expression tag	UNP A0A022MQ12
В	-19	HIS	-	expression tag	UNP A0A022MQ12
В	-18	HIS	-	expression tag	UNP A0A022MQ12
В	-17	HIS	-	expression tag	UNP A0A022MQ12
В	-16	HIS	-	expression tag	UNP A0A022MQ12
В	-15	HIS	-	expression tag	UNP A0A022MQ12
В	-14	ASP	-	expression tag	UNP A0A022MQ12
В	-13	TYR	-	expression tag	UNP A0A022MQ12
В	-12	ASP	-	expression tag	UNP A0A022MQ12
В	-11	ILE	-	expression tag	UNP A0A022MQ12
В	-10	PRO	-	expression tag	UNP A0A022MQ12
В	-9	THR	-	expression tag	UNP A0A022MQ12
В	-8	THR	-	expression tag	UNP A0A022MQ12
В	-7	GLU	-	expression tag	UNP A0A022MQ12
В	-6	ASN	-	expression tag	UNP A0A022MQ12
В	-5	LEU	-	expression tag	UNP A0A022MQ12
В	-4	TYR	-	expression tag	UNP A0A022MQ12
В	-3	PHE	-	expression tag	UNP A0A022MQ12
В	-2	GLN	-	expression tag	UNP A0A022MQ12
В	-1	GLY	-	expression tag	UNP A0A022MQ12
В	0	ALA	_	expression tag	UNP A0A022MQ12

Continued from previous page...

• Molecule 2 is BICARBONATE ION (three-letter code: BCT) (formula: CHO<sub>3</sub>).

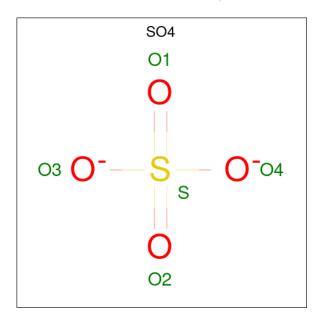






Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	А	1	Total 4	С 1	O 3	0	0

• Molecule 3 is SULFATE ION (three-letter code: SO4) (formula:  $O_4S$ ).



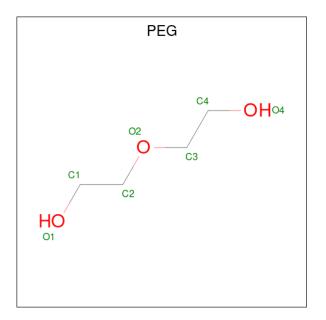
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	А	1	$\begin{array}{ccc} \text{Total} & \text{O} & \text{S} \\ 5 & 4 & 1 \end{array}$	0	0
3	А	1	$\begin{array}{ccc} \text{Total} & \text{O} & \text{S} \\ 5 & 4 & 1 \end{array}$	0	0

• Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	А	11	Total Zn 11 11	0	0
4	В	11	Total Zn 11 11	0	0

• Molecule 5 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula:  $C_4H_{10}O_3$ ).





Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	В	1	$\begin{array}{ccc} \text{Total} & \text{C} & \text{O} \\ 7 & 4 & 3 \end{array}$	0	0

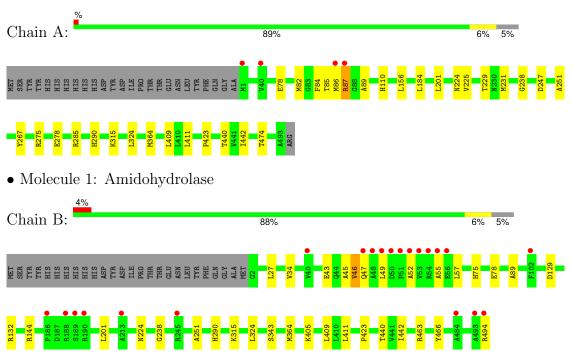
• Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	А	177	Total O 177 177	0	0
6	В	193	Total O 193 193	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: Amidohydrolase



## 4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants	81.79Å 82.75Å 141.56Å	Depositor
a, b, c, $\alpha$ , $\beta$ , $\gamma$	$90.00^{\circ}$ $90.00^{\circ}$ $90.00^{\circ}$	Depositor
Resolution (Å)	53.81 - 1.75	Depositor
	53.81 - 1.75	EDS
% Data completeness	100.0 (53.81 - 1.75)	Depositor
(in resolution range)	100.0 (53.81 - 1.75)	EDS
R <sub>merge</sub>	0.15	Depositor
R <sub>sym</sub>	(Not available)	Depositor
$< I/\sigma(I) > 1$	$1.85 (at 1.75 \text{\AA})$	Xtriage
Refinement program	REFMAC 5.8.0253	Depositor
D D	0.182 , $0.209$	Depositor
R, $R_{free}$	0.192 , $0.215$	DCC
$R_{free}$ test set	4973 reflections $(5.11%)$	wwPDB-VP
Wilson B-factor $(Å^2)$	25.6	Xtriage
Anisotropy	0.173	Xtriage
Bulk solvent $k_{sol}(e/Å^3)$ , $B_{sol}(Å^2)$	0.36, 37.1	EDS
L-test for twinning <sup>2</sup>	$< L >=0.49, < L^2>=0.33$	Xtriage
Estimated twinning fraction	0.012 for k,h,-l	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	7911	wwPDB-VP
Average B, all atoms $(Å^2)$	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 3.83% 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: BCT, PEG, SO4, ZN  $\,$ 

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	0.64	0/3823	0.77	0/5184	
1	В	0.64	0/3805	0.78	0/5161	
All	All	0.64	0/7628	0.78	0/10345	

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	3758	0	3716	27	0
1	В	3740	0	3690	18	0
2	А	4	0	0	0	0
3	А	10	0	0	0	0
4	А	11	0	0	0	0
4	В	11	0	0	0	0
5	В	7	0	10	0	0
6	А	177	0	0	1	0
6	В	193	0	0	2	0
All	All	7911	0	7416	43	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 3.

The worst 5 of 43 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:85[B]:THR:HG22	1:A:87[B]:ARG:H	1.22	1.04
1:A:85[B]:THR:CG2	1:A:87[B]:ARG:H	1.75	0.97
1:A:87[A]:ARG:O	1:A:87[A]:ARG:HG3	1.79	0.82
1:B:45:ALA:O	1:B:49:LEU:HD13	1.85	0.77
1:A:85[B]:THR:HG22	1:A:87[B]:ARG:N	2.00	0.74

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	Percentiles
1	А	495/519~(95%)	487 (98%)	7 (1%)	1 (0%)	47 29
1	В	493/519~(95%)	483 (98%)	9(2%)	1 (0%)	47 29
All	All	988/1038~(95%)	970 (98%)	16 (2%)	2(0%)	47 29

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	В	290	HIS
1	А	290	HIS

#### 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.

The Analysed column shows the number of residues for which the sidechain conformation was



Mol	Chain	Analysed	Rotameric	Outliers	Percentiles		
1	А	377/397~(95%)	373~(99%)	4 (1%)	73	60	
1	В	375/397~(94%)	370~(99%)	5 (1%)	69	54	
All	All	752/794~(95%)	743 (99%)	9 (1%)	73	56	

analysed, and the total number of residues.

5 of 9 residues with a non-rotameric side chain are listed below:

Mol	Chain	Res	Type
1	В	315	LYS
1	В	466	TYR
1	А	315	LYS
1	В	46	VAL
1	В	57	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

#### 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 26 ligands modelled in this entry, 22 are monoatomic - leaving 4 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	B	ond leng	$\operatorname{gths}$	В	Bond ang	gles
10101	туре	Unam	nes	LIIIK	Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z >2
2	BCT	А	501	-	3,3,3	1.44	0	$2,\!3,\!3$	0.84	0
5	PEG	В	501	-	6,6,6	0.15	0	$5,\!5,\!5$	0.15	0
3	SO4	А	502	-	4,4,4	0.26	0	$6,\!6,\!6$	0.09	0
3	SO4	А	503	-	4,4,4	0.31	0	$6,\!6,\!6$	0.15	0

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
5	PEG	В	501	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	В	501	PEG	O2-C3-C4-O4
5	В	501	PEG	C4-C3-O2-C2

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	$\langle RSRZ \rangle$	#RSRZ>2	$\mathbf{OWAB}(\mathrm{\AA}^2)$	Q < 0.9
1	А	493/519~(94%)	-0.12	4 (0%) 86 90	16, 27, 49, 79	0
1	В	493/519~(94%)	-0.00	21 (4%) 35 41	15, 25, 58, 100	0
All	All	986/1038~(94%)	-0.06	25 (2%) 57 63	15, 26, 52, 100	0

The worst 5 of 25 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	В	494	ARG	6.7
1	В	53	VAL	5.0
1	В	190	ARG	4.7
1	В	51	PRO	4.2
1	В	48	ALA	4.2

### 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)

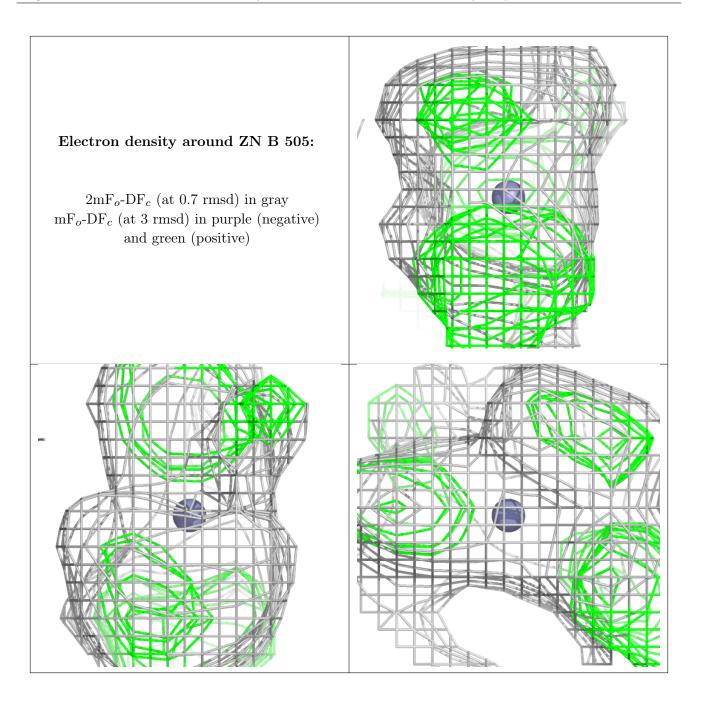
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median,  $95^{th}$  percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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



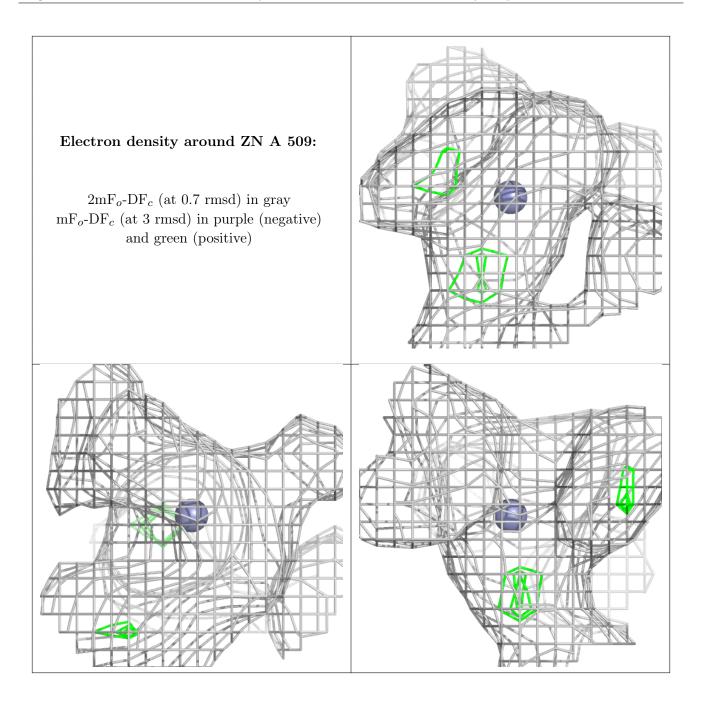
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
4	ZN	В	505	1/1	0.41	0.10	$105,\!105,\!105,\!105$	0
4	ZN	А	509	1/1	0.45	0.15	112,112,112,112	0
4	ZN	А	512	1/1	0.51	0.12	104,104,104,104	0
4	ZN	А	514	1/1	0.62	0.09	89,89,89,89	0
4	ZN	А	510	1/1	0.62	0.09	90,90,90,90	0
4	ZN	В	512	1/1	0.62	0.07	105,105,105,105	0
4	ZN	А	505	1/1	0.77	0.17	109,109,109,109	0
5	PEG	В	501	7/7	0.81	0.17	50,54,57,65	0
2	BCT	А	501	4/4	0.83	0.15	55,57,58,60	0
4	ZN	В	506	1/1	0.88	0.17	85,85,85,85	0
4	ZN	А	511	1/1	0.89	0.17	85,85,85,85	0
4	ZN	В	504	1/1	0.89	0.12	60,60,60,60	0
3	SO4	А	502	5/5	0.90	0.21	35,50,62,71	0
4	ZN	А	508	1/1	0.90	0.18	83,83,83,83	0
4	ZN	В	510	1/1	0.92	0.09	87,87,87,87	0
4	ZN	А	506	1/1	0.93	0.07	56, 56, 56, 56	0
4	ZN	А	507	1/1	0.94	0.06	64,64,64,64	0
4	ZN	А	513	1/1	0.94	0.05	$55,\!55,\!55,\!55$	0
4	ZN	В	507	1/1	0.94	0.10	$65,\!65,\!65,\!65$	0
4	ZN	В	508	1/1	0.96	0.11	83,83,83,83	0
4	ZN	В	511	1/1	0.96	0.03	69,69,69,69	0
3	SO4	А	503	5/5	0.97	0.12	38,50,52,52	0
4	ZN	В	509	1/1	0.98	0.04	77,77,77,77	0
4	ZN	В	503	1/1	0.99	0.06	43,43,43,43	0
4	ZN	В	502	1/1	1.00	0.06	21,21,21,21	0
4	ZN	А	504	1/1	1.00	0.08	22,22,22,22	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.

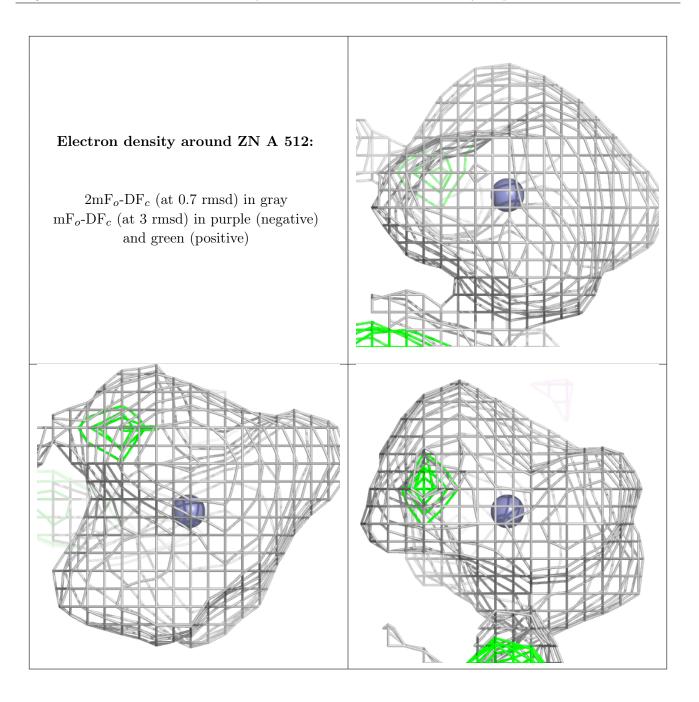




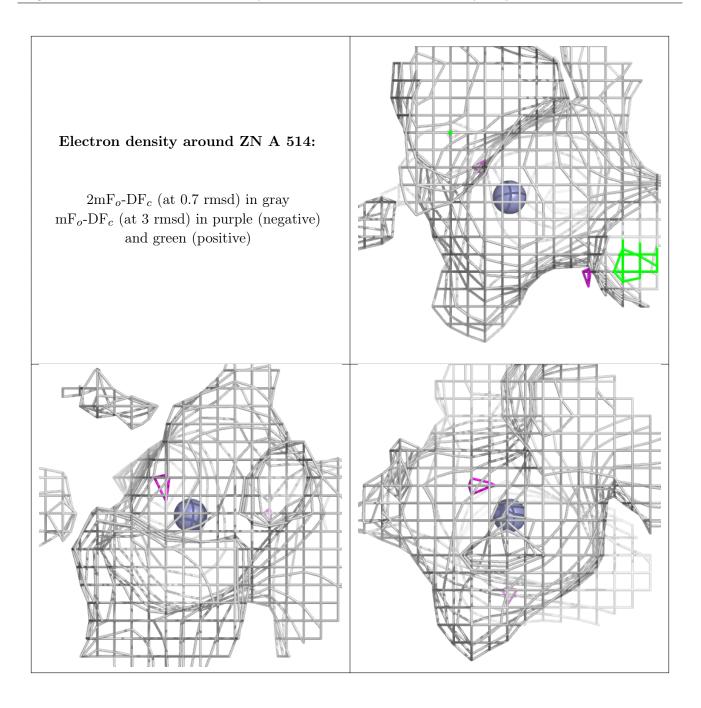




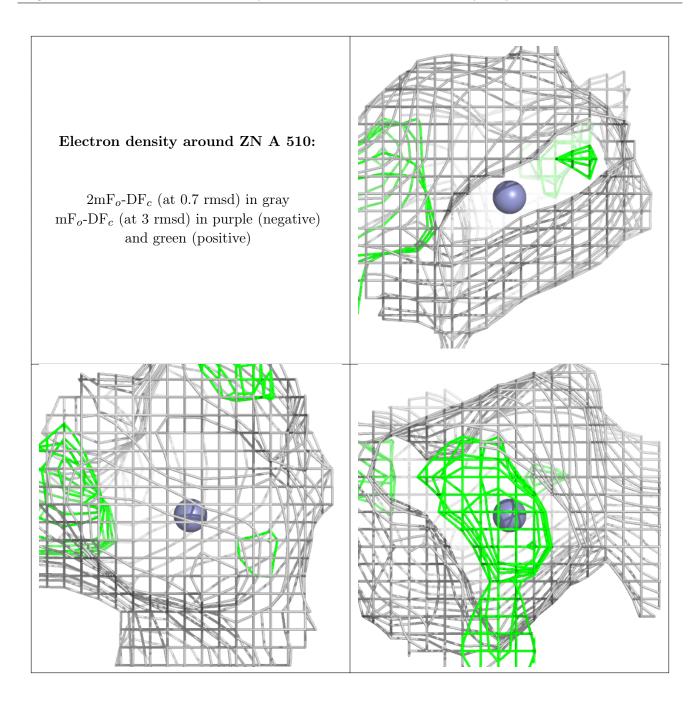




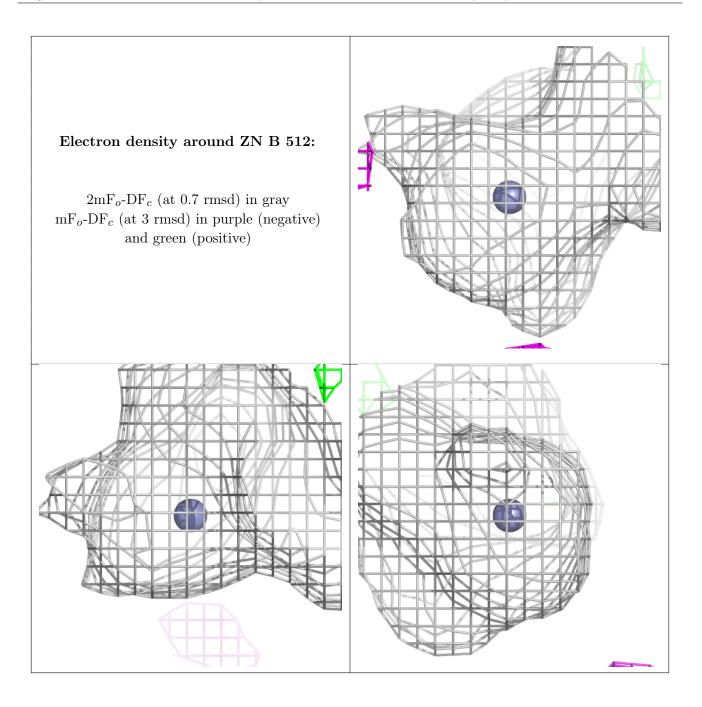




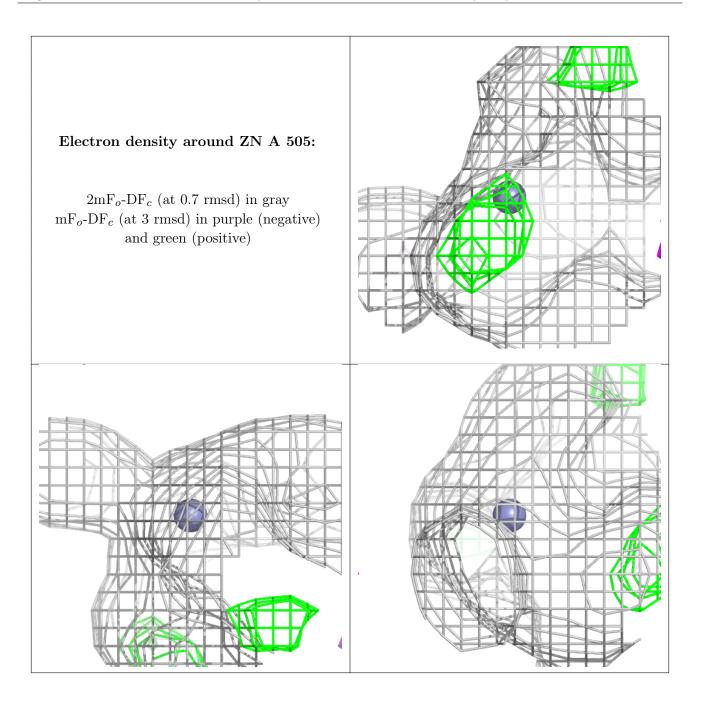




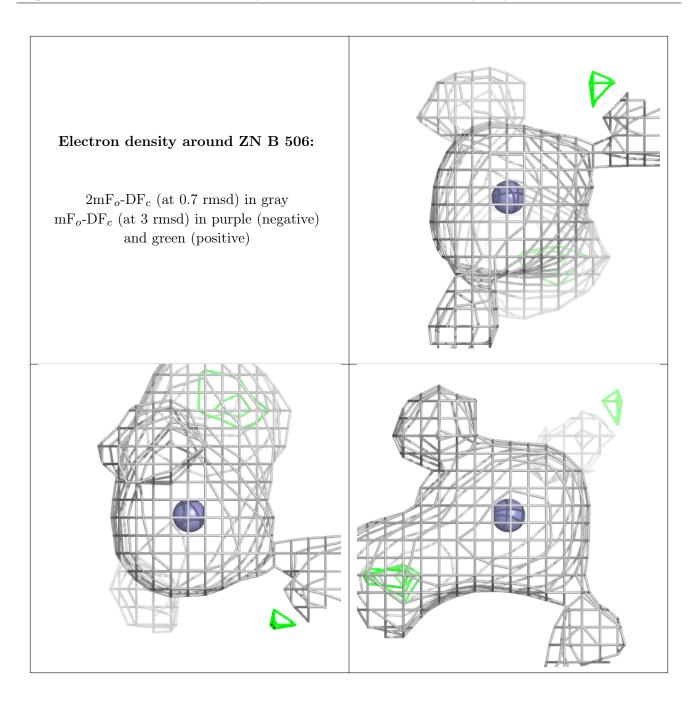




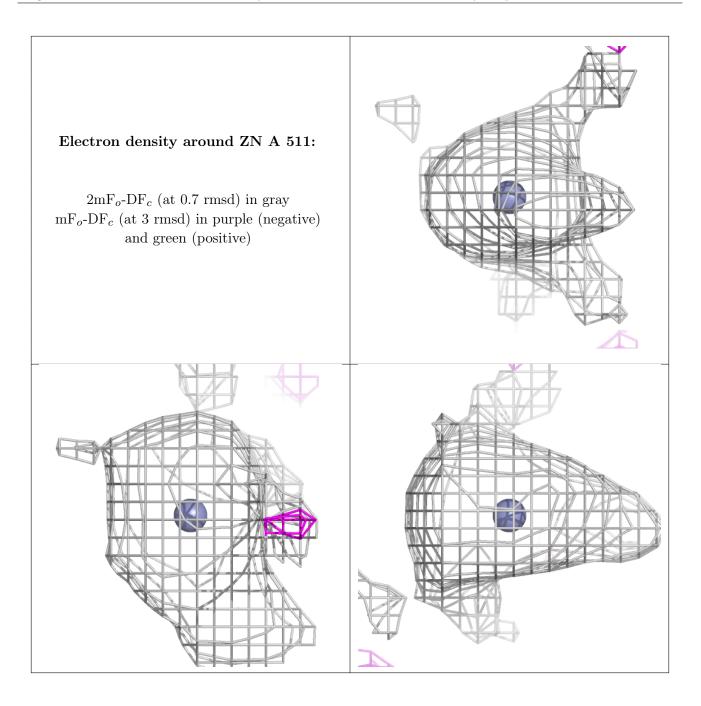




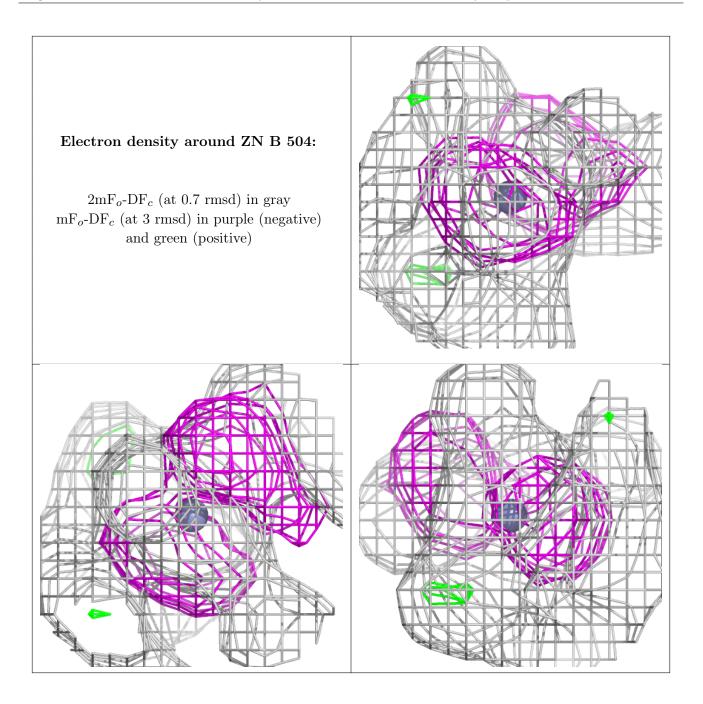




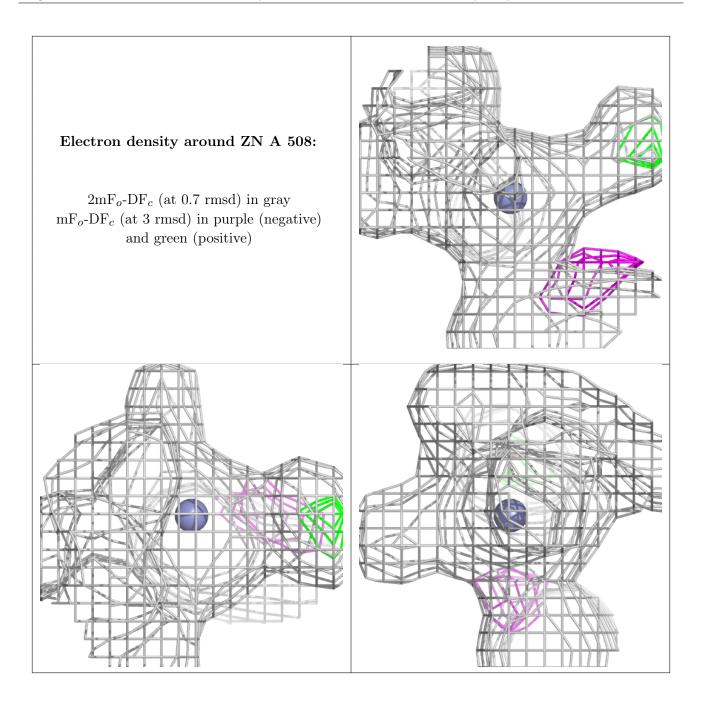




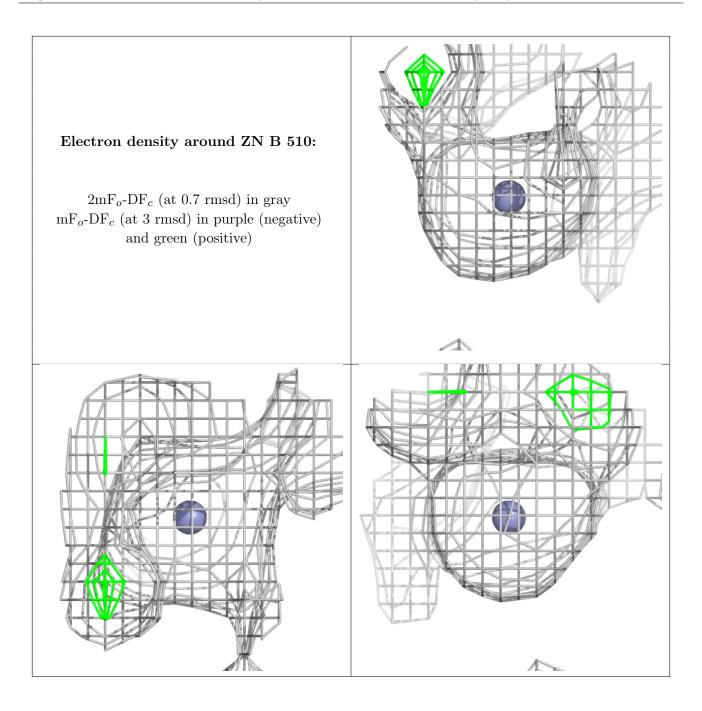




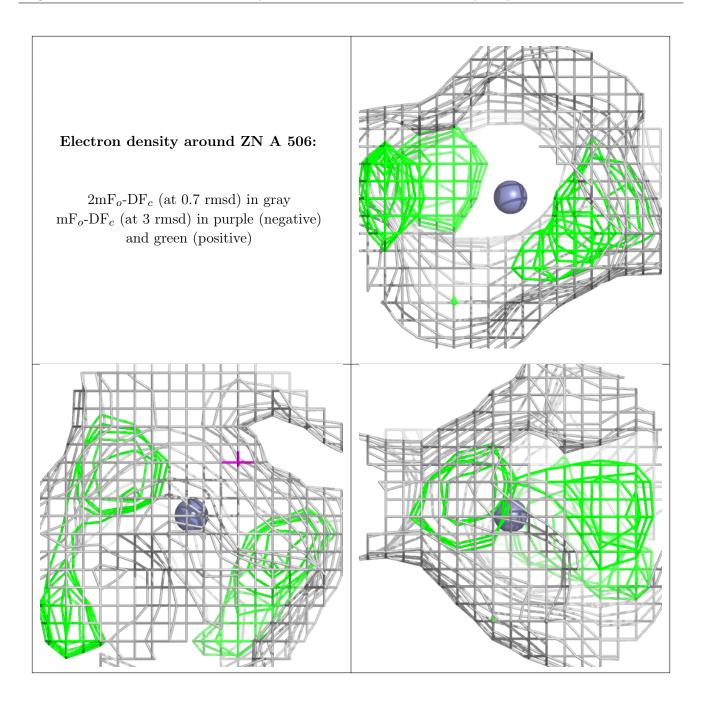




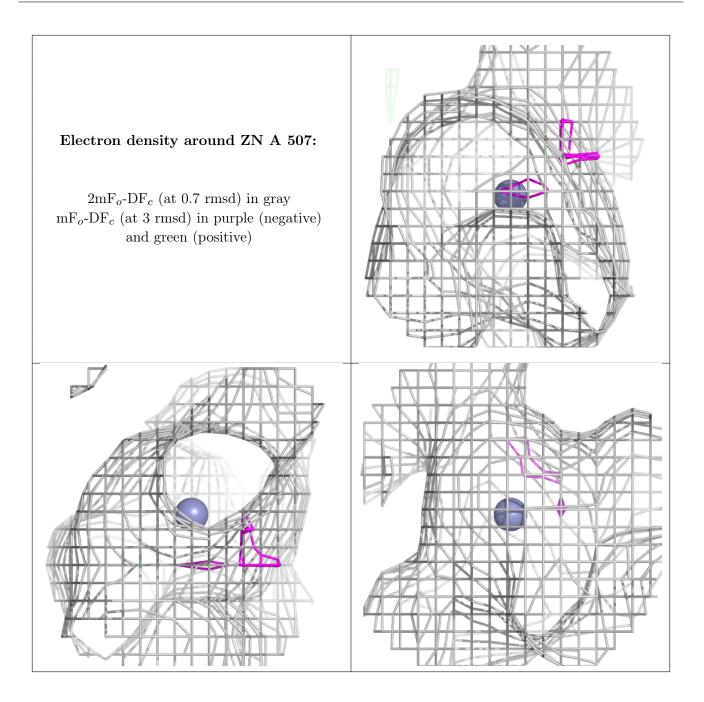




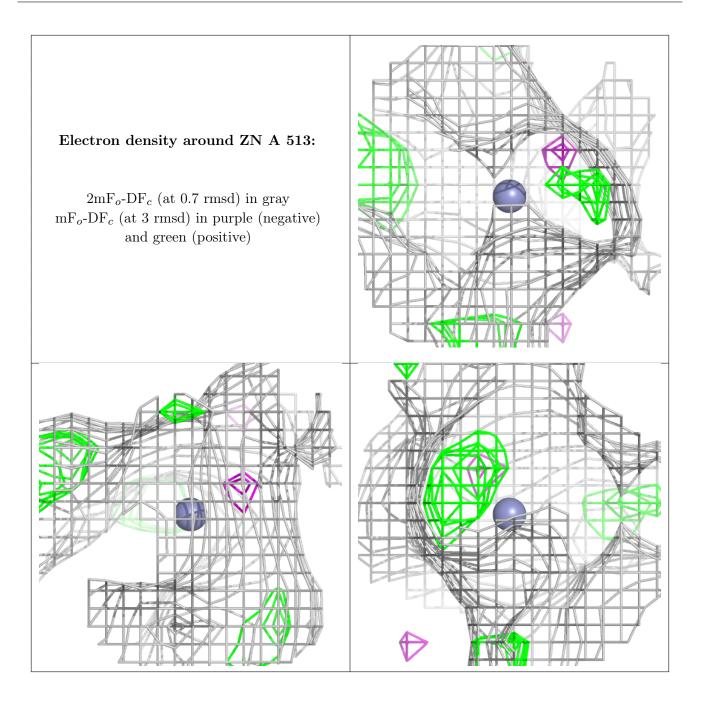




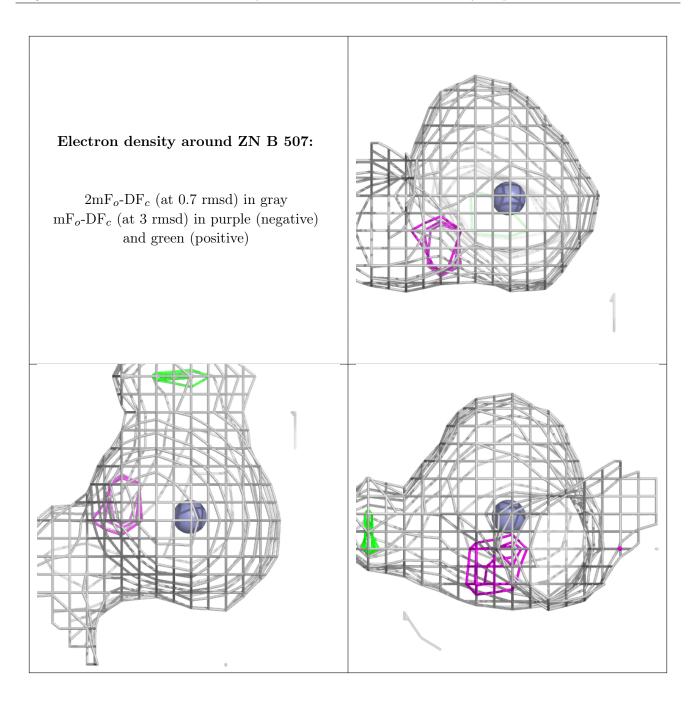




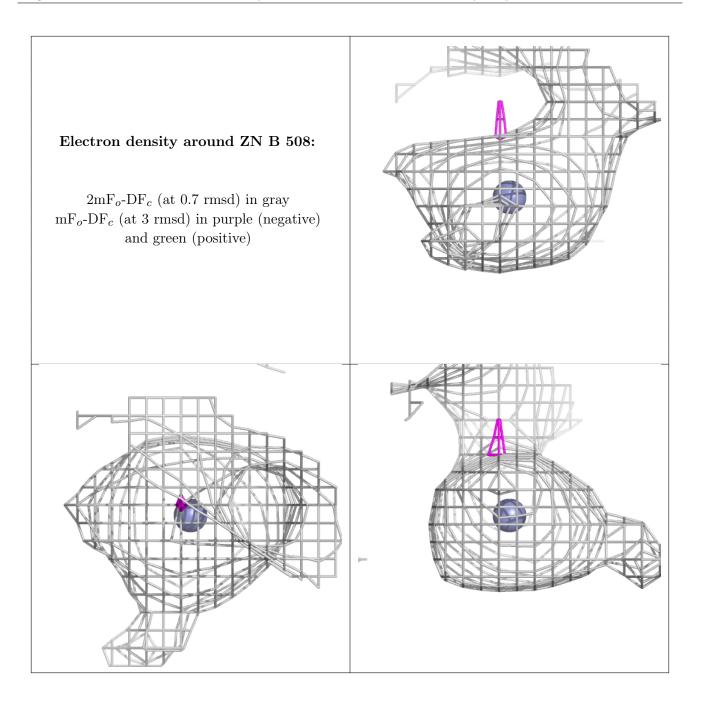




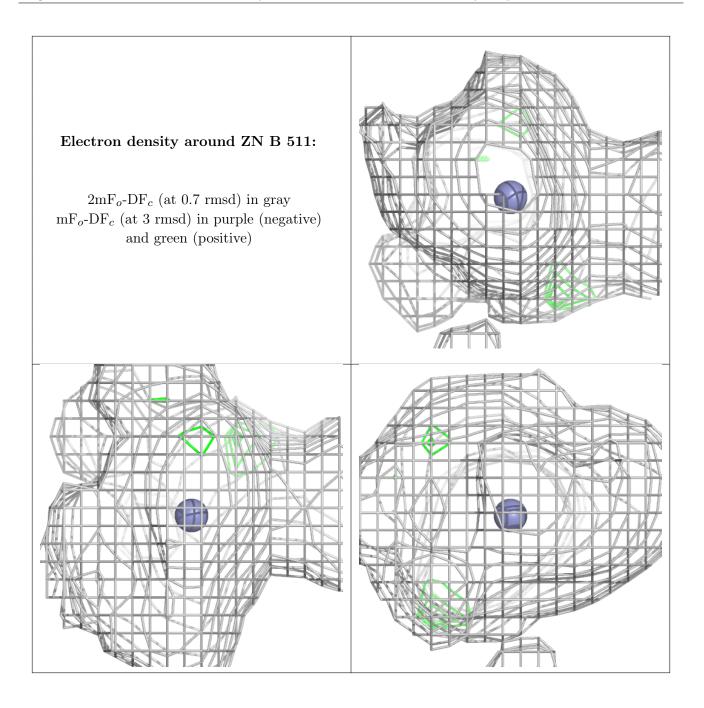




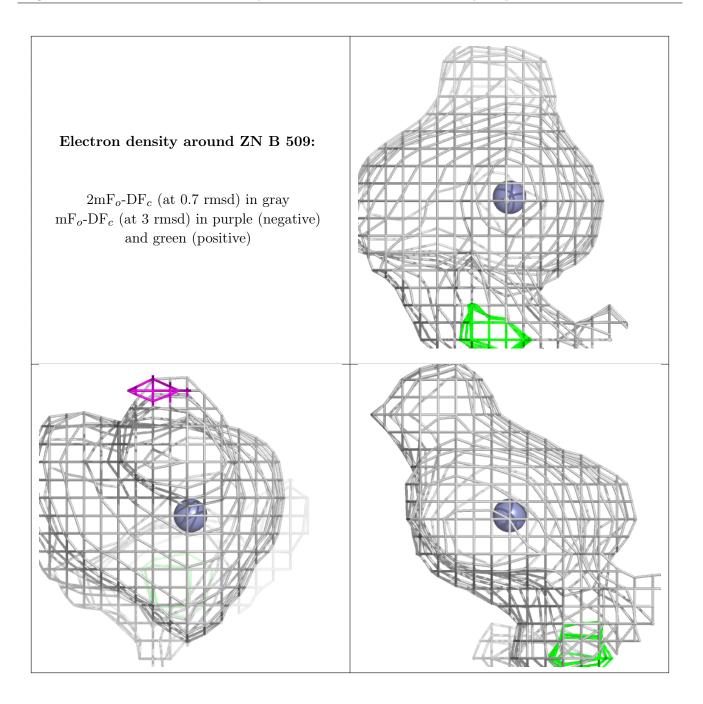




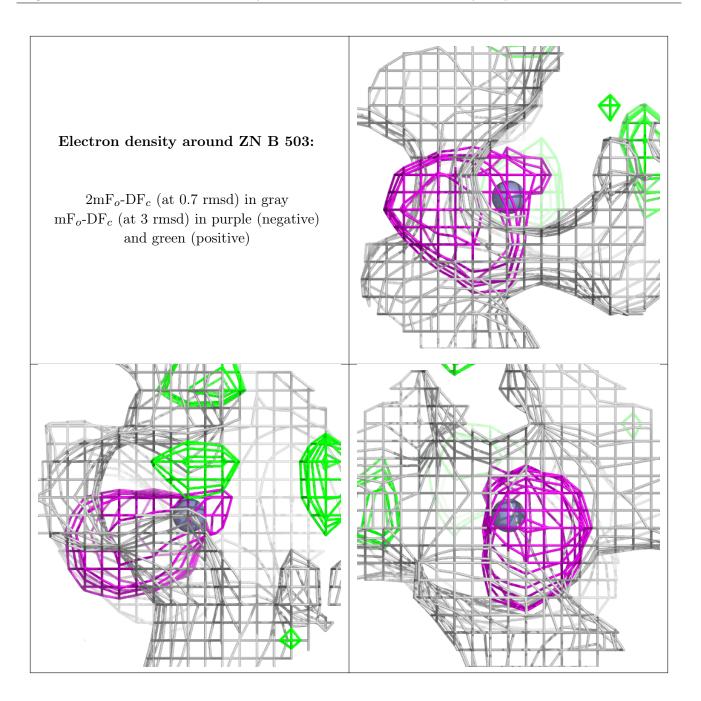




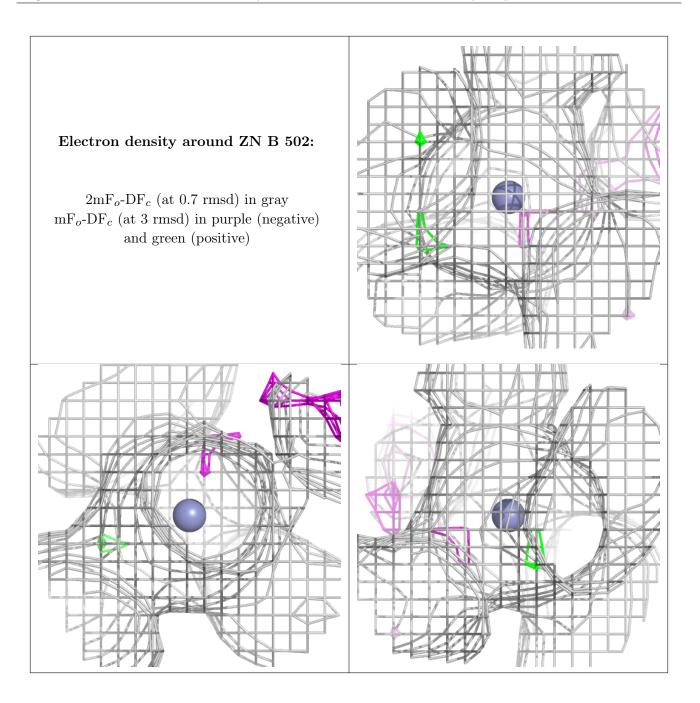




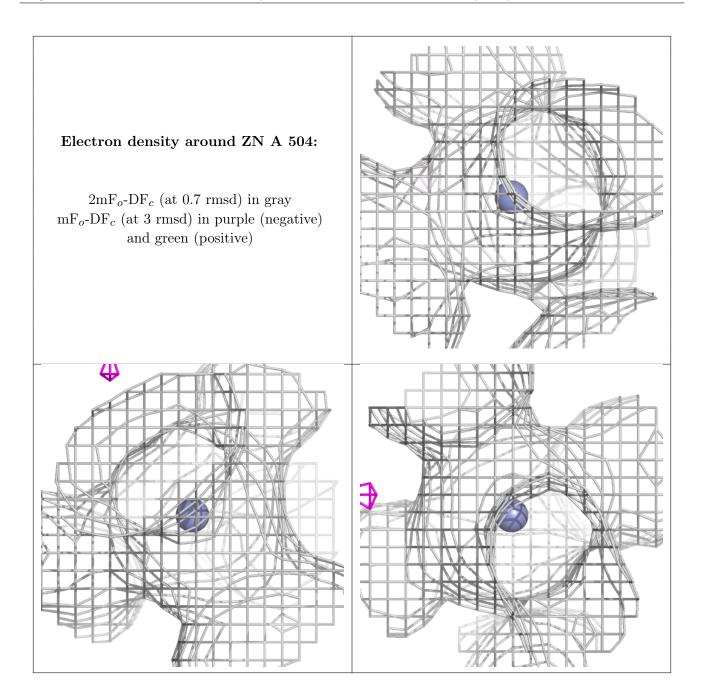












### 6.5 Other polymers (i)

There are no such residues in this entry.

