Full wwPDB X-ray Structure Validation Report (i)

May 16, 2020 – 09:15 pm BST

PDB ID	:	6OU1
Title	:	Crystal Structure of the Computationally-derived 21-Variant of the Myocilin Olfactomedin Domain
Authors	:	Hill, S.E.; Kwon, M.S.; Lieberman, R.L.
Deposited on	:	2019-05-03
Resolution	:	1.88 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 following versions of software and data (see references (i)) were used in the production of this report:

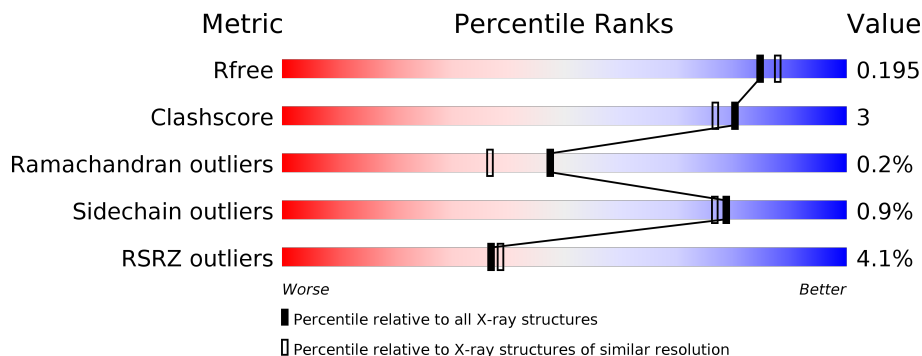|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.88 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	Percentile Ranks	Value
Rfree		0.195
Clashscore		3
Ramachandran outliers		0.2%
Sidechain outliers		0.9%
RSRZ outliers		4.1%

Worse　　　　　　　　　　　　　　　　　　　　　　　　Better

■ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	9470 (1.90-1.86)
Clashscore	141614	10282 (1.90-1.86)
Ramachandran outliers	138981	10152 (1.90-1.86)
Sidechain outliers	138945	10152 (1.90-1.86)
RSRZ outliers	127900	9303 (1.90-1.86)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for >=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	A	277	2% · 88% · 5% · 6%
1	B	277	6% · 86% · 6% · 8%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 8425 atoms, of which 3988 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 Myocilin.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	260	Total	C	H	N	O	S	0	0	0
			4100	1338	2014	340	402	6			
1	B	255	Total	C	H	N	O	S	0	0	0
			4004	1313	1958	334	393	6			

There are 42 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	255	VAL	LEU	conflict	UNP Q99972
A	277	VAL	THR	conflict	UNP Q99972
A	331	ASN	SER	conflict	UNP Q99972
A	340	GLY	GLU	conflict	UNP Q99972
A	348	ASP	GLU	conflict	UNP Q99972
A	350	LYS	ASN	conflict	UNP Q99972
A	369	TYR	PHE	conflict	UNP Q99972
A	386	ASN	ALA	conflict	UNP Q99972
A	395	ASN	ASP	conflict	UNP Q99972
A	396	GLN	GLU	conflict	UNP Q99972
A	400	ASN	ALA	conflict	UNP Q99972
A	410	THR	ASN	conflict	UNP Q99972
A	413	VAL	LEU	conflict	UNP Q99972
A	456	ASN	GLY	conflict	UNP Q99972
A	459	THR	ILE	conflict	UNP Q99972
A	462	TYR	THR	conflict	UNP Q99972
A	464	ASP	THR	conflict	UNP Q99972
A	474	ASN	SER	conflict	UNP Q99972
A	484	ARG	LYS	conflict	UNP Q99972
A	492	GLY	LEU	conflict	UNP Q99972
A	493	HIS	ASN	conflict	UNP Q99972
B	255	VAL	LEU	conflict	UNP Q99972
B	277	VAL	THR	conflict	UNP Q99972
B	331	ASN	SER	conflict	UNP Q99972
B	340	GLY	GLU	conflict	UNP Q99972

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	348	ASP	GLU	conflict	UNP Q99972
B	350	LYS	ASN	conflict	UNP Q99972
B	369	TYR	PHE	conflict	UNP Q99972
B	386	ASN	ALA	conflict	UNP Q99972
B	395	ASN	ASP	conflict	UNP Q99972
B	396	GLN	GLU	conflict	UNP Q99972
B	400	ASN	ALA	conflict	UNP Q99972
B	410	THR	ASN	conflict	UNP Q99972
B	413	VAL	LEU	conflict	UNP Q99972
B	456	ASN	GLY	conflict	UNP Q99972
B	459	THR	ILE	conflict	UNP Q99972
B	462	TYR	THR	conflict	UNP Q99972
B	464	ASP	THR	conflict	UNP Q99972
B	474	ASN	SER	conflict	UNP Q99972
B	484	ARG	LYS	conflict	UNP Q99972
B	492	GLY	LEU	conflict	UNP Q99972
B	493	HIS	ASN	conflict	UNP Q99972

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca) (labeled as "Ligand of Interest" by author).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total 1	Ca 1	0	0
2	A	1	Total 1	Ca 1	0	0

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total 1	Na 1	0	0
3	A	1	Total 1	Na 1	0	0

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

GOL

C3

OHO3

C1   C2

HO

O1

OH

O2

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	H	O	0	0
			14	3	8	3		
4	B	1	Total	C	H	O	0	0
			14	3	8	3		

• Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	165	Total	O	0	0
			165	165		
5	B	124	Total	O	0	0
			124	124		

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1: Myocilin

Chain A:

- Molecule 1: Myocilin

Chain B:
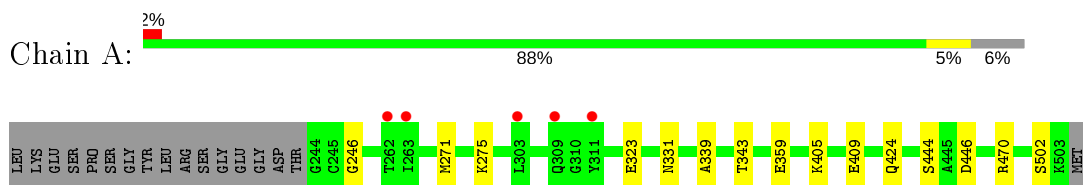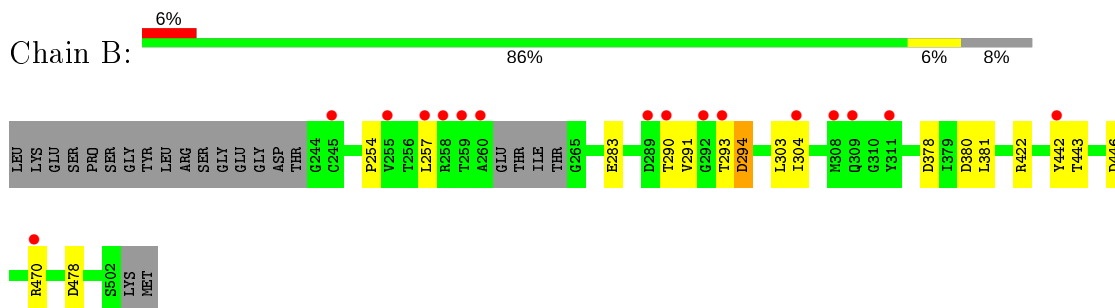
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	43.36Å 66.89Å 81.00Å 90.00° 95.19° 90.00°	Depositor
Resolution (Å)	36.71 – 1.88 36.71 – 1.88	Depositor EDS
% Data completeness (in resolution range)	91.0 (36.71-1.88) 90.7 (36.71-1.88)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.10 (at 1.88Å)	Xtriage
Refinement program	PHENIX (1.13_2998: ???)	Depositor
R, R_{free}	0.168 , 0.195 0.168 , 0.195	Depositor DCC
R_{free} test set	2000 reflections (5.55%)	wwPDB-VP
Wilson B-factor (Å^2)	21.2	Xtriage
Anisotropy	0.824	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.45 , 57.9	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	8425	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 6.94% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, CA, NA

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#$|Z|$>5	RMSZ	#$|Z|$>5
1	A	0.33	0/2142	0.51	0/2921
1	B	0.31	0/2101	0.48	0/2864
All	All	0.32	0/4243	0.50	0/5785

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2086	2014	2017	10	0
1	B	2046	1958	1972	12	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	6	8	8	0	0
4	B	6	8	8	0	0
5	A	165	0	0	6	1
5	B	124	0	0	3	2
All	All	4437	3988	4005	22	2

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.

All (22) 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:470:ARG:NH1	5:A:702:HOH:O	2.04	0.88
1:B:380:ASP:OD1	5:B:701:HOH:O	1.95	0.84
1:A:409:GLU:OE1	5:A:701:HOH:O	1.96	0.83
1:A:343:THR:OG1	1:A:359:GLU:OE2	1.96	0.83
1:B:291:VAL:HG23	1:B:293:THR:HG23	1.70	0.74
1:B:478:ASP:OD1	5:B:701:HOH:O	2.14	0.64
1:A:405:LYS:HE2	5:A:755:HOH:O	2.02	0.60
1:B:291:VAL:HG23	1:B:293:THR:CG2	2.32	0.59
1:B:381:LEU:O	5:B:701:HOH:O	2.17	0.59
1:A:444:SER:OG	1:A:446:ASP:OD1	2.20	0.59
1:B:290:THR:O	1:B:291:VAL:HG13	2.06	0.55
1:A:424:GLN:OE1	5:A:704:HOH:O	2.18	0.55
1:B:442:TYR:O	1:B:443:THR:HB	2.11	0.50
1:A:275:LYS:NZ	5:A:703:HOH:O	2.14	0.48
1:A:331:ASN:ND2	5:A:705:HOH:O	2.23	0.47
1:A:246:GLY:HA3	1:A:502:SER:O	2.15	0.47
1:B:294:ASP:OD1	1:B:294:ASP:N	2.44	0.45
1:B:283:GLU:HG3	1:B:304:ILE:HG13	2.01	0.43
1:B:254:PRO:O	1:B:470:ARG:NH2	2.51	0.43
1:B:443:THR:HB	1:B:446:ASP:O	2.18	0.42
1:B:283:GLU:HG2	1:B:303:LEU:HA	2.03	0.41
1:A:323:GLU:HB2	1:A:339:ALA:HA	2.02	0.40

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:B:813:HOH:O	5:B:814:HOH:O[2_444]	1.96	0.24
5:A:840:HOH:O	5:B:717:HOH:O[2_454]	1.98	0.22

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	A	258/277 (93%)	250 (97%)	8 (3%)	0	100	100
1	B	251/277 (91%)	241 (96%)	9 (4%)	1 (0%)	34	22
All	All	509/554 (92%)	491 (96%)	17 (3%)	1 (0%)	47	37

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	294	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.

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	226/240 (94%)	225 (100%)	1 (0%)	91	90
1	B	221/240 (92%)	218 (99%)	3 (1%)	67	62
All	All	447/480 (93%)	443 (99%)	4 (1%)	78	76

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	271	MET
1	B	257	LEU
1	B	378	ASP
1	B	422	ARG

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

Mol	Chain	Res	Type
1	A	424	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 carbohydrates in this entry.

5.6   Ligand geometry (i)

Of 6 ligands modelled in this entry, 4 are monoatomic - leaving 2 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#$|Z| > 2$	Counts	RMSZ	#$|Z| > 2$
4	GOL	B	603	-	5,5,5	0.81	0	5,5,5	1.00	0
4	GOL	A	603	-	5,5,5	0.77	0	5,5,5	1.17	1 (20%)

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	B	603	-	-	0/4/4/4	-
4	GOL	A	603	-	-	2/4/4/4	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^\circ$)	Ideal($^\circ$)
4	A	603	GOL	C3-C2-C1	-2.12	103.45	111.70

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	603	GOL	C1-C2-C3-O3
4	A	603	GOL	O2-C2-C3-O3

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, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	260/277 (93%)	0.00	5 (1%)	66	68	15, 23, 45, 72	0
1	B	255/277 (92%)	0.38	16 (6%)	20	21	19, 28, 54, 90	0
All	All	515/554 (92%)	0.19	21 (4%)	37	39	15, 25, 49, 90	0

All (21) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	292	GLY	5.1
1	A	311	TYR	4.5
1	B	311	TYR	4.5
1	B	260	ALA	4.0
1	A	263	ILE	3.8
1	B	289	ASP	3.7
1	B	259	THR	3.6
1	B	257	LEU	3.2
1	B	304	ILE	3.1
1	B	258	ARG	3.0
1	B	308	MET	3.0
1	B	309	GLN	2.9
1	B	293	THR	2.8
1	B	290	THR	2.5
1	B	255	VAL	2.4
1	A	262	THR	2.4
1	A	309	GLN	2.3
1	B	442	TYR	2.2
1	A	303	LEU	2.2
1	B	470	ARG	2.2
1	B	245	CYS	2.2

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 carbohydrates 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, 95th 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	CA	B	601	1/1	0.85	0.11	49,49,49,49	0
3	NA	B	602	1/1	0.89	0.17	34,34,34,34	0
4	GOL	B	603	6/6	0.89	0.19	30,36,43,43	0
4	GOL	A	603	6/6	0.93	0.11	22,36,46,46	0
3	NA	A	602	1/1	0.96	0.07	16,16,16,16	0
2	CA	A	601	1/1	0.96	0.07	7,7,7,7	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.
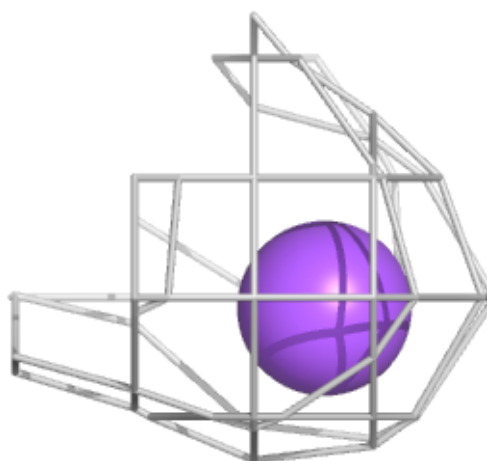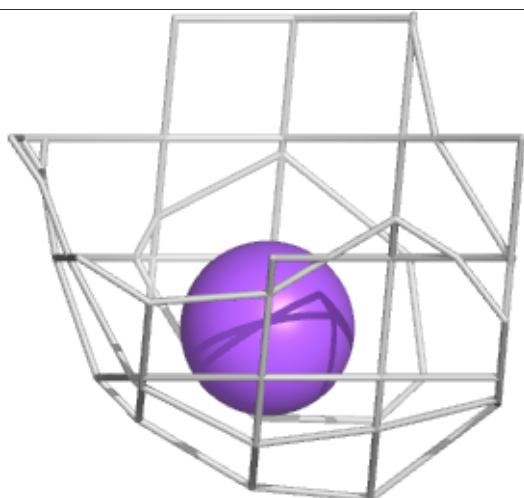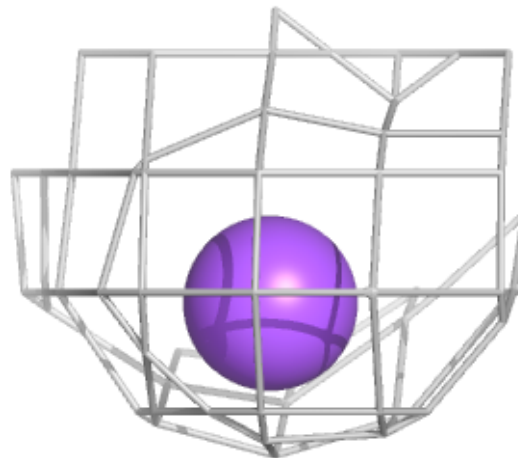
Electron density around CA B 601:

2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
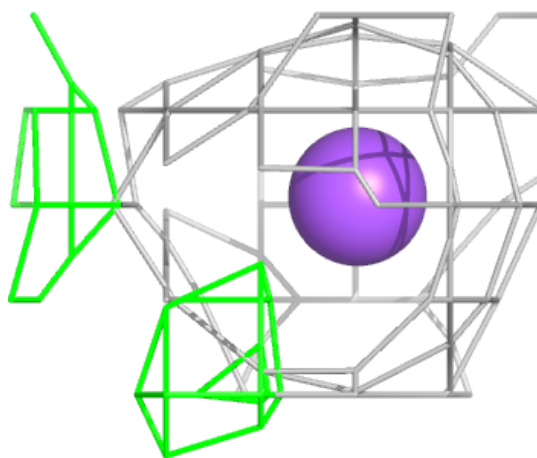mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative)
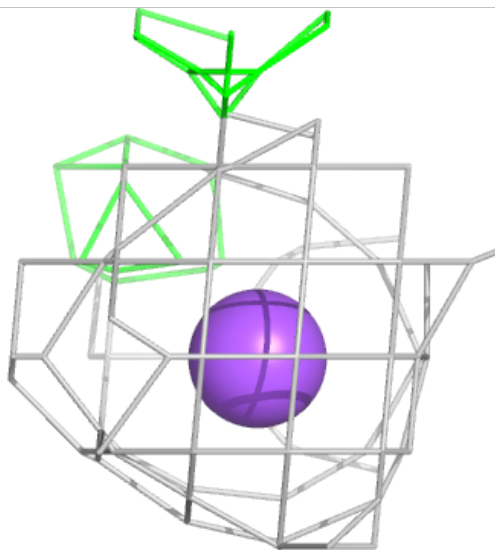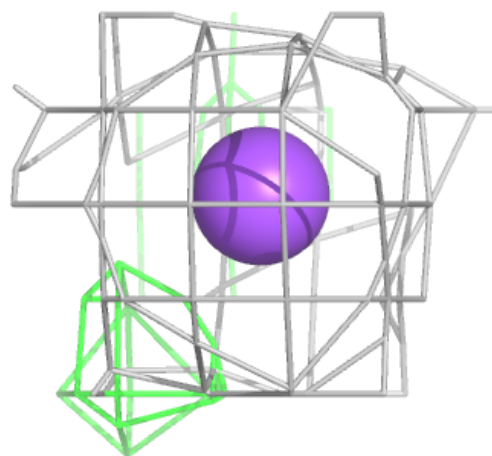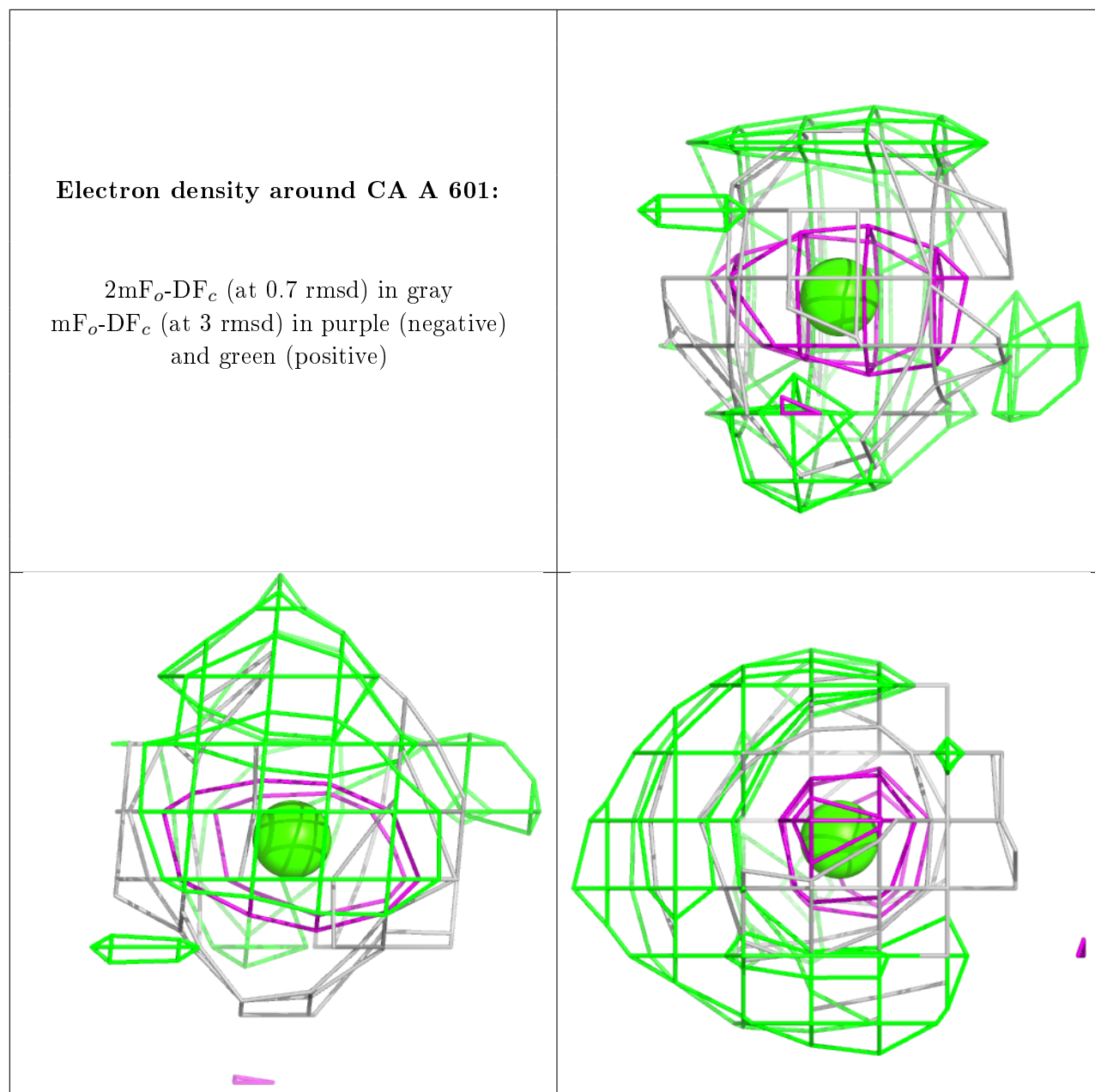and green (positive)
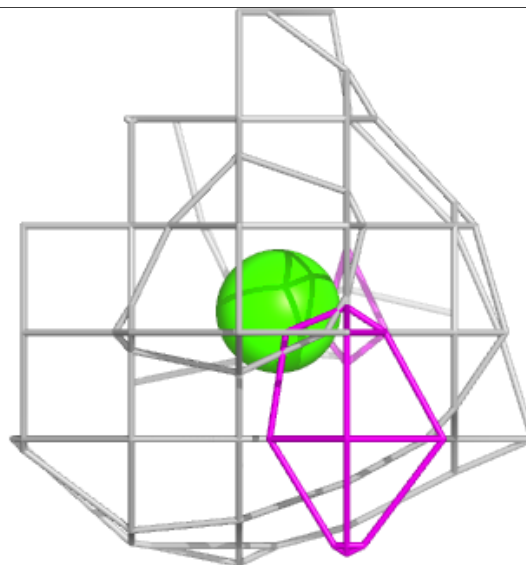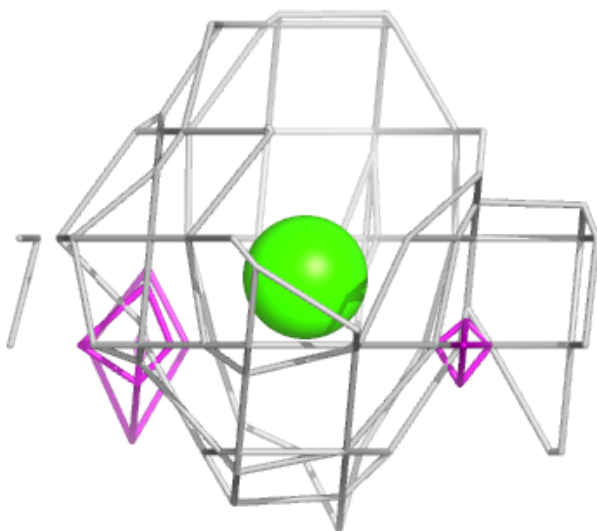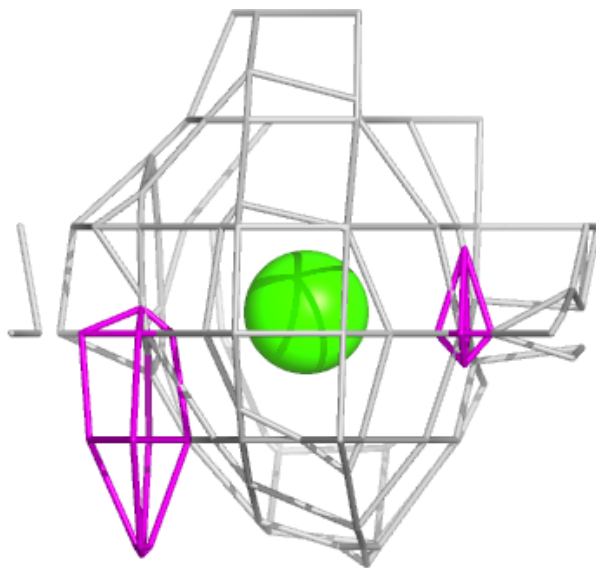
Electron density around NA B 602:

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

Electron density around NA A 602:

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

Electron density around CA A 601:

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

6.5 Other polymers (i)

There are no such residues in this entry.