



wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 23, 2024 – 05:34 AM EDT

PDB ID : 6I5Q
Title : Papaver somniferum O-methyltransferase 1
Authors : Cabry, M.P.; Offen, W.A.; Winzer, T.; Li, Y.; Graham, I.A.; Davies, G.J.; Saleh, P.
Deposited on : 2018-11-14
Resolution : 3.05 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 ⓘ 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 ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.20.1
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

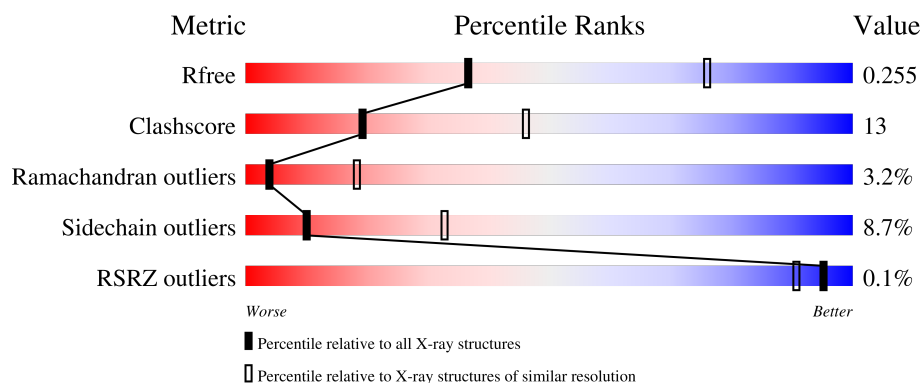
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.05 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	1754 (3.10-3.02)
Clashscore	141614	1864 (3.10-3.02)
Ramachandran outliers	138981	1794 (3.10-3.02)
Sidechain outliers	138945	1793 (3.10-3.02)
RSRZ outliers	127900	1713 (3.10-3.02)

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 $\leq 5\%$. 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	393	<div> <div>60%</div> <div>24%</div> <div>•</div> <div>13%</div> </div>
1	B	393	<div> <div>56%</div> <div>28%</div> <div>•</div> <div>13%</div> </div>
1	C	393	<div> <div>66%</div> <div>20%</div> <div>•</div> <div>13%</div> </div>
1	D	393	<div> <div>64%</div> <div>21%</div> <div>•</div> <div>12%</div> </div>

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 10068 atoms, of which 0 are hydrogens and 0 are deuteriums.

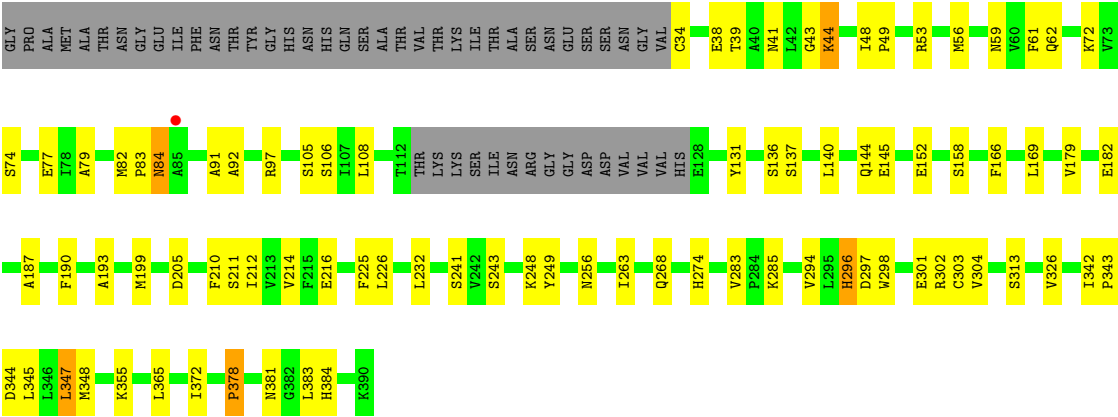
In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called O-methyltransferase 1.

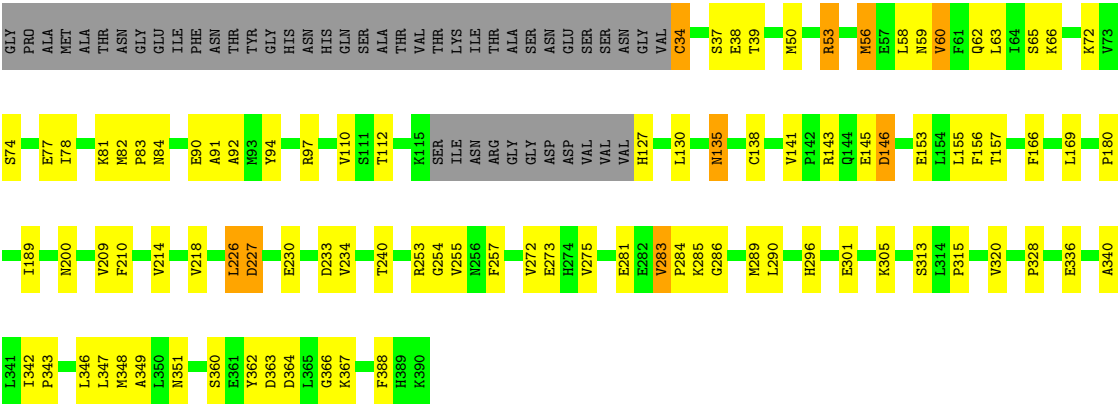
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	342	Total	C	N	O	S	0	0	0
			2546	1646	413	471	16			
1	B	342	Total	C	N	O	S	0	0	0
			2479	1588	407	467	17			
1	C	342	Total	C	N	O	S	0	0	0
			2512	1610	416	471	15			
1	D	346	Total	C	N	O	S	0	0	0
			2531	1616	421	478	16			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	GLY	-	expression tag	UNP I3PLQ5
A	-1	PRO	-	expression tag	UNP I3PLQ5
A	0	ALA	-	expression tag	UNP I3PLQ5
B	-2	GLY	-	expression tag	UNP I3PLQ5
B	-1	PRO	-	expression tag	UNP I3PLQ5
B	0	ALA	-	expression tag	UNP I3PLQ5
C	-2	GLY	-	expression tag	UNP I3PLQ5
C	-1	PRO	-	expression tag	UNP I3PLQ5
C	0	ALA	-	expression tag	UNP I3PLQ5
D	-2	GLY	-	expression tag	UNP I3PLQ5
D	-1	PRO	-	expression tag	UNP I3PLQ5
D	0	ALA	-	expression tag	UNP I3PLQ5



● Molecule 1: O-methyltransferase 1



4 Data and refinement statistics

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α , β , γ	112.00Å 112.00Å 304.82Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	97.00 – 3.05 97.00 – 3.05	Depositor EDS
% Data completeness (in resolution range)	100.0 (97.00-3.05) 100.0 (97.00-3.05)	Depositor EDS
R_{merge}	0.15	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.98 (at 3.07Å)	Xtriage
Refinement program	REFMAC 5.8.0238	Depositor
R, R_{free}	0.210 , 0.256 0.215 , 0.255	Depositor DCC
R_{free} test set	2067 reflections (4.78%)	wwPDB-VP
Wilson B-factor (Å ²)	85.6	Xtriage
Anisotropy	0.054	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 92.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	0.033 for -h,-k,l	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	10068	wwPDB-VP
Average B, all atoms (Å ²)	65.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.86% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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.

5 Model quality [i](#)

5.1 Standard geometry [i](#)

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	0.75	0/2599	0.86	0/3533
1	B	0.77	0/2530	0.86	0/3440
1	C	0.76	0/2562	0.90	0/3483
1	D	0.80	0/2584	0.89	0/3513
All	All	0.77	0/10275	0.88	0/13969

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	2546	0	2460	67	0
1	B	2479	0	2291	103	0
1	C	2512	0	2359	54	0
1	D	2531	0	2354	55	0
All	All	10068	0	9464	252	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 13.

The worst 5 of 252 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:311:TRP:CD1	1:B:390:LYS:HE3	1.63	1.32
1:B:379:ILE:HD11	1:B:385:VAL:CG2	1.70	1.20
1:B:363:ASP:OD1	1:B:374:THR:OG1	1.69	1.09
1:B:379:ILE:CD1	1:B:385:VAL:HG23	1.83	1.08
1:B:379:ILE:HD11	1:B:385:VAL:HG23	1.10	1.05

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	338/393 (86%)	272 (80%)	55 (16%)	11 (3%)	4	17
1	B	338/393 (86%)	260 (77%)	63 (19%)	15 (4%)	2	12
1	C	338/393 (86%)	295 (87%)	37 (11%)	6 (2%)	8	30
1	D	342/393 (87%)	284 (83%)	46 (14%)	12 (4%)	3	17
All	All	1356/1572 (86%)	1111 (82%)	201 (15%)	44 (3%)	4	18

5 of 44 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	227	ASP
1	B	60	VAL
1	B	70	ASP
1	D	59	ASN
1	D	91	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	258/333 (78%)	235 (91%)	23 (9%)	9	31
1	B	238/333 (72%)	215 (90%)	23 (10%)	8	27
1	C	245/333 (74%)	228 (93%)	17 (7%)	15	42
1	D	246/333 (74%)	223 (91%)	23 (9%)	9	29
All	All	987/1332 (74%)	901 (91%)	86 (9%)	10	33

5 of 86 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	296	HIS
1	D	127	HIS
1	C	304	VAL
1	D	56	MET
1	D	189	ILE

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

Mol	Chain	Res	Type
1	A	41	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

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	342/393 (87%)	-0.30	0	100	100	36, 60, 95, 124	52 (15%)
1	B	342/393 (87%)	-0.37	0	100	100	42, 72, 102, 121	95 (27%)
1	C	342/393 (87%)	-0.35	1 (0%)	94	85	38, 63, 92, 121	124 (36%)
1	D	346/393 (88%)	-0.41	0	100	100	36, 63, 89, 118	139 (40%)
All	All	1372/1572 (87%)	-0.36	1 (0%)	95	91	36, 65, 97, 124	410 (29%)

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	85	ALA	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 monosaccharides in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.