

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

#### Mar 3, 2024 – 01:04 AM EST

PDB ID	:	6ARG
Title	:	Aspergillus fumigatus Cytosolic Thiolase: Apo enzyme in complex with rubid-
		ium ions
Authors	:	Marshall, A.C.; Bond, C.S.; Bruning, J.B.
Deposited on		
Resolution	:	1.78  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:

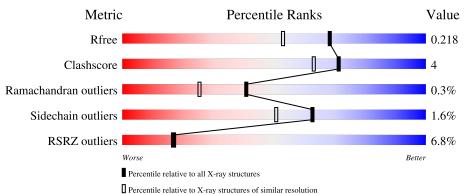
MolProbity Mogul Xtriage (Phenix) EDS	:	4.02b-467 1.8.5 (274361), CSD as541be (2020) 1.13 2.36
buster-report Percentile statistics Refmac	: : :	1.1.7 (2018) 20191225.v01 (using entries in the PDB archive December 25th 2019) 5.8.0158 7.0.044 (Gargrove)
Ideal geometry (DNA, RNA) Validation Pipeline (wwPDB-VP)		Parkinson et al. (1996) 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 1.78 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	9185 (1.80-1.76)
Clashscore	141614	10184 (1.80-1.76)
Ramachandran outliers	138981	10051 (1.80-1.76)
Sidechain outliers	138945	10050 (1.80-1.76)
RSRZ outliers	127900	9032 (1.80-1.76)

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	399	85%	13% ••
1	В	399	92%	7% •
1	С	399	<sup>2%</sup> 93%	6% •
1	D	399	3% 94%	5% •



#### 6ARG

# 2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 13089 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	Δ	205	Total	С	Ν	Ο	$\mathbf{S}$	0	6	0
	А	395	2860	1806	492	551	11	0	6	0
1	В	396	Total	С	Ν	0	S	0	2	0
		390	2851	1793	494	552	12			
1	С	396	Total	С	Ν	0	S	0	8	0
			2877	1817	493	555	12			0
1	1 D	205	Total	С	Ν	0	S	0	1	0
	395	2841	1790	488	551	12	0	4	0	

• Molecule 1 is a protein called Acetyl-CoA acetyltransferase.

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
А	0	GLY	-	expression tag	UNP Q4WCL5
В	0	GLY	-	expression tag	UNP Q4WCL5
С	0	GLY	-	expression tag	UNP Q4WCL5
D	0	GLY	-	expression tag	UNP Q4WCL5

• Molecule 2 is RUBIDIUM ION (three-letter code: RB) (formula: Rb) (labeled as "Ligand of Interest" by depositor).

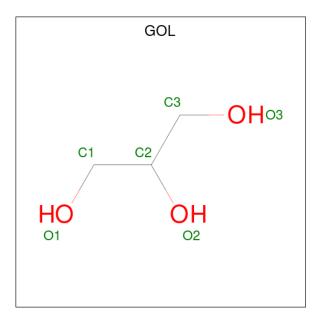
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	А	3	Total Rb 3 3	0	0
2	В	3	Total Rb 3 3	0	0
2	С	3	Total Rb 3 3	0	0
2	D	4	Total Rb 4 4	0	0

• Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	А	1	Total Cl 1 1	0	0
3	В	1	Total Cl 1 1	0	0
3	С	1	Total Cl 1 1	0	0
3	D	1	Total Cl 1 1	0	0

• Molecule 4 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	С	1	$\begin{array}{ccc} \text{Total} & \text{C} & \text{O} \\ 6 & 3 & 3 \end{array}$	0	0
4	С	1	$\begin{array}{ccc} \text{Total} & \text{C} & \text{O} \\ 6 & 3 & 3 \end{array}$	0	0
4	D	1	$\begin{array}{ccc} \text{Total} & \text{C} & \text{O} \\ 6 & 3 & 3 \end{array}$	0	0

• Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	А	215	Total O 215 215	0	0
5	В	249	Total         O           249         249	0	0
5	С	565	Total O 565 565	0	0

Continued on next page...



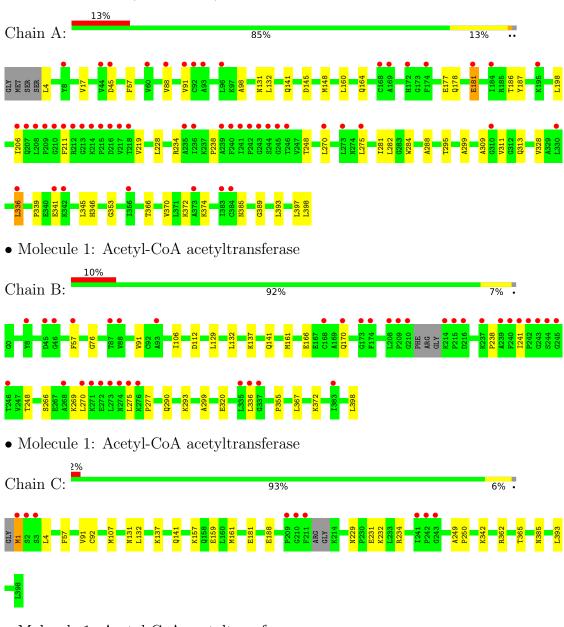
Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	D	596	Total O 596 596	0	0



# 3 Residue-property plots (i)

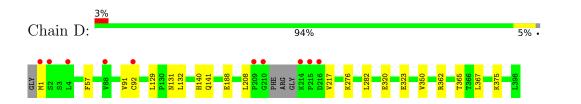
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: Acetyl-CoA acetyltransferase

 $\bullet$  Molecule 1: Acetyl-CoA acetyltransferase







## 4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants	71.26Å 105.37Å 110.26Å	Depositor
a, b, c, $\alpha$ , $\beta$ , $\gamma$	$90.00^{\circ}$ $108.69^{\circ}$ $90.00^{\circ}$	Depositor
Resolution (Å)	36.20 - 1.78	Depositor
Resolution (A)	36.16 - 1.78	EDS
% Data completeness	98.6 (36.20-1.78)	Depositor
(in resolution range)	98.6(36.16-1.78)	EDS
R <sub>merge</sub>	0.17	Depositor
R <sub>sym</sub>	(Not available)	Depositor
$< I/\sigma(I) > 1$	$1.43 (at 1.78 \text{\AA})$	Xtriage
Refinement program	PHENIX dev_2686	Depositor
B B.	0.182 , $0.220$	Depositor
$R, R_{free}$	0.181 , $0.218$	DCC
$R_{free}$ test set	7332 reflections $(5.06\%)$	wwPDB-VP
Wilson B-factor $(Å^2)$	17.0	Xtriage
Anisotropy	0.249	Xtriage
Bulk solvent $k_{sol}(e/Å^3)$ , $B_{sol}(Å^2)$	0.30 , $50.6$	EDS
L-test for twinning <sup>2</sup>	$<  L  > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.021 for h,-k,-h-l	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	13089	wwPDB-VP
Average B, all atoms $(Å^2)$	32.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 3.78% 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: GOL, CL, RB

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	0.29	0/2923	0.50	0/3966	
1	В	0.31	0/2900	0.51	0/3932	
1	С	0.43	0/2944	0.57	0/3995	
1	D	0.45	0/2896	0.57	0/3932	
All	All	0.38	0/11663	0.54	0/15825	

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	2860	0	2923	38	0
1	В	2851	0	2913	20	0
1	С	2877	0	2953	17	0
1	D	2841	0	2891	14	0
2	А	3	0	0	0	0
2	В	3	0	0	0	0
2	С	3	0	0	0	0
2	D	4	0	0	0	0
3	А	1	0	0	0	0

Continued on next page...



Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	В	1	0	0	0	0
3	С	1	0	0	0	0
3	D	1	0	0	0	0
4	С	12	0	16	1	0
4	D	6	0	8	0	0
5	А	215	0	0	6	0
5	В	249	0	0	6	0
5	С	565	0	0	7	2
5	D	596	0	0	2	1
All	All	13089	0	11704	82	2

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 4.

The worst 5 of 82 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:B:299:ALA:HA	1:B:336:LEU:HD21	1.64	0.80
1:C:1:MET:HA	1:C:4:LEU:HD12	1.65	0.79
1:B:137:LYS:NZ	5:B:503:HOH:O	2.11	0.78
1:A:4:LEU:N	5:A:501:HOH:O	2.18	0.75
1:C:181:GLU:OE2	1:C:234:ARG:NH2	2.22	0.72

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
5:C:921:HOH:O	5:D:837:HOH:O[1_455]	2.14	0.06
5:C:510:HOH:O	5:C:704:HOH:O[2_547]	2.16	0.04

### 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	399/399~(100%)	390~(98%)	8 (2%)	1 (0%)	41	25
1	В	394/399~(99%)	381~(97%)	12 (3%)	1 (0%)	41	25
1	С	400/399~(100%)	393~(98%)	6 (2%)	1 (0%)	41	25
1	D	395/399~(99%)	387~(98%)	7 (2%)	1 (0%)	41	25
All	All	1588/1596~(100%)	1551~(98%)	33~(2%)	4 (0%)	41	25

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	В	91	VAL
1	А	91	VAL
1	С	91	VAL
1	D	91	VAL

#### 5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent side chain 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 side chain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Rotameric Outliers		Percentiles		
1	А	295/298~(99%)	287~(97%)	8(3%)	44	28		
1	В	295/298~(99%)	293~(99%)	2(1%)	84	79		
1	С	299/298~(100%)	292~(98%)	7~(2%)	50	34		
1	D	293/298~(98%)	290 (99%)	3 (1%)	76	68		
All	All	1182/1192~(99%)	1162 (98%)	20 (2%)	62	48		

 $5~{\rm of}~20$  residues with a non-rotameric side chain are listed below:

Mol	Chain	Res	Type
1	С	231[B]	GLU
1	D	57	PHE
1	D	276	LYS
1	D	92	CYS
1	А	372	LYS

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



sidechains are listed below:

Mol	Chain	Res	Type
1	В	164	GLN

#### 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 20 ligands modelled in this entry, 17 are monoatomic - leaving 3 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	Res	Link	B	ond leng	gths	B	ond ang	gles
	Type	Unain	nes		Counts	RMSZ	# Z >2	Counts	RMSZ	# Z >2
4	GOL	D	406	-	$5,\!5,\!5$	1.03	0	$5,\!5,\!5$	0.97	0
4	GOL	С	406	-	$5,\!5,\!5$	0.94	0	$5,\!5,\!5$	1.51	1 (20%)
4	GOL	С	405	-	$5,\!5,\!5$	0.82	0	$5,\!5,\!5$	1.09	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
4	GOL	D	406	-	-	2/4/4/4	-

Continued on next page...



Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	GOL	С	406	-	-	2/4/4/4	-
4	GOL	С	405	-	-	0/4/4/4	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
4	С	406	GOL	C3-C2-C1	-2.64	101.45	111.70

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	D	406	GOL	C1-C2-C3-O3
4	С	406	GOL	O1-C1-C2-C3
4	D	406	GOL	O2-C2-C3-O3
4	С	406	GOL	O1-C1-C2-O2

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	С	406	GOL	1	0

### 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	$OWAB(Å^2)$	$\mathbf{Q}{<}0.9$
1	А	395/399~(98%)	0.84	51 (12%) 3 3	23, 47, 73, 130	0
1	В	396/399~(99%)	0.50	38 (9%) 8 7	20, 39, 67, 98	0
1	С	396/399~(99%)	-0.16	9 (2%) 60 60	7, 17, 38, 94	0
1	D	395/399~(98%)	-0.18	10 (2%) 57 56	9, 16, 34, 99	0
All	All	1582/1596~(99%)	0.25	108 (6%) 17 17	7, 27, 67, 130	0

The worst 5 of 108 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	А	211	PHE	10.2
1	А	213	GLY	7.4
1	В	215	PRO	6.8
1	В	214	LYS	6.8
1	В	210	GLY	6.1

## 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)

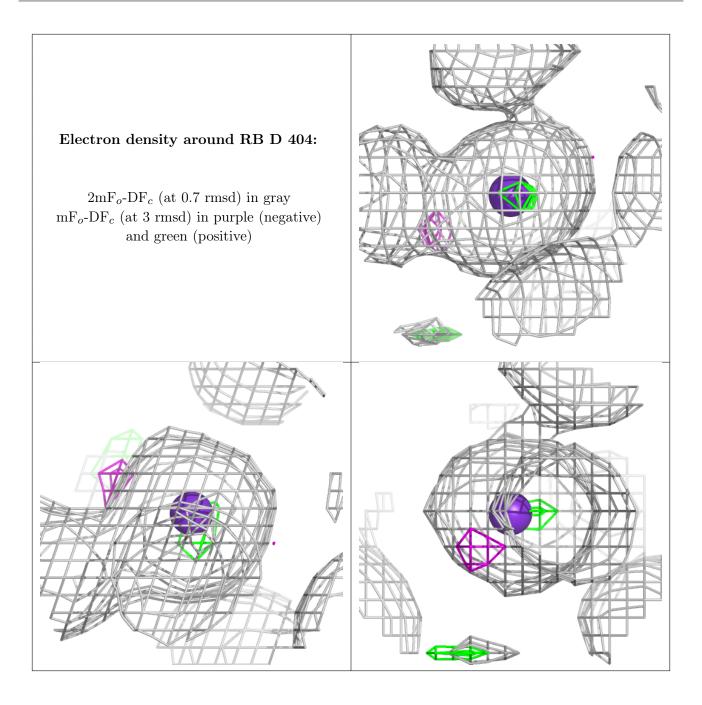
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median,  $95^{th}$  percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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



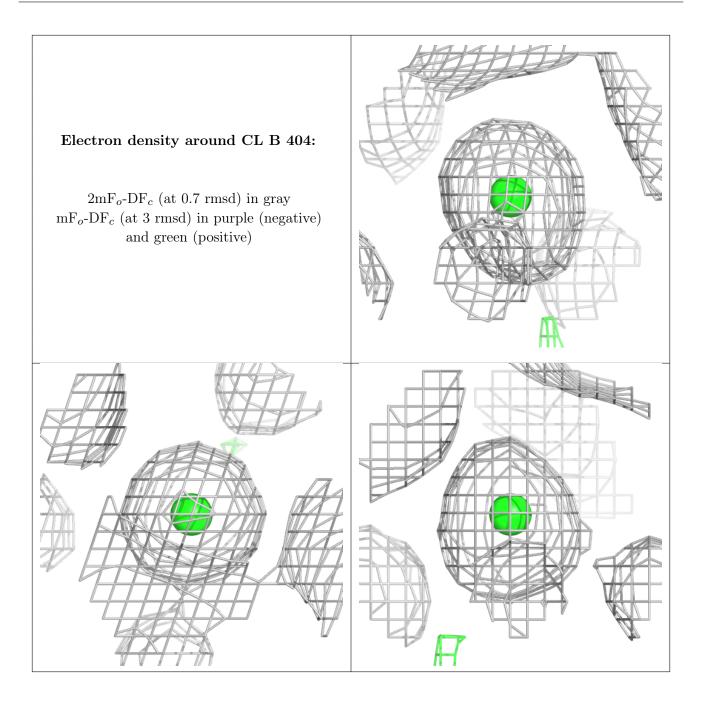
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
4	GOL	С	406	6/6	0.75	0.26	20,20,20,20	0
4	GOL	С	405	6/6	0.88	0.18	20,20,20,20	0
4	GOL	D	406	6/6	0.96	0.10	24,33,45,51	0
2	RB	D	404	1/1	0.97	0.07	48,48,48,48	1
3	CL	В	404	1/1	0.98	0.10	30,30,30,30	1
2	RB	В	403	1/1	0.98	0.09	88,88,88,88	0
2	RB	А	402	1/1	0.98	0.04	66,66,66,66	1
3	CL	А	404	1/1	0.98	0.09	30,30,30,30	1
2	RB	В	402	1/1	0.99	0.02	51,51,51,51	0
3	CL	С	404	1/1	0.99	0.07	$25,\!25,\!25,\!25$	0
3	CL	D	405	1/1	0.99	0.09	21,21,21,21	0
2	RB	D	403	1/1	1.00	0.04	23,23,23,23	1
2	RB	В	401	1/1	1.00	0.04	40,40,40,40	0
2	RB	А	401	1/1	1.00	0.04	40,40,40,40	1
2	RB	А	403	1/1	1.00	0.02	$65,\!65,\!65,\!65$	1
2	RB	С	401	1/1	1.00	0.06	18,18,18,18	0
2	RB	С	402	1/1	1.00	0.04	19,19,19,19	1
2	RB	С	403	1/1	1.00	0.04	20,20,20,20	1
2	RB	D	401	1/1	1.00	0.05	16,16,16,16	0
2	RB	D	402	1/1	1.00	0.05	$19,\!19,\!19,\!19$	1

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.

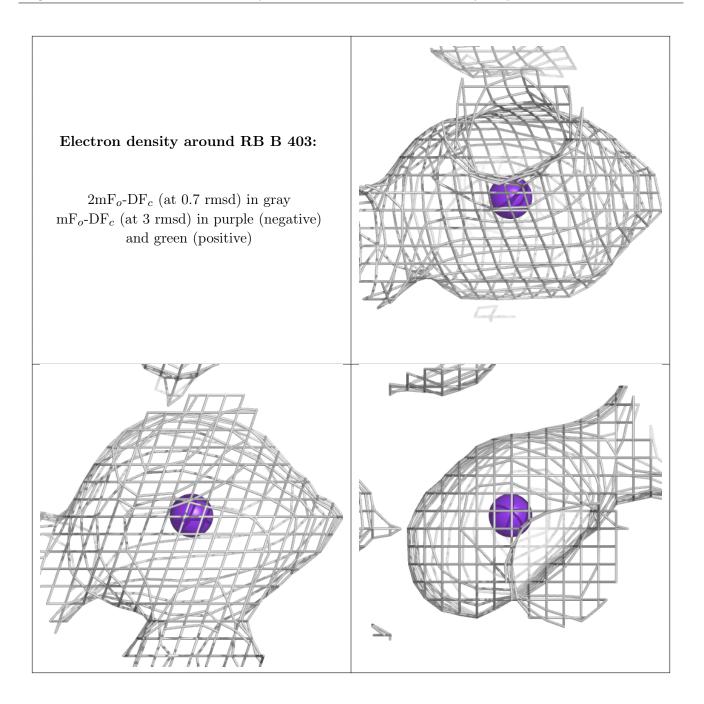




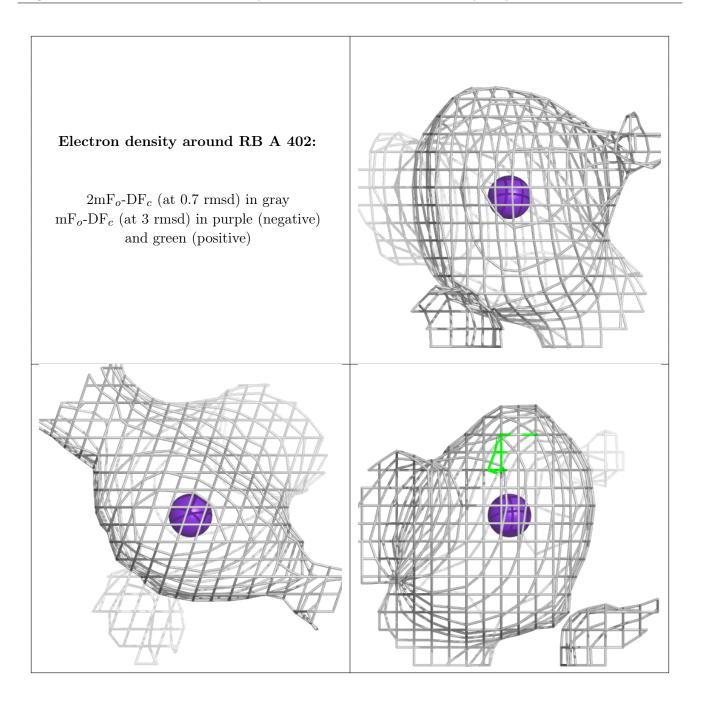




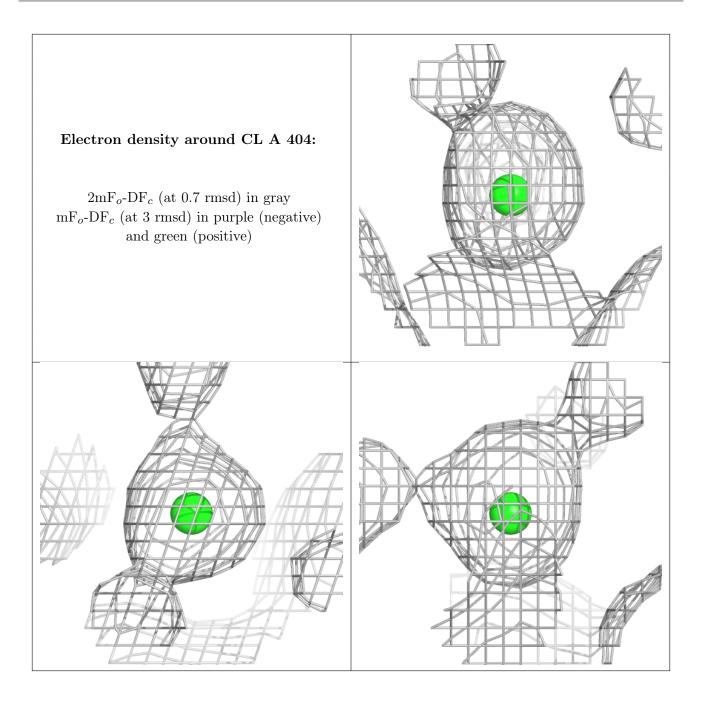




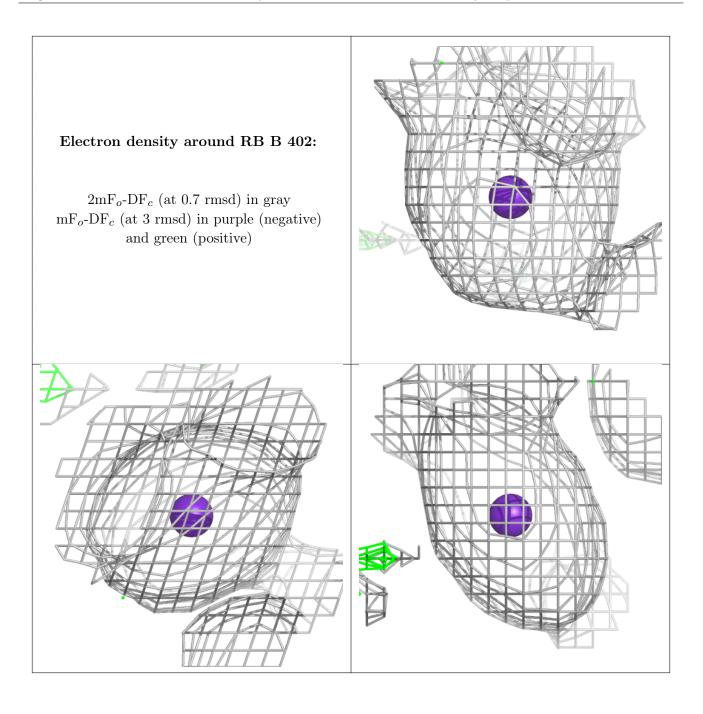




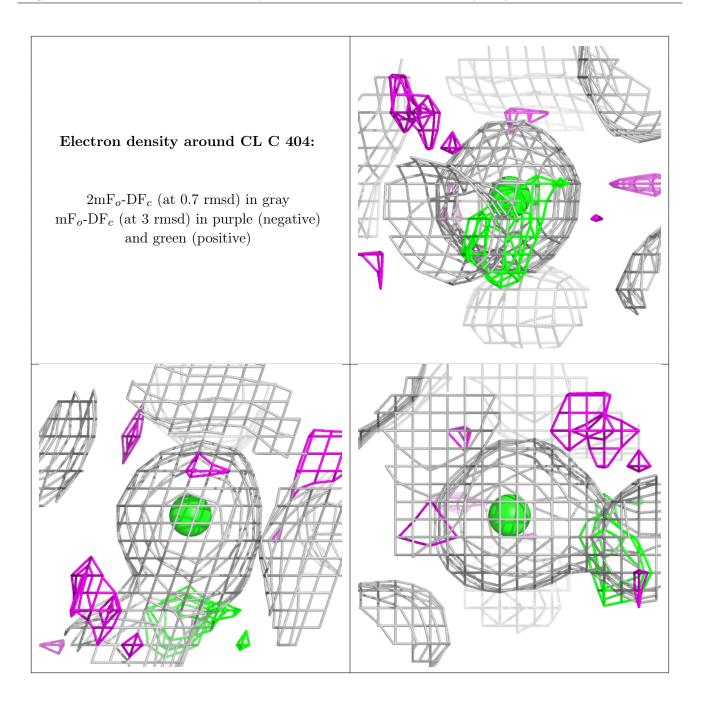




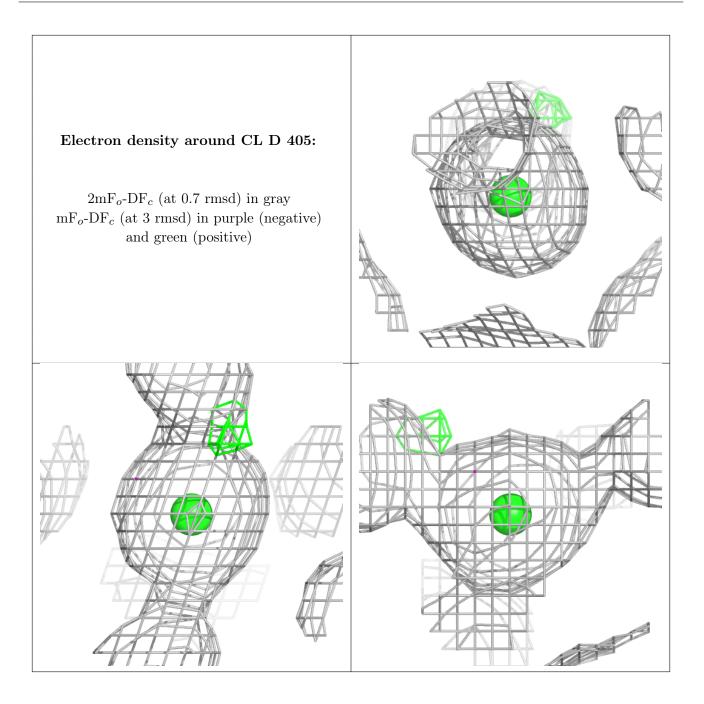




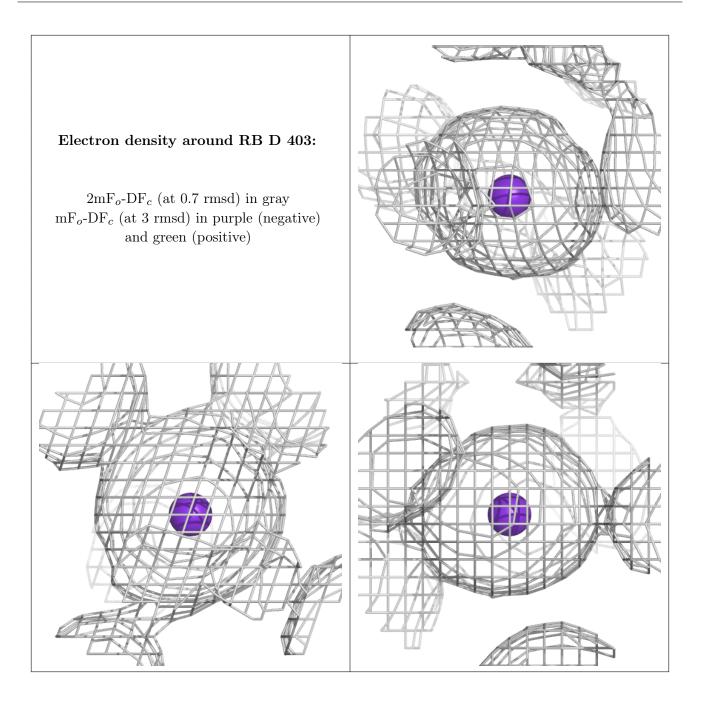




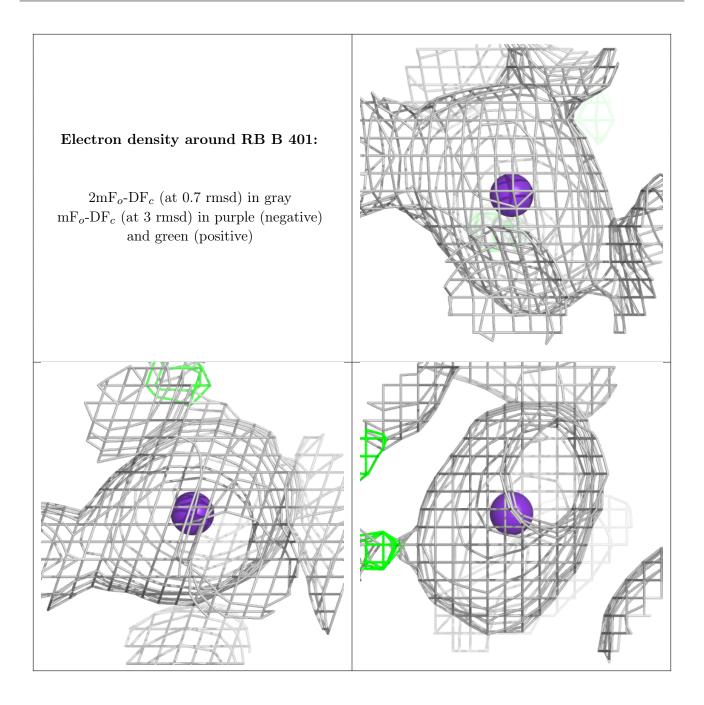




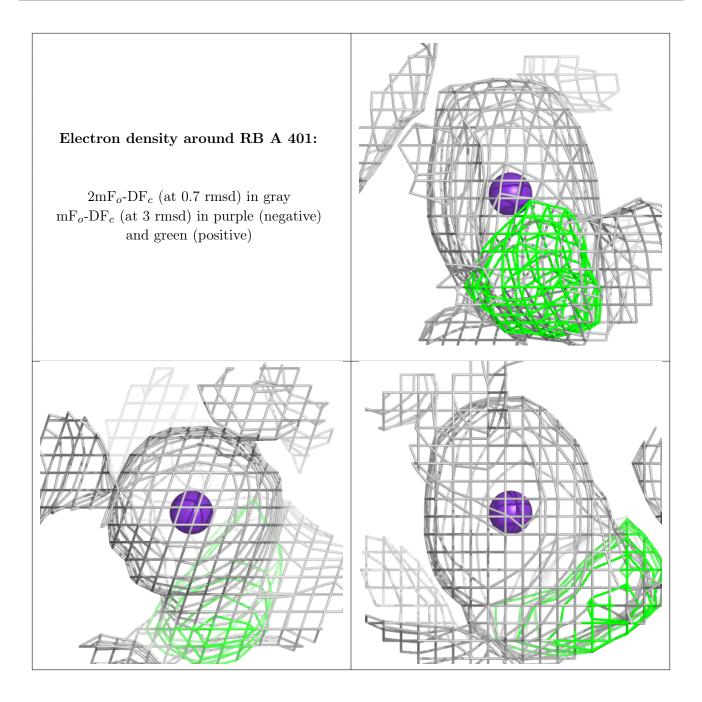




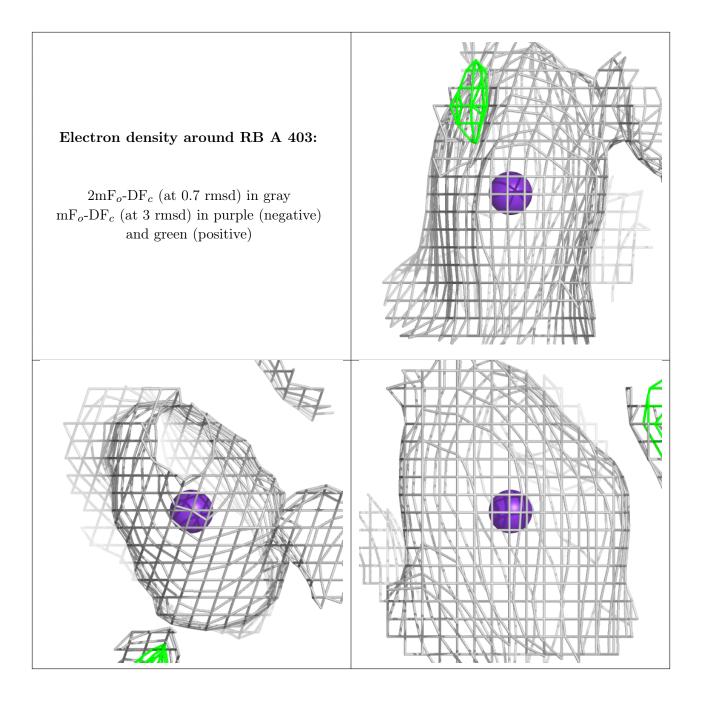




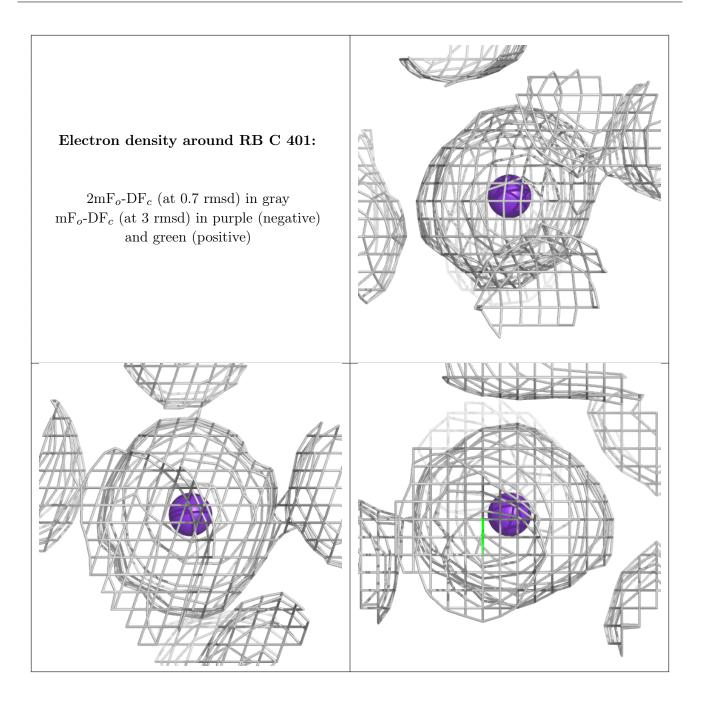




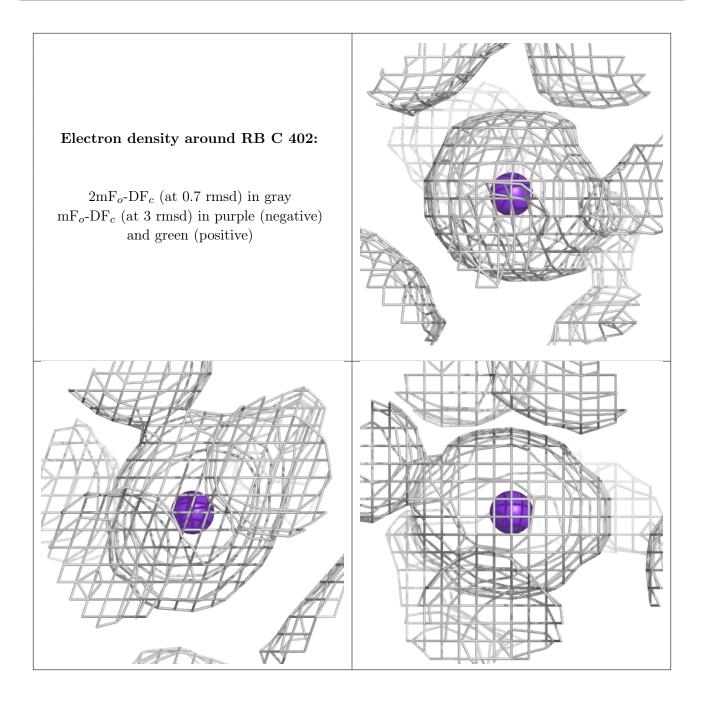




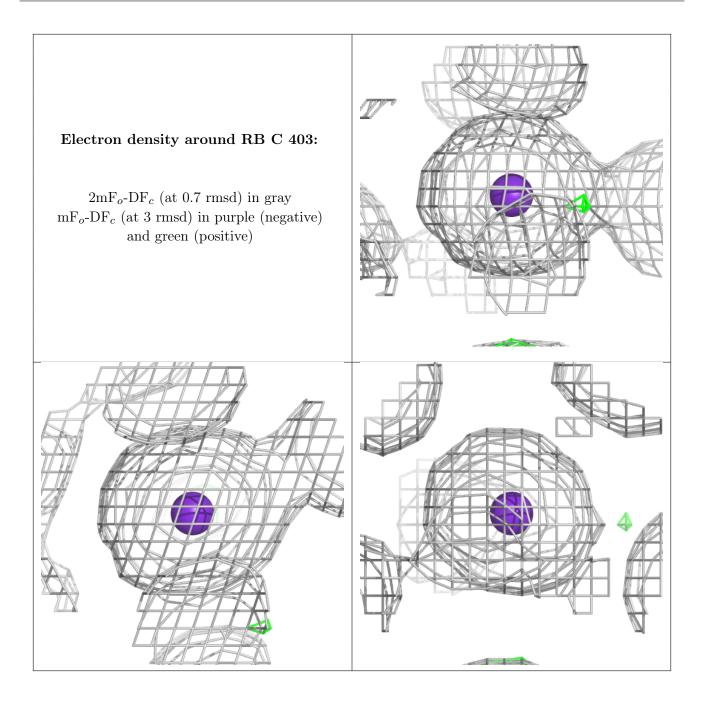




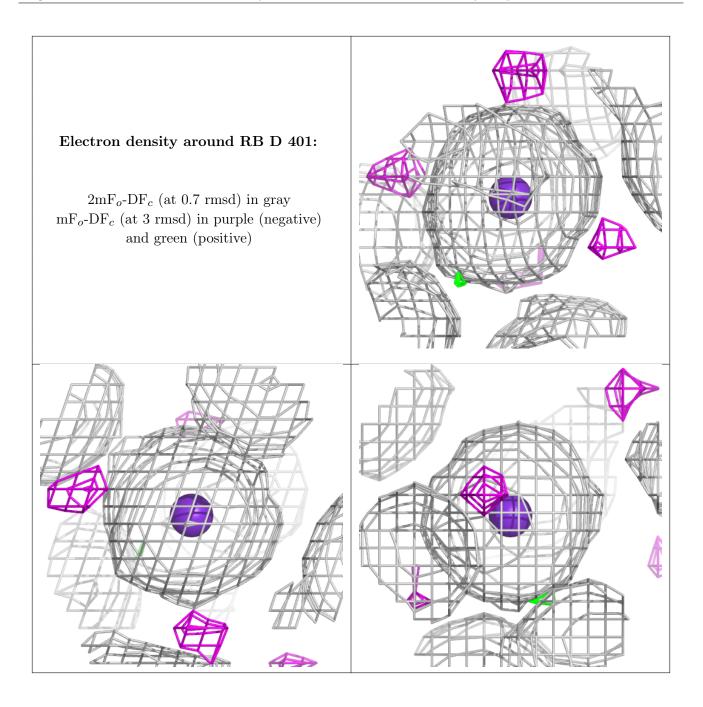




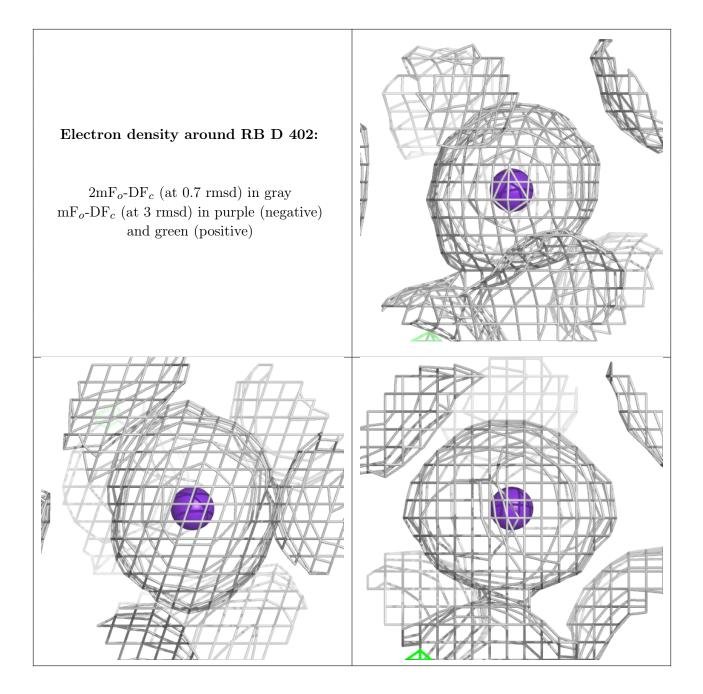












## 6.5 Other polymers (i)

There are no such residues in this entry.

