wwPDB X-ray Structure Validation Summary Report (i)

May 21, 2020 – 01:41 pm BST

PDB ID	:	2OZV
Title	:	Crystal structure of a predicted O-methyltransferase, protein Atu636 from Agrobacterium tumefaciens.
Authors	:	Cuff, M.E.; Xu, X.; Zheng, X.; Edwards, A.; Savchenko, A.; Joachimiak, A.; Midwest Center for Structural Genomics (MCSG)
Deposited on	:	2007-02-27
Resolution	:	1.70 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 following versions of software and data (see references (1)) were used in the production of this report:

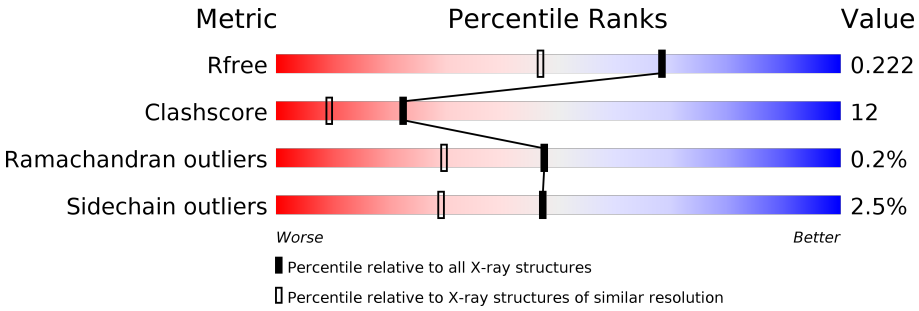
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, 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 1.70 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	4298 (1.70-1.70)
Clashscore	141614	4695 (1.70-1.70)
Ramachandran outliers	138981	4610 (1.70-1.70)
Sidechain outliers	138945	4610 (1.70-1.70)

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%

Mol	Chain	Length	Quality of chain
1	A	260	67% 13% 20%
1	B	260	64% 15% • 19%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 4033 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 Hypothetical protein Atu0636.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	208	Total	C	N	O	S	Se	0	23	0
			1759	1089	326	332	2	10			
1	B	210	Total	C	N	O	S	Se	0	23	0
			1763	1097	325	328	2	11			

There are 66 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	9	MSE	-	CLONING ARTIFACT	UNP Q8UHP4
A	10	GLY	-	CLONING ARTIFACT	UNP Q8UHP4
A	11	SER	-	CLONING ARTIFACT	UNP Q8UHP4
A	12	SER	-	CLONING ARTIFACT	UNP Q8UHP4
A	13	HIS	-	CLONING ARTIFACT	UNP Q8UHP4
A	14	HIS	-	CLONING ARTIFACT	UNP Q8UHP4
A	15	HIS	-	CLONING ARTIFACT	UNP Q8UHP4
A	16	HIS	-	CLONING ARTIFACT	UNP Q8UHP4
A	17	HIS	-	CLONING ARTIFACT	UNP Q8UHP4
A	18	HIS	-	CLONING ARTIFACT	UNP Q8UHP4
A	19	SER	-	CLONING ARTIFACT	UNP Q8UHP4
A	20	SER	-	CLONING ARTIFACT	UNP Q8UHP4
A	21	GLY	-	CLONING ARTIFACT	UNP Q8UHP4
A	22	ARG	-	CLONING ARTIFACT	UNP Q8UHP4
A	23	GLU	-	CLONING ARTIFACT	UNP Q8UHP4
A	24	ASN	-	CLONING ARTIFACT	UNP Q8UHP4
A	25	LEU	-	CLONING ARTIFACT	UNP Q8UHP4
A	26	TYR	-	CLONING ARTIFACT	UNP Q8UHP4
A	27	PHE	-	CLONING ARTIFACT	UNP Q8UHP4
A	28	GLN	-	CLONING ARTIFACT	UNP Q8UHP4
A	29	GLY	-	CLONING ARTIFACT	UNP Q8UHP4
A	30	HIS	-	CLONING ARTIFACT	UNP Q8UHP4
A	31	MSE	MET	MODIFIED RESIDUE	UNP Q8UHP4
A	34	MSE	MET	MODIFIED RESIDUE	UNP Q8UHP4
A	60	MSE	MET	MODIFIED RESIDUE	UNP Q8UHP4

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	80	MSE	MET	MODIFIED RESIDUE	UNP Q8UHP4
A	129	MSE	MET	MODIFIED RESIDUE	UNP Q8UHP4
A	152	MSE	MET	MODIFIED RESIDUE	UNP Q8UHP4
A	168	MSE	MET	MODIFIED RESIDUE	UNP Q8UHP4
A	213	MSE	MET	MODIFIED RESIDUE	UNP Q8UHP4
A	233	MSE	MET	MODIFIED RESIDUE	UNP Q8UHP4
A	267	GLY	-	CLONING ARTIFACT	UNP Q8UHP4
A	268	SER	-	CLONING ARTIFACT	UNP Q8UHP4
B	9	MSE	-	CLONING ARTIFACT	UNP Q8UHP4
B	10	GLY	-	CLONING ARTIFACT	UNP Q8UHP4
B	11	SER	-	CLONING ARTIFACT	UNP Q8UHP4
B	12	SER	-	CLONING ARTIFACT	UNP Q8UHP4
B	13	HIS	-	CLONING ARTIFACT	UNP Q8UHP4
B	14	HIS	-	CLONING ARTIFACT	UNP Q8UHP4
B	15	HIS	-	CLONING ARTIFACT	UNP Q8UHP4
B	16	HIS	-	CLONING ARTIFACT	UNP Q8UHP4
B	17	HIS	-	CLONING ARTIFACT	UNP Q8UHP4
B	18	HIS	-	CLONING ARTIFACT	UNP Q8UHP4
B	19	SER	-	CLONING ARTIFACT	UNP Q8UHP4
B	20	SER	-	CLONING ARTIFACT	UNP Q8UHP4
B	21	GLY	-	CLONING ARTIFACT	UNP Q8UHP4
B	22	ARG	-	CLONING ARTIFACT	UNP Q8UHP4
B	23	GLU	-	CLONING ARTIFACT	UNP Q8UHP4
B	24	ASN	-	CLONING ARTIFACT	UNP Q8UHP4
B	25	LEU	-	CLONING ARTIFACT	UNP Q8UHP4
B	26	TYR	-	CLONING ARTIFACT	UNP Q8UHP4
B	27	PHE	-	CLONING ARTIFACT	UNP Q8UHP4
B	28	GLN	-	CLONING ARTIFACT	UNP Q8UHP4
B	29	GLY	-	CLONING ARTIFACT	UNP Q8UHP4
B	30	HIS	-	CLONING ARTIFACT	UNP Q8UHP4
B	31	MSE	MET	MODIFIED RESIDUE	UNP Q8UHP4
B	34	MSE	MET	MODIFIED RESIDUE	UNP Q8UHP4
B	60	MSE	MET	MODIFIED RESIDUE	UNP Q8UHP4
B	80	MSE	MET	MODIFIED RESIDUE	UNP Q8UHP4
B	129	MSE	MET	MODIFIED RESIDUE	UNP Q8UHP4
B	152	MSE	MET	MODIFIED RESIDUE	UNP Q8UHP4
B	168	MSE	MET	MODIFIED RESIDUE	UNP Q8UHP4
B	213	MSE	MET	MODIFIED RESIDUE	UNP Q8UHP4
B	233	MSE	MET	MODIFIED RESIDUE	UNP Q8UHP4
B	267	GLY	-	CLONING ARTIFACT	UNP Q8UHP4
B	268	SER	-	CLONING ARTIFACT	UNP Q8UHP4

- Molecule 2 is water.

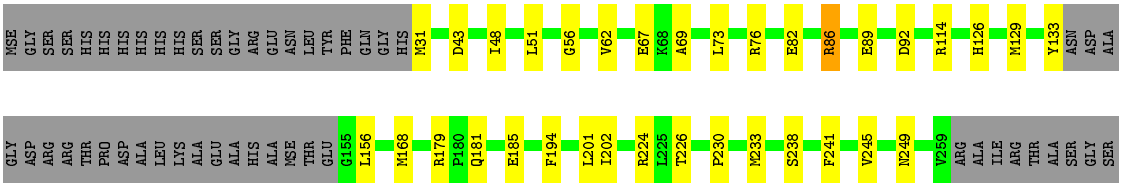
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	266	Total O 266 266	0	0
2	B	245	Total O 245 245	0	0

3 Residue-property plots (i)

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. 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. 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.

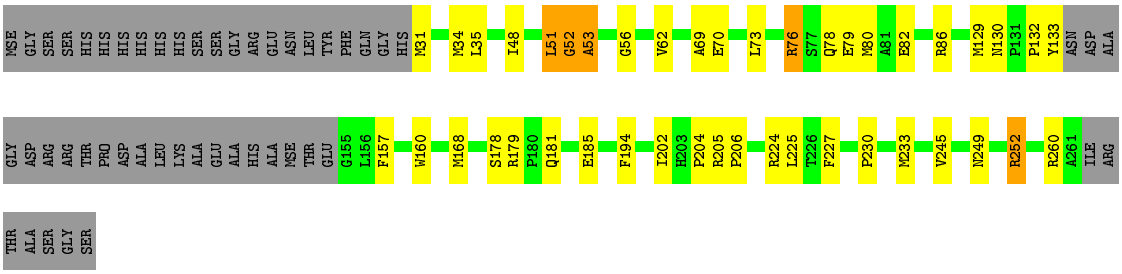
- Molecule 1: Hypothetical protein Atu0636

Chain A:



- Molecule 1: Hypothetical protein Atu0636

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	48.54Å 78.05Å 62.26Å 90.00° 94.55° 90.00°	Depositor
Resolution (Å)	31.03 – 1.70 31.03 – 1.70	Depositor EDS
% Data completeness (in resolution range)	95.5 (31.03-1.70) 95.5 (31.03-1.70)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.08	Depositor
$< I/\sigma(I) >$ [1]	2.24 (at 1.70Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.178 , 0.214 0.189 , 0.222	Depositor DCC
R_{free} test set	2490 reflections (5.12%)	wwPDB-VP
Wilson B-factor (Å^2)	26.5	Xtriage
Anisotropy	0.385	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 36.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	4033	wwPDB-VP
Average B, all atoms (Å^2)	49.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.66% 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

5.1 Standard geometry

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	0.65	0/1786	0.72	0/2395
1	B	0.64	0/1796	0.73	0/2408
All	All	0.64	0/3582	0.72	0/4803

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	B	0	2

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	51	LEU	Peptide
1	B	52[B]	GLY	Peptide

5.2 Too-close contacts

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	1759	0	1735	34	1
1	B	1763	0	1756	55	2

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	266	0	0	10	1
2	B	245	0	0	9	0
All	All	4033	0	3491	82	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 12.

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:76:ARG:HG3	2:B:499:HOH:O	1.45	1.17
1:B:133[B]:TYR:CD2	1:B:179[B]:ARG:NH2	2.16	1.12
1:B:179[A]:ARG:NH1	1:B:181[A]:GLN:OE1	1.92	1.03
1:B:252:ARG:HG2	1:B:252:ARG:HH11	1.31	0.93
1:B:133[B]:TYR:OH	1:B:178[B]:SER:OG	1.91	0.89

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)
1:A:76:ARG:NH1	1:B:79[A]:GLU:OE2[1_556]	2.03	0.17
1:B:79[A]:GLU:OE1	2:A:471:HOH:O[1_554]	2.03	0.17

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	227/260 (87%)	222 (98%)	5 (2%)	0	100 100
1	B	228/260 (88%)	222 (97%)	4 (2%)	2 (1%)	17 5
All	All	455/520 (88%)	444 (98%)	9 (2%)	2 (0%)	47 18

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	53[A]	ALA
1	B	53[B]	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	179/188 (95%)	175 (98%)	4 (2%)	52	34
1	B	179/188 (95%)	174 (97%)	5 (3%)	43	25
All	All	358/376 (95%)	349 (98%)	9 (2%)	47	29

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

Mol	Chain	Res	Type
1	B	31	MSE
1	B	252	ARG
1	B	80	MSE
1	A	92	ASP
1	B	76	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	B	250	ASN

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

There are no carbohydrates in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.