W O R L D W I D E
wwPDB
PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 16, 2021 – 08:08 AM GMT

|		
PDB ID	:	6ZA7
Title	:	Structure of the apo transcriptional repressor Atu1419 (VanR) from agrobacterium fabrum
Authors	:	Morera, S.; Vigouroux, A.; Legrand, P.
Deposited on	:	2020-06-04
Resolution	:	2.34 Å(reported)

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 2.34 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.
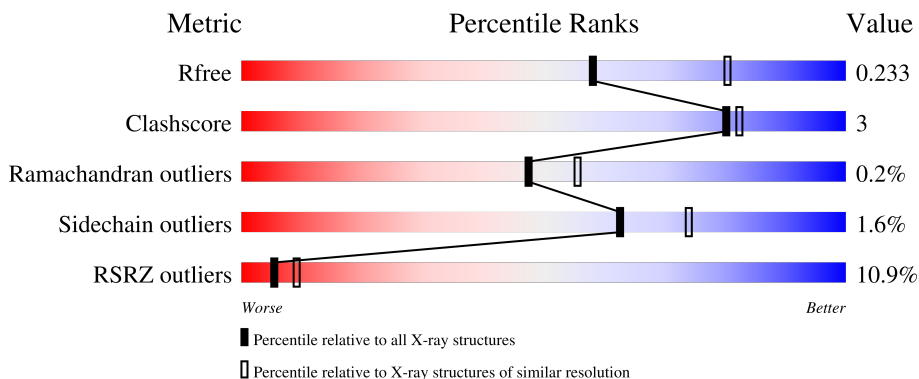
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2096 (2.36-2.32)
Clashscore	141614	2193 (2.36-2.32)
Ramachandran outliers	138981	2159 (2.36-2.32)
Sidechain outliers	138945	2160 (2.36-2.32)
RSRZ outliers	127900	2067 (2.36-2.32)

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	A	250	5% 88% 8% ·
1	B	250	16% 84% 8% · 6%
1	C	250	8% 86% 8% 5%
1	D	250	12% 88% 6% 6%

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 7347 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 Transcriptional regulator, GntR family.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	240	Total	C	N	O	S	0	0	0
			1826	1135	342	343	6			
1	B	234	Total	C	N	O	S	0	0	0
			1780	1104	334	336	6			
1	C	238	Total	C	N	O	S	0	0	0
			1819	1131	342	340	6			
1	D	236	Total	C	N	O	S	0	0	0
			1801	1121	338	336	6			

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	245	HIS	-	expression tag	UNP A9CJ36
A	246	HIS	-	expression tag	UNP A9CJ36
A	247	HIS	-	expression tag	UNP A9CJ36
A	248	HIS	-	expression tag	UNP A9CJ36
A	249	HIS	-	expression tag	UNP A9CJ36
A	250	HIS	-	expression tag	UNP A9CJ36
B	245	HIS	-	expression tag	UNP A9CJ36
B	246	HIS	-	expression tag	UNP A9CJ36
B	247	HIS	-	expression tag	UNP A9CJ36
B	248	HIS	-	expression tag	UNP A9CJ36
B	249	HIS	-	expression tag	UNP A9CJ36
B	250	HIS	-	expression tag	UNP A9CJ36
C	245	HIS	-	expression tag	UNP A9CJ36
C	246	HIS	-	expression tag	UNP A9CJ36
C	247	HIS	-	expression tag	UNP A9CJ36
C	248	HIS	-	expression tag	UNP A9CJ36
C	249	HIS	-	expression tag	UNP A9CJ36
C	250	HIS	-	expression tag	UNP A9CJ36
D	245	HIS	-	expression tag	UNP A9CJ36
D	246	HIS	-	expression tag	UNP A9CJ36
D	247	HIS	-	expression tag	UNP A9CJ36

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	248	HIS	-	expression tag	UNP A9CJ36
D	249	HIS	-	expression tag	UNP A9CJ36
D	250	HIS	-	expression tag	UNP A9CJ36

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total 1	Zn 1	0	0
2	A	1	Total 1	Zn 1	0	0
2	D	1	Total 1	Zn 1	0	0
2	C	1	Total 1	Zn 1	0	0

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).
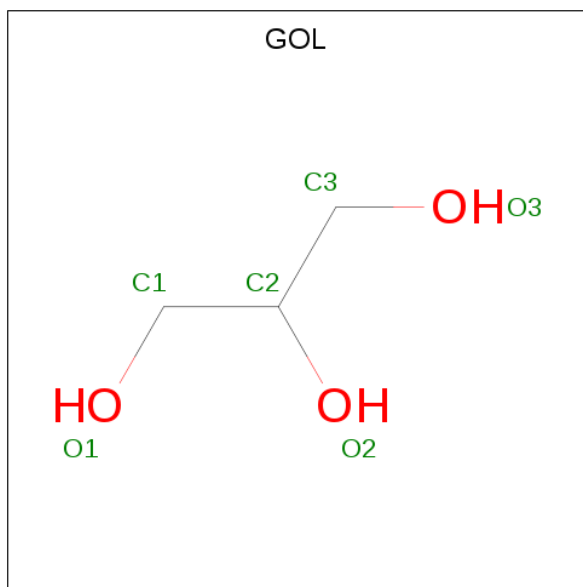
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total 6	C 3	O 3	0	0
3	B	1	Total 6	C 3	O 3	0	0
3	C	1	Total 6	C 3	O 3	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	D	1	Total 6	C 3	O 3	0	0

- Molecule 4 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	B	1	Total 7	C 4	O 3	0	0

- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula: O_4S).
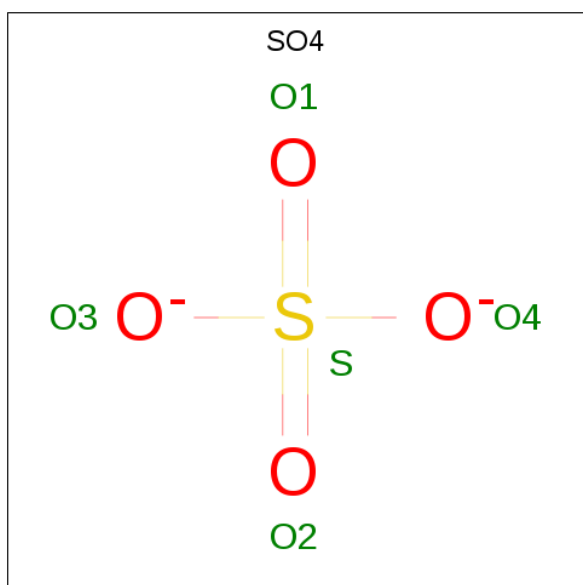
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	C	1	Total 5	O 4	S 1	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	28	Total 28	O 28	0	0
6	B	16	Total 16	O 16	0	0
6	C	20	Total 20	O 20	0	0
6	D	17	Total 17	O 17	0	0

3 Residue-property plots ⓘ

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.
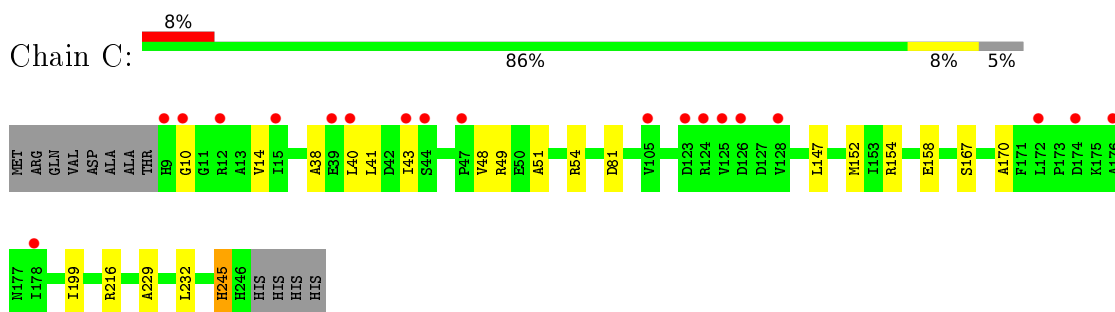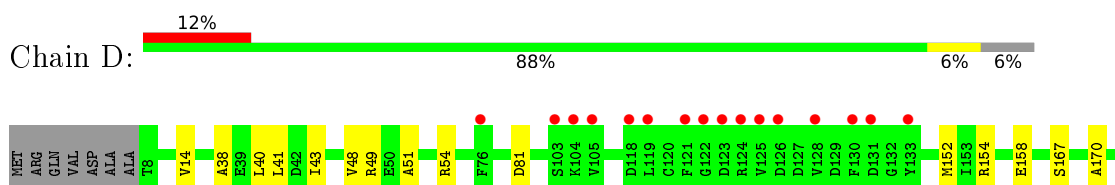
- Molecule 1: Transcriptional regulator, GntR family
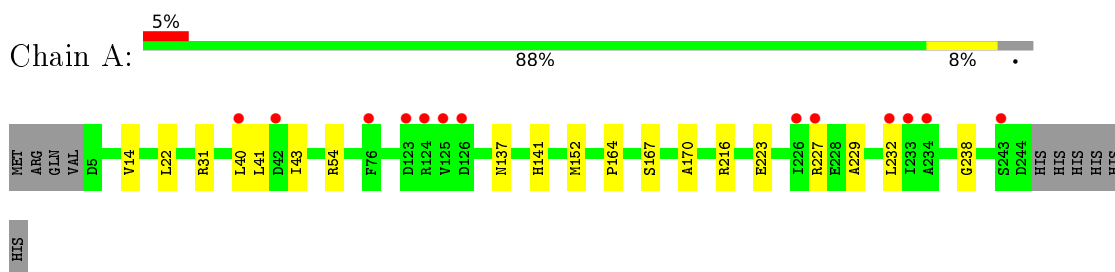
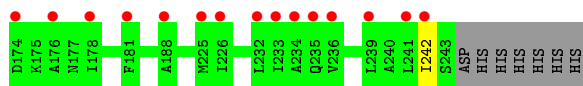- Molecule 1: Transcriptional regulator, GntR family
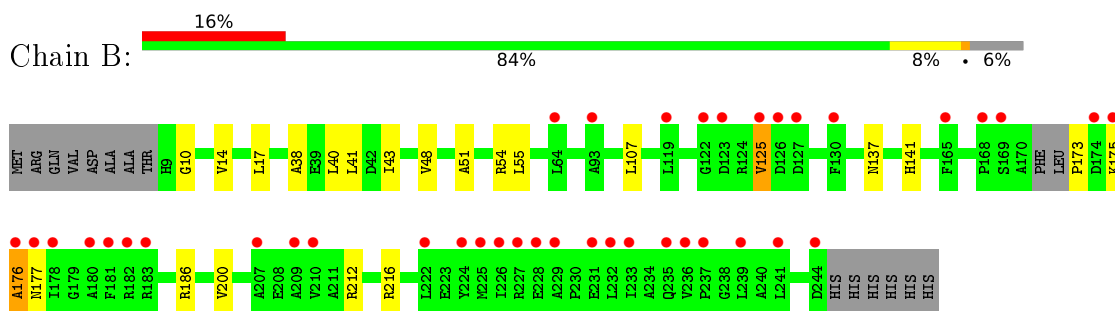
- Molecule 1: Transcriptional regulator, GntR family

- Molecule 1: Transcriptional regulator, GntR family

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	52.11Å 79.01Å 140.61Å 90.00° 95.08° 90.00°	Depositor
Resolution (Å)	43.38 – 2.34 43.59 – 2.34	Depositor EDS
% Data completeness (in resolution range)	71.5 (43.38-2.34) 71.5 (43.59-2.34)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.40 (at 2.34Å)	Xtriage
Refinement program	BUSTER 2.10.3	Depositor
R, R_{free}	0.194 , 0.223 0.204 , 0.233	Depositor DCC
R_{free} test set	1718 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	54.0	Xtriage
Anisotropy	0.082	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 39.3	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	0.95	EDS
Total number of atoms	7347	wwPDB-VP
Average B, all atoms (Å^2)	73.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.82% 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.

wwPDB X-ray Structure Validation Summary Report

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, ZN, PEG, SO4

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	0.45	0/1850	0.64	0/2496
1	B	0.46	0/1801	0.63	0/2426
1	C	0.47	0/1844	0.65	0/2487
1	D	0.46	0/1825	0.62	0/2461
All	All	0.46	0/7320	0.64	0/9870

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	1826	0	1839	8	0
1	B	1780	0	1795	13	0
1	C	1819	0	1829	12	0
1	D	1801	0	1818	9	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
3	A	6	0	8	0	0

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	6	0	8	0	0
3	C	6	0	8	0	0
3	D	6	0	8	0	0
4	B	7	0	10	2	0
5	C	5	0	0	0	0
6	A	28	0	0	0	0
6	B	16	0	0	0	0
6	C	20	0	0	0	0
6	D	17	0	0	0	0
All	All	7347	0	7323	39	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 39 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:17:LEU:HD23	1:B:55:LEU:HD12	1.77	0.66
1:B:137:ASN:O	1:B:141:HIS:HD2	1.83	0.61
1:A:22:LEU:HD22	1:A:164:PRO:HG3	1.83	0.60
1:A:223:GLU:O	1:A:227:ARG:HG2	2.03	0.58
1:C:167:SER:HB3	1:C:170:ALA:HB2	1.85	0.57

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	A	238/250 (95%)	234 (98%)	4 (2%)	0	100	100
1	B	230/250 (92%)	225 (98%)	4 (2%)	1 (0%)	34	38

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	236/250 (94%)	231 (98%)	4 (2%)	1 (0%)	34	38
1	D	234/250 (94%)	231 (99%)	3 (1%)	0	100	100
All	All	938/1000 (94%)	921 (98%)	15 (2%)	2 (0%)	47	55

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	245	HIS
1	B	176	ALA

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	184/195 (94%)	180 (98%)	4 (2%)	52	63
1	B	180/195 (92%)	177 (98%)	3 (2%)	60	72
1	C	184/195 (94%)	181 (98%)	3 (2%)	62	74
1	D	182/195 (93%)	180 (99%)	2 (1%)	73	83
All	All	730/780 (94%)	718 (98%)	12 (2%)	62	74

5 of 12 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	125	VAL
1	B	216	ARG
1	C	245	HIS
1	B	40	LEU
1	C	216	ARG

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 10 ligands modelled in this entry, 4 are monoatomic - leaving 6 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	#$|Z| > 2$
3	GOL	D	302	-	5,5,5	0.16	0	5,5,5	0.37	0
4	PEG	B	303	-	6,6,6	0.35	0	5,5,5	0.42	0
3	GOL	C	302	2	5,5,5	0.10	0	5,5,5	0.20	0
3	GOL	A	302	2	5,5,5	0.22	0	5,5,5	0.37	0
5	SO4	C	303	-	4,4,4	0.15	0	6,6,6	0.11	0
3	GOL	B	302	2	5,5,5	0.10	0	5,5,5	0.16	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
3	GOL	C	302	2	-	0/4/4/4	-
3	GOL	A	302	2	-	0/4/4/4	-
3	GOL	D	302	-	-	0/4/4/4	-
3	GOL	B	302	2	-	0/4/4/4	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	PEG	B	303	-	-	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
4	B	303	PEG	O2-C3-C4-O4
4	B	303	PEG	C4-C3-O2-C2

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	303	PEG	2	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, 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(Å²)	Q<0.9
1	A	240/250 (96%)	0.47	13 (5%)	25	36	33, 59, 110, 129	0
1	B	234/250 (93%)	0.82	40 (17%)	1	2	39, 67, 134, 159	0
1	C	238/250 (95%)	0.49	19 (7%)	12	18	40, 68, 112, 155	0
1	D	236/250 (94%)	0.75	31 (13%)	3	5	43, 74, 126, 164	0
All	All	948/1000 (94%)	0.63	103 (10%)	5	9	33, 67, 121, 164	0

The worst 5 of 103 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	176	ALA	9.9
1	D	233	ILE	9.3
1	B	226	ILE	7.2
1	B	178	ILE	7.1
1	A	232	LEU	7.0

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

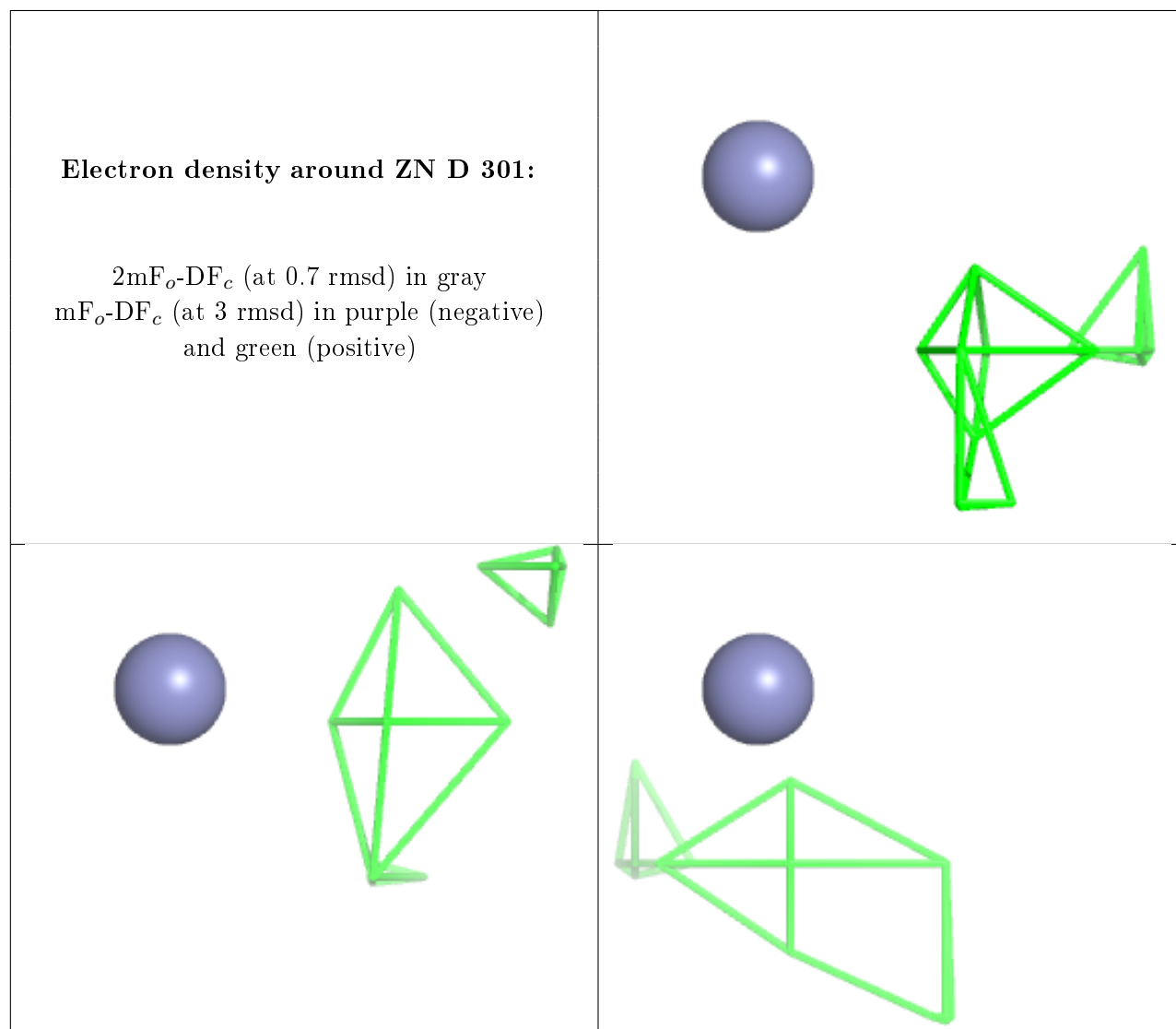There are no non-standard protein/DNA/RNA residues in this entry.

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, 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
5	SO4	C	303	5/5	0.69	0.22	147,147,147,148	0
4	PEG	B	303	7/7	0.83	0.24	51,54,59,61	0
3	GOL	B	302	6/6	0.89	0.14	62,66,69,70	0
3	GOL	D	302	6/6	0.95	0.20	52,55,61,64	0
3	GOL	C	302	6/6	0.96	0.20	58,61,64,65	0
3	GOL	A	302	6/6	0.96	0.24	38,44,49,51	0
2	ZN	D	301	1/1	0.99	0.17	66,66,66,66	0
2	ZN	C	301	1/1	0.99	0.14	69,69,69,69	0
2	ZN	A	301	1/1	0.99	0.18	46,46,46,46	0
2	ZN	B	301	1/1	0.99	0.17	63,63,63,63	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.
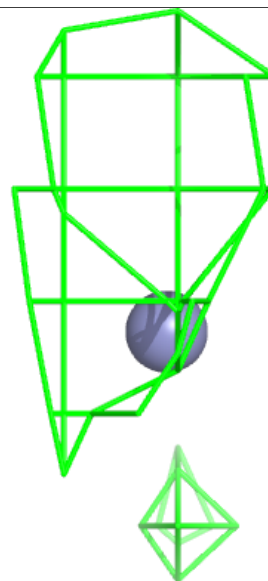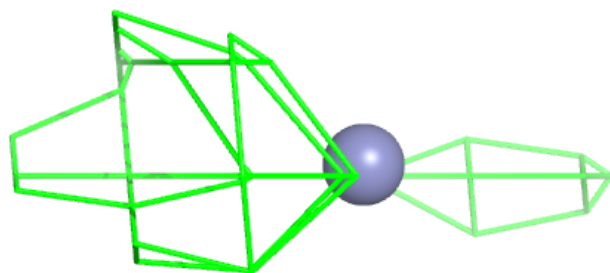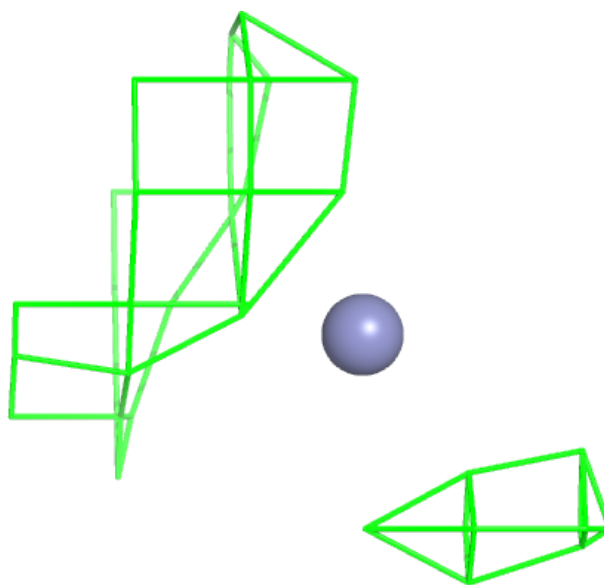

Electron density around ZN D 301:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)
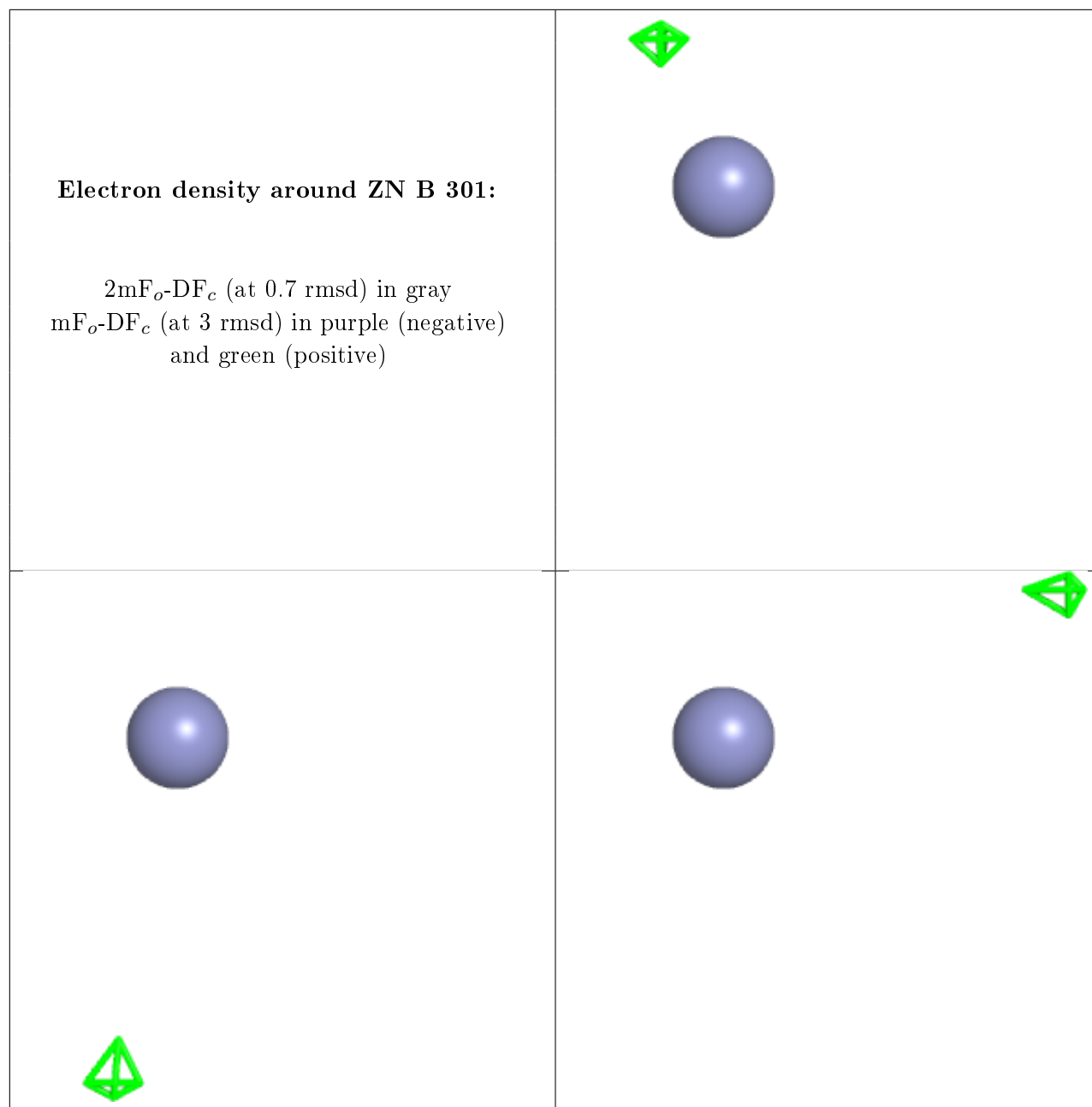
Electron density around ZN C 301:

2mF_o-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)
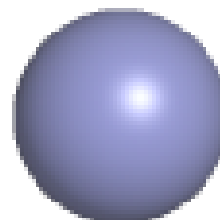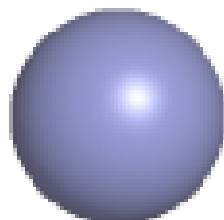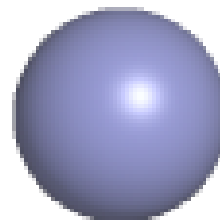
Electron density around ZN A 301:

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

Electron density around ZN B 301:

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

6.5 Other polymers (i)

There are no such residues in this entry.