



wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 26, 2017 – 06:50 PM EDT

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 : unknown
Resolution : 2.50 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

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

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

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20030345
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20030345

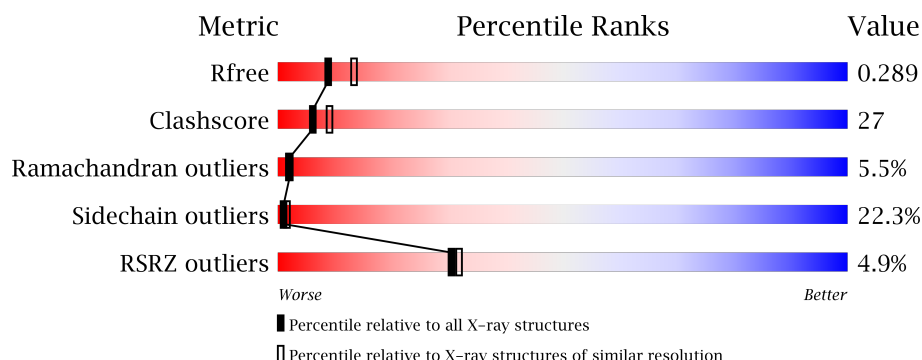
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 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}	100719	3846 (2.50-2.50)
Clashscore	112137	4554 (2.50-2.50)
Ramachandran outliers	110173	4463 (2.50-2.50)
Sidechain outliers	110143	4465 (2.50-2.50)
RSRZ outliers	101464	3876 (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. 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	149	<div> <div>3%</div> <div>34%</div> <div>38%</div> <div>19%</div> <div>9%</div> </div>
1	B	149	<div> <div>5%</div> <div>27%</div> <div>47%</div> <div>19%</div> <div>7%</div> </div>
1	C	149	<div> <div>5%</div> <div>35%</div> <div>47%</div> <div>12%</div> <div>6%</div> </div>
1	D	149	<div> <div>6%</div> <div>28%</div> <div>42%</div> <div>21%</div> <div>9%</div> </div>

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 4917 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 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

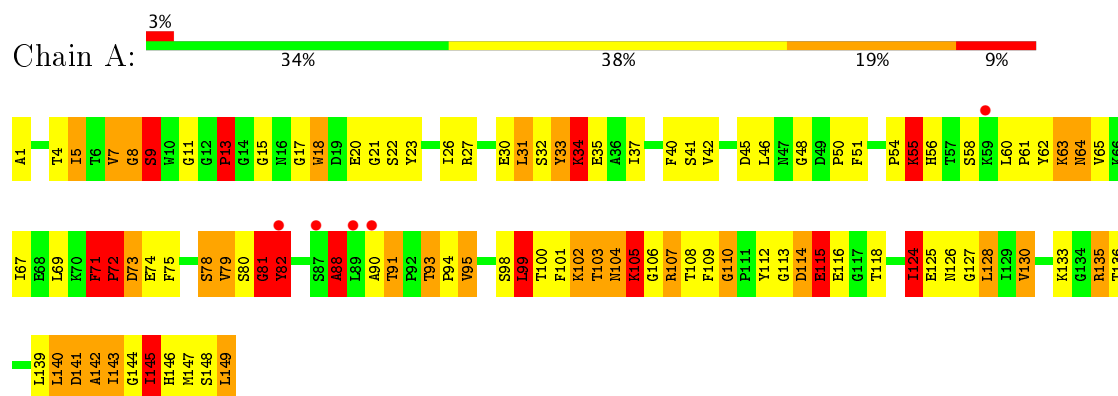
- Molecule 2 is water.

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

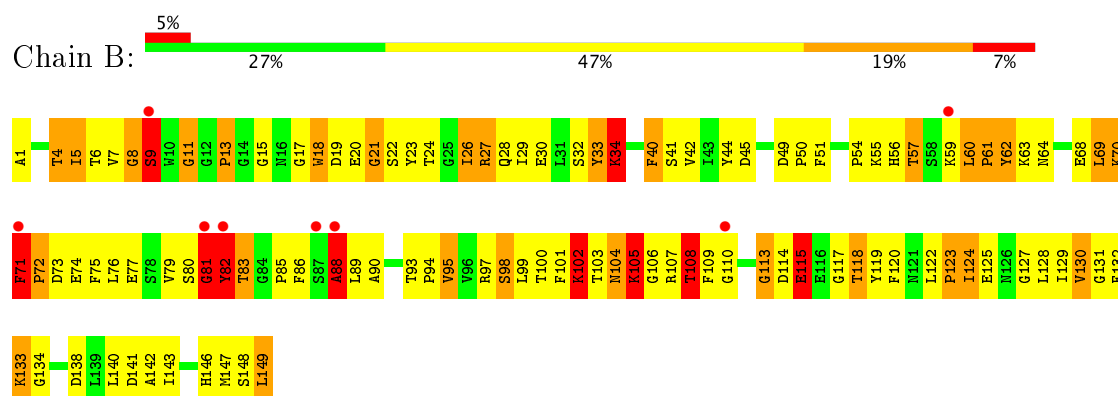
3 Residue-property plots [i](#)

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 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.

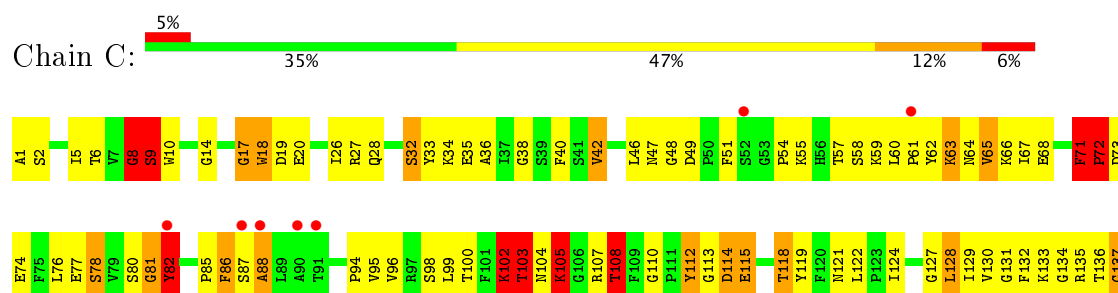
• Molecule 1: Artocarpin

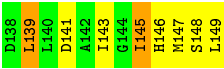


• Molecule 1: Artocarpin

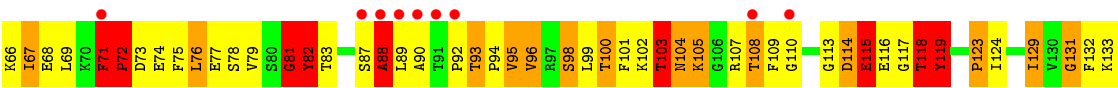


• Molecule 1: Artocarpin





● Molecule 1: Artocarpin



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)	0.25 , 57.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	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.

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

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 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	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

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 27.

The worst 5 of 239 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: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 [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	146 / 149 (98%)	124 (85%)	13 (9%)	9 (6%)	2	1
1	B	147 / 149 (99%)	130 (88%)	11 (8%)	6 (4%)	3	4
1	C	147 / 149 (99%)	129 (88%)	9 (6%)	9 (6%)	2	1
1	D	146 / 149 (98%)	128 (88%)	10 (7%)	8 (6%)	2	2
All	All	586 / 596 (98%)	511 (87%)	43 (7%)	32 (6%)	2	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 [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	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%)	1	2
1	D	118/122 (97%)	92 (78%)	26 (22%)	1	2
All	All	471/488 (96%)	366 (78%)	105 (22%)	1	1

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 [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](#)

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](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	2
1	D	2
1	B	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	D	1:ALA	C	2:SER	N	3.92
1	A	1:ALA	C	2:SER	N	3.83
1	B	70:LYS	C	71:PHE	N	1.75
1	D	82:TYR	C	83:THR	N	1.15
1	A	82:TYR	C	83:THR	N	1.08

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	149/149 (100%)	0.21	5 (3%) 46 48	2, 7, 20, 47	0
1	B	149/149 (100%)	0.28	8 (5%) 26 27	2, 9, 28, 51	0
1	C	149/149 (100%)	0.24	7 (4%) 32 34	2, 7, 26, 48	0
1	D	149/149 (100%)	0.35	9 (6%) 23 23	2, 9, 30, 56	0
All	All	596/596 (100%)	0.27	29 (4%) 30 32	2, 8, 27, 56	0

The worst 5 of 29 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	91	THR	9.3
1	C	91	THR	6.1
1	B	9	SER	4.8
1	B	87	SER	4.7
1	D	90	ALA	4.6

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](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.