



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 14, 2017 – 09:55 am GMT

PDB ID : 2D3R  
Title : Cratylia folibunda seed lectin at acidic pH  
Authors : Del Sol, F.G.; Cavada, B.S.; Calvete, J.J.  
Deposited on : 2005-09-30  
Resolution : 2.90 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](mailto: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 : trunk28620  
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) : recalc28949

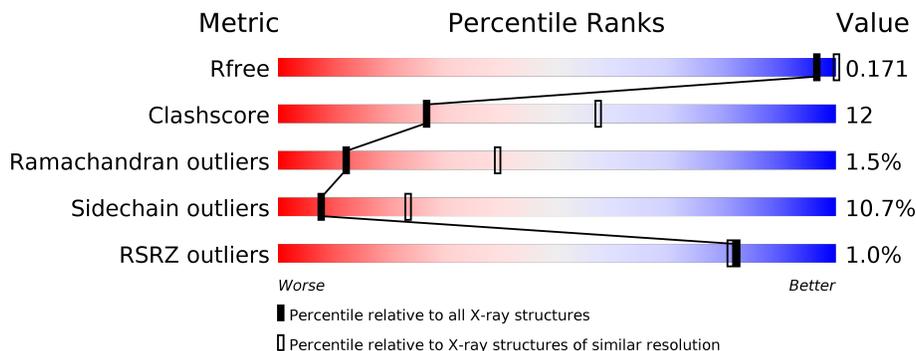
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.90 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	1586 (2.90-2.90)
Clashscore	112137	1807 (2.90-2.90)
Ramachandran outliers	110173	1768 (2.90-2.90)
Sidechain outliers	110143	1770 (2.90-2.90)
RSRZ outliers	101464	1596 (2.90-2.90)

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  $\geq 3$ , 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	236	<p>72% 24%</p>
1	B	236	<p>2% 67% 28%</p>
1	C	236	<p>0% 70% 23% 6%</p>
1	D	236	<p>70% 25% 5%</p>

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 7258 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 Lectin alpha chain.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
1	A	236	1794	1126	305	363	0	0	0
1	B	236	1794	1126	305	363	0	0	0
1	C	236	1794	1126	305	363	0	0	0
1	D	236	1794	1126	305	363	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	211	PHE	TRP	SEE REMARK 999	UNP P81517
B	211	PHE	TRP	SEE REMARK 999	UNP P81517
C	211	PHE	TRP	SEE REMARK 999	UNP P81517
D	211	PHE	TRP	SEE REMARK 999	UNP P81517

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

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

- Molecule 3 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total Mn 1 1	0	0
3	A	1	Total Mn 1 1	0	0
3	D	1	Total Mn 1 1	0	0
3	C	1	Total Mn 1 1	0	0

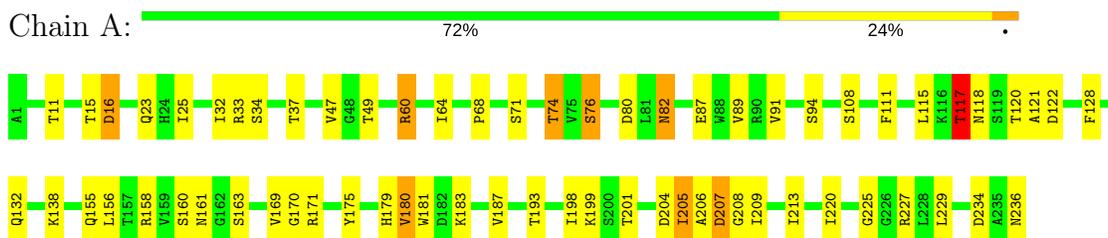
- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	22	Total O 22 22	0	0
4	B	21	Total O 21 21	0	0
4	C	18	Total O 18 18	0	0
4	D	13	Total O 13 13	0	0

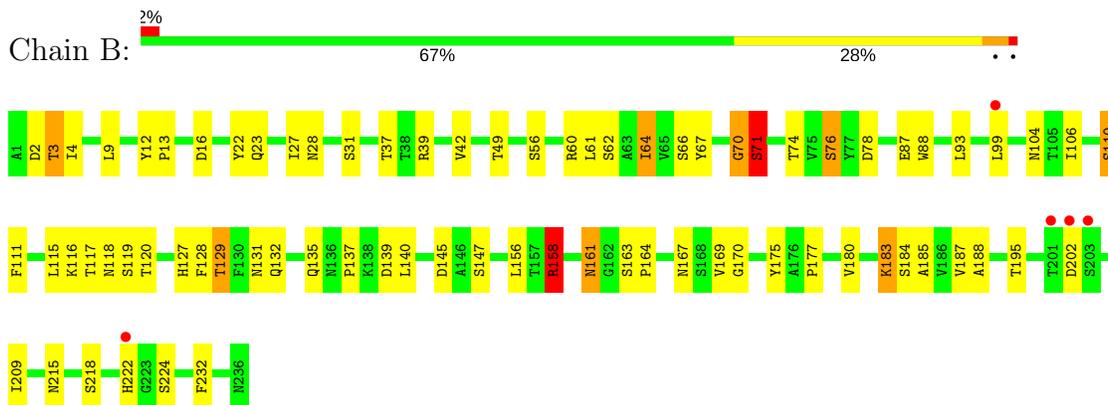
### 3 Residue-property plots

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.

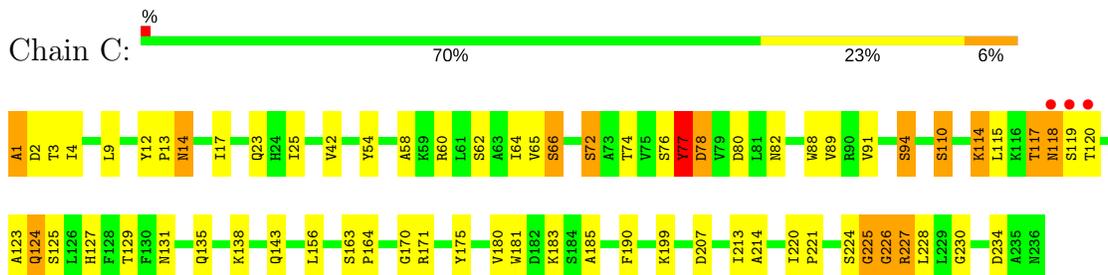
- Molecule 1: Lectin alpha chain



- Molecule 1: Lectin alpha chain

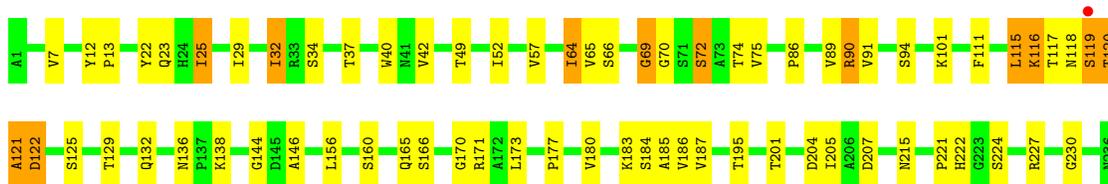


- Molecule 1: Lectin alpha chain



- Molecule 1: Lectin alpha chain





## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	60.21Å 125.44Å 126.14Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.80 – 2.90 49.86 – 2.80	Depositor EDS
% Data completeness (in resolution range)	(Not available) (49.80-2.90) 68.3 (49.86-2.80)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.77 (at 2.81Å)	Xtrriage
Refinement program	REFMAC 5.2.0005	Depositor
R, $R_{free}$	0.178 , 0.220 0.184 , 0.171	Depositor DCC
$R_{free}$ test set	1393 reflections (9.79%)	DCC
Wilson B-factor (Å <sup>2</sup> )	45.3	Xtrriage
Anisotropy	0.315	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 33.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtrriage
Estimated twinning fraction	0.036 for -h,l,k	Xtrriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	7258	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	34.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.13% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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

Bond lengths and bond angles in the following residue types are not validated in this section: CA, MN

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	0.52	0/1833	0.70	2/2500 (0.1%)
1	B	0.53	0/1833	0.66	1/2500 (0.0%)
1	C	0.64	1/1833 (0.1%)	0.88	9/2500 (0.4%)
1	D	0.53	0/1833	0.73	1/2500 (0.0%)
All	All	0.56	1/7332 (0.0%)	0.75	13/10000 (0.1%)

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	0
1	C	0	2
All	All	1	2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	227	ARG	CB-CG	7.38	1.72	1.52

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	2	ASP	CA-C-N	-13.64	87.20	117.20
1	D	121	ALA	N-CA-C	-10.89	81.58	111.00
1	C	2	ASP	O-C-N	9.32	137.62	122.70
1	C	1	ALA	C-N-CA	8.92	144.01	121.70
1	C	2	ASP	C-N-CA	7.96	141.61	121.70

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	117	THR	CB

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	77	TYR	Mainchain,Peptide

## 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	1794	0	1736	41	0
1	B	1794	0	1736	49	12
1	C	1794	0	1735	51	12
1	D	1794	0	1736	47	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	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
4	A	22	0	0	0	0
4	B	21	0	0	1	0
4	C	18	0	0	0	0
4	D	13	0	0	0	0
All	All	7258	0	6943	164	12

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 164 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:D:117:THR:CG2	1:D:185:ALA:HA	1.67	1.23
1:D:117:THR:HG22	1:D:185:ALA:CA	1.75	1.14
1:D:117:THR:HB	1:D:184:SER:O	1.51	1.11
1:C:3:THR:O	1:C:4:ILE:HG12	1.49	1.10
1:C:110:SER:HB3	1:C:129:THR:HG22	1.03	1.02

The worst 5 of 12 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:B:222:HIS:CG	1:C:1:ALA:N[1_655]	1.05	1.15
1:B:222:HIS:CE1	1:C:1:ALA:CB[1_655]	1.07	1.13
1:B:222:HIS:NE2	1:C:1:ALA:CA[1_655]	1.36	0.84
1:B:222:HIS:CD2	1:C:1:ALA:CA[1_655]	1.46	0.74
1:B:222:HIS:ND1	1:C:1:ALA:N[1_655]	1.58	0.62

## 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	234/236 (99%)	218 (93%)	15 (6%)	1 (0%)	38 72
1	B	234/236 (99%)	212 (91%)	16 (7%)	6 (3%)	6 24
1	C	234/236 (99%)	214 (92%)	17 (7%)	3 (1%)	14 43
1	D	234/236 (99%)	213 (91%)	17 (7%)	4 (2%)	11 36
All	All	936/944 (99%)	857 (92%)	65 (7%)	14 (2%)	12 39

5 of 14 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	221	PRO
1	D	121	ALA

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	184	SER
1	C	226	GLY
1	D	69	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	199/199 (100%)	176 (88%)	23 (12%)	6	19
1	B	199/199 (100%)	178 (89%)	21 (11%)	8	24
1	C	199/199 (100%)	181 (91%)	18 (9%)	11	33
1	D	199/199 (100%)	176 (88%)	23 (12%)	6	19
All	All	796/796 (100%)	711 (89%)	85 (11%)	8	23

5 of 85 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	161	ASN
1	C	72	SER
1	D	183	LYS
1	B	169	VAL
1	B	218	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	131	ASN
1	B	135	GLN
1	C	124	GLN
1	B	104	ASN
1	C	118	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](#)

Of 8 ligands modelled in this entry, 8 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

### 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	236/236 (100%)	-0.34	0 <b>100</b> <b>100</b>	19, 32, 44, 48	0
1	B	236/236 (100%)	-0.19	5 (2%) 64 60	20, 35, 50, 54	0
1	C	236/236 (100%)	-0.13	3 (1%) 77 76	23, 38, 52, 57	0
1	D	236/236 (100%)	-0.26	1 (0%) 92 92	20, 34, 44, 51	0
All	All	944/944 (100%)	-0.23	9 (0%) 82 81	19, 35, 48, 57	0

The worst 5 of 9 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	120	THR	4.2
1	B	99	LEU	3.8
1	C	119	SER	2.9
1	B	222	HIS	2.9
1	B	203	SER	2.6

### 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](#)

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
2	CA	A	237	1/1	0.97	0.07	-2.33	30,30,30,30	0
2	CA	B	241	1/1	0.96	0.06	-2.45	37,37,37,37	0
3	MN	C	244	1/1	0.98	0.10	-2.89	34,34,34,34	0
2	CA	D	239	1/1	0.97	0.07	-3.13	32,32,32,32	0
3	MN	B	242	1/1	0.97	0.08	-3.16	32,32,32,32	0
3	MN	D	240	1/1	0.97	0.07	-3.84	45,45,45,45	0
3	MN	A	238	1/1	0.99	0.08	-4.16	21,21,21,21	0
2	CA	C	243	1/1	0.93	0.06	-4.21	49,49,49,49	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.