



wwPDB X-ray Structure Validation Summary Report

Feb 26, 2014 – 04:18 PM GMT

PDB ID : 1J4S
Title : Structure of Artocarpin: a Lectin with Mannose Specificity (Form 1)
Authors : Pratap, J.V.; Jeyaprakash, A.A.; Rani, P.G.; Sekar, K.; Surolia, A.; Vijayan, M.
Deposited on : 2001-10-30
Resolution : 2.50 Å(reported)

This is a wwPDB validation summary report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

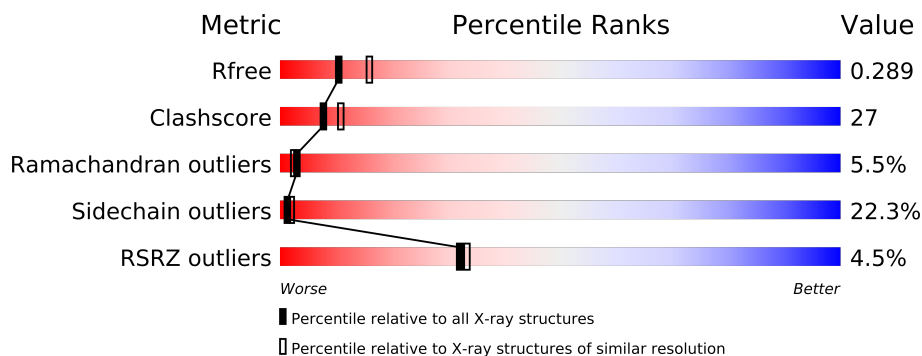
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

1 Overall quality at a glance

The reported resolution of this entry is 2.50 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}	66092	2784 (2.50-2.50)
Clashscore	79885	3562 (2.50-2.50)
Ramachandran outliers	78287	3480 (2.50-2.50)
Sidechain outliers	78261	3482 (2.50-2.50)
RSRZ outliers	66119	2785 (2.50-2.50)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	149	
1	B	149	
1	C	149	
1	D	149	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 4917 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Artocarpin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	149	Total	C	N	O	S	0	0	0
			1119	718	181	219	1			
1	B	149	Total	C	N	O	S	0	0	0
			1118	717	181	219	1			
1	C	149	Total	C	N	O	S	0	0	0
			1126	722	182	221	1			
1	D	149	Total	C	N	O	S	0	0	0
			1123	720	181	221	1			

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	9	SER	PRO	CONFLICT	UNP Q7M1T4
A	20	GLU	ASP	CONFLICT	UNP Q7M1T4
A	49	ASP	GLU	CONFLICT	UNP Q7M1T4
A	70	LYS	ARG	CONFLICT	UNP Q7M1T4
A	84	GLY	ALA	CONFLICT	UNP Q7M1T4
A	145	ILE	VAL	CONFLICT	UNP Q7M1T4
A	148	SER	ALA	CONFLICT	UNP Q7M1T4
B	9	SER	PRO	CONFLICT	UNP Q7M1T4
B	20	GLU	ASP	CONFLICT	UNP Q7M1T4
B	49	ASP	GLU	CONFLICT	UNP Q7M1T4
B	70	LYS	ARG	CONFLICT	UNP Q7M1T4
B	84	GLY	ALA	CONFLICT	UNP Q7M1T4
B	145	ILE	VAL	CONFLICT	UNP Q7M1T4
B	148	SER	ALA	CONFLICT	UNP Q7M1T4
C	9	SER	PRO	CONFLICT	UNP Q7M1T4
C	20	GLU	ASP	CONFLICT	UNP Q7M1T4
C	49	ASP	GLU	CONFLICT	UNP Q7M1T4
C	70	LYS	ARG	CONFLICT	UNP Q7M1T4
C	84	GLY	ALA	CONFLICT	UNP Q7M1T4
C	145	ILE	VAL	CONFLICT	UNP Q7M1T4
C	148	SER	ALA	CONFLICT	UNP Q7M1T4

Continued on next page...

Continued from previous page...

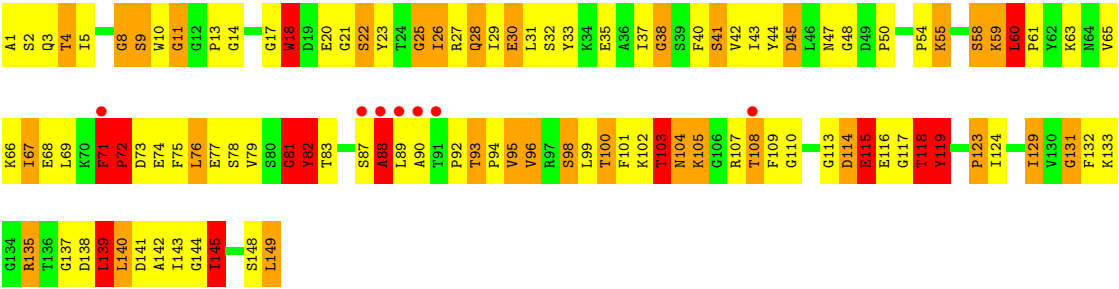
Chain	Residue	Modelled	Actual	Comment	Reference
D	9	SER	PRO	CONFLICT	UNP Q7M1T4
D	20	GLU	ASP	CONFLICT	UNP Q7M1T4
D	49	ASP	GLU	CONFLICT	UNP Q7M1T4
D	70	LYS	ARG	CONFLICT	UNP Q7M1T4
D	84	GLY	ALA	CONFLICT	UNP Q7M1T4
D	145	ILE	VAL	CONFLICT	UNP Q7M1T4
D	148	SER	ALA	CONFLICT	UNP Q7M1T4

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	102	Total	O	0	0
			102	102		
2	B	112	Total	O	0	0
			112	112		
2	C	105	Total	O	0	0
			105	105		
2	D	112	Total	O	0	0
			112	112		

● Molecule 1: Artocarpin

Chain D: 



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	69.88Å 73.74Å 60.64Å 90.00° 95.06° 90.00°	Depositor
Resolution (Å)	20.00 – 2.50 10.00 – 2.45	Depositor EDS
% Data completeness (in resolution range)	(Not available) (20.00-2.50) 78.9 (10.00-2.45)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.96 (at 2.44Å)	Xtriage
Refinement program	X-PLOR 3.1	Depositor
R, R_{free}	0.199 , 0.262 0.230 , 0.289	Depositor DCC
R_{free} test set	866 reflections (4.93%)	DCC
Wilson B-factor (Å ²)	19.4	Xtriage
Anisotropy	0.610	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.24 , 29.6	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 17578 reflections	Xtriage
F_o, F_c correlation	0.86	EDS
Total number of atoms	4917	wwPDB-VP
Average B, all atoms (Å ²)	11.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.33% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

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	3.37	3/1149 (0.3%)	2.62	51/1561 (3.3%)
1	B	1.01	3/1149 (0.3%)	2.60	75/1563 (4.8%)
1	C	0.84	2/1157 (0.2%)	2.11	59/1572 (3.8%)
1	D	3.21	2/1153 (0.2%)	2.56	75/1566 (4.8%)
All	All	2.41	10/4608 (0.2%)	2.48	260/6262 (4.2%)

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	A	1	5
1	B	1	7
1	C	1	5
1	D	1	4
All	All	4	21

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1	ALA	C-O	109.76	3.31	1.23
1	D	1	ALA	C-O	105.32	3.23	1.23
1	B	70	LYS	C-N	17.96	1.75	1.34
1	A	82	TYR	C-N	-11.34	1.07	1.34
1	D	82	TYR	C-N	-8.24	1.15	1.34

The worst 5 of 260 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	1	ALA	CA-C-O	-42.77	30.29	120.10
1	D	1	ALA	CA-C-O	-41.92	32.06	120.10
1	A	82	TYR	O-C-N	-31.99	71.51	122.70

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	81	GLY	O-C-N	-28.68	76.81	122.70
1	B	71	PHE	C-N-CD	-27.11	60.95	120.60

All (4) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	1	ALA	CA
1	B	1	ALA	CA
1	C	1	ALA	CA
1	D	1	ALA	CA

5 of 21 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	114	ASP	Mainchain
1	A	81	GLY	Mainchain,Peptide
1	A	82	TYR	Mainchain
1	A	9	SER	Mainchain
1	B	8	GLY	Peptide

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1119	0	1060	73	0
1	B	1118	0	1059	68	0
1	C	1126	0	1076	46	0
1	D	1123	0	1064	65	0
2	A	102	0	0	9	0
2	B	112	0	0	7	0
2	C	105	0	0	6	0
2	D	112	0	0	9	0
All	All	4917	0	4259	239	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 27.

The worst 5 of 239 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:70:LYS:C	1:B:71:PHE:N	1.75	1.36
1:B:82:TYR:HE2	1:B:115:GLU:HG3	1.10	1.14
1:B:82:TYR:CE2	1:B:115:GLU:HG3	1.92	1.05
1:C:72:PRO:HD3	2:C:171:HOH:O	1.62	0.99
1:A:82:TYR:N	1:A:82:TYR:CD1	2.26	0.99

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone ⓘ

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	146/149 (98%)	124 (85%)	13 (9%)	9 (6%)	2	2
1	B	147/149 (99%)	130 (88%)	11 (8%)	6 (4%)	4	5
1	C	147/149 (99%)	129 (88%)	9 (6%)	9 (6%)	2	2
1	D	146/149 (98%)	128 (88%)	10 (7%)	8 (6%)	3	2
All	All	586/596 (98%)	511 (87%)	43 (7%)	32 (6%)	3	2

5 of 32 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	71	PHE
1	A	72	PRO
1	A	82	TYR
1	A	99	LEU
1	A	110	GLY

5.3.2 Protein sidechains ⓘ

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	117/122 (96%)	88 (75%)	29 (25%)	1	1
1	B	117/122 (96%)	91 (78%)	26 (22%)	1	2
1	C	119/122 (98%)	95 (80%)	24 (20%)	2	3
1	D	118/122 (97%)	92 (78%)	26 (22%)	1	2
All	All	471/488 (96%)	366 (78%)	105 (22%)	1	2

5 of 105 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	118	THR
1	C	58	SER
1	D	114	ASP
1	B	122	LEU
1	B	148	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	47	ASN
1	B	104	ASN
1	C	121	ASN
1	B	16	ASN
1	C	104	ASN

5.3.3 RNA ⓘ

There are no RNA chains 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 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 ⓘ

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	149/149 (100%)	0.29	6 (4%) 36 37	2, 7, 20, 47	0
1	B	149/149 (100%)	0.35	8 (5%) 25 25	2, 9, 28, 51	0
1	C	149/149 (100%)	0.30	6 (4%) 36 37	2, 7, 26, 48	0
1	D	149/149 (100%)	0.41	7 (4%) 30 31	2, 9, 30, 56	0
All	All	596/596 (100%)	0.34	27 (4%) 32 33	2, 8, 27, 56	0

The worst 5 of 27 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	91	THR	8.5
1	C	91	THR	6.0
1	B	9	SER	4.7
1	D	87	SER	4.4
1	B	87	SER	4.3

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

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

6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.