



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 23, 2020 – 10:07 pm BST

PDB ID : 2VS6  
Title : K173A, R174A, K177A-trichosanthin  
Authors : Too, P.H.; Ma, M.K.; Mak, A.N.; Tung, C.K.; Zhu, G.; Au, S.W.; Wong, K.B.;  
Shaw, P.C.; Ng, A.  
Deposited on : 2008-04-21  
Resolution : 2.40 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

<https://www.wwpdb.org/validation/2017/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)	:	1.13
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

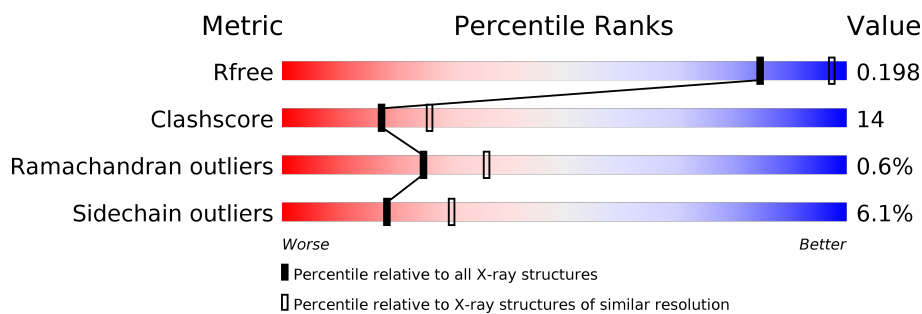
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.40 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	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)

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 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  $\leq 5\%$ .

Mol	Chain	Length	Quality of chain
1	A	248	
1	B	248	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 4184 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 RIBOSOME-INACTIVATING PROTEIN ALPHA-TRICHOSANTHIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	247	Total	C	N	O	S	0	0	0
			1896	1202	320	369	5			
1	B	248	Total	C	N	O	S	0	0	0
			1907	1208	324	370	5			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	173	ALA	LYS	engineered mutation	UNP P09989
A	174	ALA	ARG	engineered mutation	UNP P09989
A	177	ALA	LYS	engineered mutation	UNP P09989
B	173	ALA	LYS	engineered mutation	UNP P09989
B	174	ALA	ARG	engineered mutation	UNP P09989
B	177	ALA	LYS	engineered mutation	UNP P09989

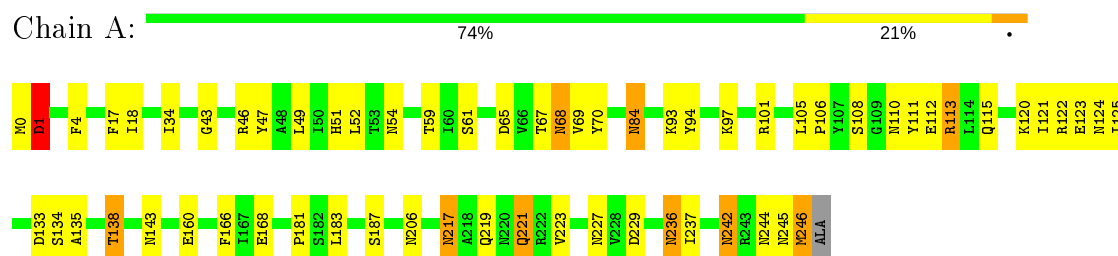
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	215	Total	O	0	0
			215	215		
2	B	166	Total	O	0	0
			166	166		

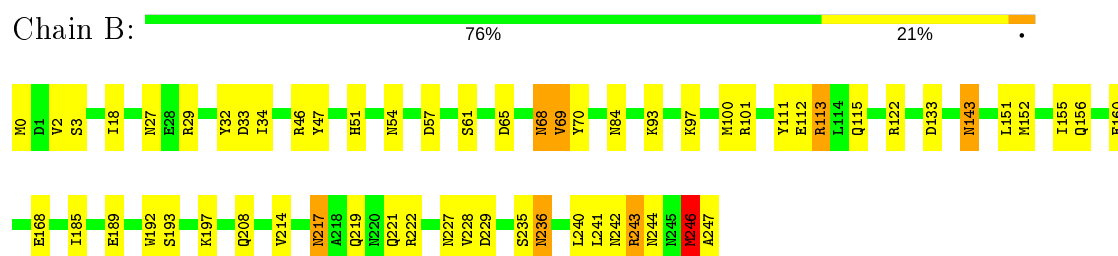
### 3 Residue-property plots [i](#)

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: RIBOSOME-INACTIVATING PROTEIN ALPHA-TRICHOSANTHIN



#### • Molecule 1: RIBOSOME-INACTIVATING PROTEIN ALPHA-TRICHOSANTHIN



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	70.03Å 95.84Å 101.83Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 2.40 30.54 – 1.90	Depositor EDS
% Data completeness (in resolution range)	99.0 (50.00-2.40) 98.8 (30.54-1.90)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.60 (at 1.89Å)	Xtriage
Refinement program	REFMAC 5.0	Depositor
R, $R_{free}$	0.191 , 0.236 0.195 , 0.198	Depositor DCC
$R_{free}$ test set	2753 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	19.0	Xtriage
Anisotropy	0.135	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 58.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	4184	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	21.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.50% 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

### 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	9.22	5/1928 (0.3%)	1.22	5/2623 (0.2%)
1	B	0.66	7/1939 (0.4%)	0.68	4/2637 (0.2%)
All	All	6.53	12/3867 (0.3%)	0.99	9/5260 (0.2%)

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	2	1
1	B	1	0
All	All	3	1

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	246	MET	C-O	351.49	7.91	1.23
1	A	246	MET	CA-C	200.03	6.73	1.52
1	B	246	MET	C-O	-13.86	0.97	1.23
1	B	0	MET	C-O	-8.40	1.07	1.23
1	B	0	MET	N-CA	-8.17	1.30	1.46

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	246	MET	CA-C-O	-50.85	13.32	120.10
1	A	246	MET	N-CA-C	-13.62	74.23	111.00
1	B	0	MET	CG-SD-CE	9.59	115.55	100.20
1	B	246	MET	CB-CG-SD	-7.30	90.49	112.40
1	A	1	ASP	O-C-N	-6.41	112.44	122.70

All (3) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	1	ASP	CA
1	A	246	MET	CA
1	B	246	MET	CA

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1	ASP	Mainchain

## 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	1896	0	1901	59	0
1	B	1907	0	1917	45	0
2	A	215	0	0	0	0
2	B	166	0	0	2	0
All	All	4184	0	3818	104	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 14.

The worst 5 of 104 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:A:68:ASN:HB3	1:A:246:MET:C	1.27	1.49
1:A:245:ASN:ND2	1:A:246:MET:C	1.84	1.28
1:A:246:MET:HA	1:A:246:MET:O	1.39	1.22
1:A:68:ASN:CB	1:A:246:MET:C	2.12	1.15
1:A:120:LYS:HD3	1:A:125:ILE:HD11	1.37	1.07

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	245/248 (99%)	235 (96%)	10 (4%)	0	100	100
1	B	246/248 (99%)	238 (97%)	5 (2%)	3 (1%)	13	19
All	All	491/496 (99%)	473 (96%)	15 (3%)	3 (1%)	25	36

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	246	MET
1	B	29	ARG
1	B	69	VAL

### 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	205/206 (100%)	192 (94%)	13 (6%)	18	28
1	B	206/206 (100%)	194 (94%)	12 (6%)	20	32
All	All	411/412 (100%)	386 (94%)	25 (6%)	18	30

5 of 25 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	236	ASN
1	B	57	ASP
1	B	243	ARG

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	A	242	ASN
1	B	68	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 36 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	244	ASN
1	B	51	HIS
1	B	236	ASN
1	B	20	ASN
1	B	68	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](#)

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.