Full wwPDB NMR Structure Validation Report (i)

Jun 13, 2024 – 09:28 AM EDT

PDB ID	:	1MIT
Title	:	RECOMBINANT CUCURBITA MAXIMA TRYPSIN INHIBITOR V (RCMTI-V) (NMR, MINIMIZED AVERAGE STRUCTURE)
Authors	:	Cai, M.; Gong, Y.; Huang, Y.; Liu, J.; Prakash, O.; Wen, L.; Wen, J.J.; Huang, J.-K.; Krishnamoorthi, R.
Deposited on	:	1995-10-26

This is a Full wwPDB NMR Structure Validation 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/NMRValidationReportHelp
with specific help available everywhere you see the (i) 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 (i)) were used in the production of this report:

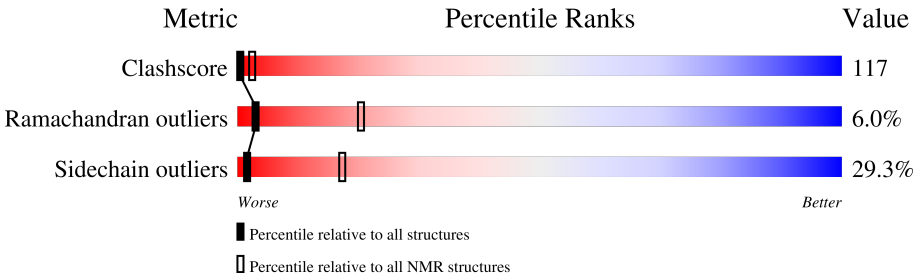
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
SOLUTION NMR

The overall completeness of chemical shifts assignment was not calculated.

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)	NMR archive (#Entries)
Clashscore	158937	12864
Ramachandran outliers	154571	11451
Sidechain outliers	154315	11428

The table below summarises the geometric issues observed across the polymeric chains and their fit to the experimental data. The red, orange, yellow and green segments indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. A cyan segment indicates the fraction of residues that are not part of the well-defined cores, and 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 <=5%

Mol	Chain	Length	Quality of chain
1	A	69	13% 67% 16% •

2 Ensemble composition and analysis (i)

This entry contains 1 models. Identification of well-defined residues and clustering analysis are not possible.

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1068 atoms, of which 546 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called TRYPSIN INHIBITOR V.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	69	1068	328	546	101	91	2	0

4 Residue-property plots (i)

These plots are provided for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic is the same as shown in the summary in section 1 of this report. The second graphic shows the sequence where residues are colour-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 outliers are shown as green connectors. Residues which are classified as ill-defined in the NMR ensemble, are shown in cyan with an underline colour-coded according to the previous scheme. Residues which were present in the experimental sample, but not modelled in the final structure are shown in grey.

- Molecule 1: TRYPSIN INHIBITOR V

Chain A: 13% 67% 16% •

G1A S1 S2 C3 P4 G5 K6 S7 S8 W9 P10 H11 L12 V13 G14 V15 G16 G17 S18 V19 A20 K21 A22 I23 I24 E25 R26 Q27 N28 V31 K32 A33 V34 I35 L36 E37 E38 G39 T40 P41 V42 T43 K44 D45 F46 R47 C48 N49 R50 V51 R52 I53 W54 V55 N56 K57 R58 G59 L60

V61 V62 S63 P64 P65 R66 I67 G68

5 Refinement protocol and experimental data overview (i)

Of the ? calculated structures, 1 were deposited, based on the following criterion: *?*.

The authors did not provide any information on software used for structure solution, optimization or refinement.

No chemical shift data was provided.

6 Model quality (i)

6.1 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 (average) 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	1.11	3/533 (0.6%)	1.53	8/721 (1.1%)
All	All	1.11	3/533 (0.6%)	1.53	8/721 (1.1%)

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	Planarity
1	A	0	3
All	All	0	3

All bond outliers are listed below. They are sorted according to the Z-score.

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	54	TRP	CG-CD2	-6.01	1.33	1.43
1	A	9	TRP	CG-CD2	-5.95	1.33	1.43
1	A	11	HIS	CG-ND1	-5.23	1.27	1.38

All angle outliers are listed below. They are sorted according to the Z-score.

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	9	TRP	NE1-CE2-CZ2	8.10	139.31	130.40
1	A	54	TRP	NE1-CE2-CZ2	7.82	139.00	130.40
1	A	54	TRP	CD1-NE1-CE2	6.76	115.09	109.00
1	A	9	TRP	CD1-NE1-CE2	6.74	115.07	109.00
1	A	54	TRP	CG-CD1-NE1	-6.55	103.55	110.10
1	A	9	TRP	NE1-CE2-CD2	-6.47	100.83	107.30
1	A	9	TRP	CG-CD1-NE1	-6.45	103.65	110.10
1	A	54	TRP	NE1-CE2-CD2	-6.35	100.95	107.30

There are no chirality outliers.

All planar outliers are listed below.

Mol	Chain	Res	Type	Group
1	A	47	ARG	Sidechain
1	A	52	ARG	Sidechain
1	A	66	ARG	Sidechain

6.2 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 each 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 averaged over the ensemble.

Mol	Chain	Non-H	H(model)	H(added)	Clashes
1	A	522	546	546	125
All	All	522	546	546	125

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 117.

All clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)
1:A:21:LYS:NZ	1:A:33:ALA:HB3	0.94	1.78
1:A:21:LYS:HZ3	1:A:33:ALA:HB3	0.91	1.25
1:A:6:LYS:HD2	1:A:27:GLN:HB3	0.81	1.52
1:A:42:VAL:HG21	1:A:52:ARG:NH2	0.79	1.92
1:A:17:GLY:N	1:A:61:VAL:HG21	0.79	1.93
1:A:53:ILE:HG23	1:A:63:SER:HB3	0.79	1.52
1:A:5:GLY:HA3	1:A:28:ASN:ND2	0.77	1.93
1:A:37:GLU:HA	1:A:55:VAL:HG22	0.77	1.56
1:A:3:CYS:HB3	1:A:4:PRO:HD2	0.76	1.58
1:A:21:LYS:HG3	1:A:33:ALA:HB3	0.75	1.58
1:A:56:ASN:ND2	1:A:58:ARG:HB2	0.74	1.96
1:A:24:ILE:HG21	1:A:51:VAL:HB	0.72	1.60
1:A:9:TRP:CH2	1:A:51:VAL:HG21	0.70	2.20
1:A:28:ASN:HB3	1:A:31:VAL:HG22	0.70	1.63
1:A:5:GLY:HA3	1:A:28:ASN:HD21	0.70	1.44
1:A:9:TRP:O	1:A:65:PRO:HB3	0.70	1.87
1:A:12:LEU:O	1:A:15:VAL:HG23	0.70	1.86
1:A:13:VAL:HA	1:A:61:VAL:O	0.69	1.88
1:A:35:ILE:C	1:A:36:LEU:HD23	0.69	2.07
1:A:28:ASN:HB3	1:A:31:VAL:CG2	0.68	2.19

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)
1:A:27:GLN:HA	1:A:27:GLN:OE1	0.67	1.89
1:A:42:VAL:HG21	1:A:52:ARG:HH22	0.67	1.46
1:A:6:LYS:HE3	1:A:27:GLN:NE2	0.67	2.03
1:A:35:ILE:O	1:A:36:LEU:HD23	0.67	1.90
1:A:28:ASN:O	1:A:31:VAL:HG22	0.66	1.90
1:A:24:ILE:CD1	1:A:53:ILE:HD11	0.66	2.21
1:A:6:LYS:CD	1:A:27:GLN:HB3	0.65	2.19
1:A:34:VAL:HB	1:A:52:ARG:HD2	0.65	1.67
1:A:34:VAL:CB	1:A:52:ARG:HD2	0.64	2.22
1:A:63:SER:O	1:A:65:PRO:HD3	0.64	1.93
1:A:53:ILE:HG23	1:A:63:SER:CB	0.64	2.23
1:A:37:GLU:HG2	1:A:55:VAL:CG2	0.63	2.22
1:A:21:LYS:HG3	1:A:33:ALA:CB	0.63	2.22
1:A:24:ILE:HG12	1:A:51:VAL:HG11	0.62	1.70
1:A:40:THR:O	1:A:42:VAL:HG22	0.61	1.95
1:A:53:ILE:CG2	1:A:63:SER:HB3	0.61	2.26
1:A:19:VAL:O	1:A:23:ILE:HG13	0.60	1.95
1:A:24:ILE:HG22	1:A:31:VAL:HG23	0.60	1.73
1:A:12:LEU:C	1:A:15:VAL:HG23	0.60	2.17
1:A:9:TRP:CH2	1:A:51:VAL:CG2	0.60	2.85
1:A:62:VAL:O	1:A:64:PRO:HD3	0.60	1.97
1:A:24:ILE:CG2	1:A:51:VAL:HB	0.59	2.28
1:A:36:LEU:O	1:A:54:TRP:HA	0.58	1.98
1:A:3:CYS:HB3	1:A:4:PRO:CD	0.58	2.26
1:A:12:LEU:HA	1:A:15:VAL:HG21	0.58	1.76
1:A:49:ASN:ND2	1:A:50:ARG:HG3	0.58	2.13
1:A:21:LYS:HZ3	1:A:33:ALA:CB	0.56	2.08
1:A:42:VAL:CG2	1:A:52:ARG:NH2	0.56	2.68
1:A:42:VAL:CG2	1:A:52:ARG:HH22	0.56	2.14
1:A:34:VAL:O	1:A:53:ILE:HD12	0.55	2.01
1:A:28:ASN:ND2	1:A:31:VAL:CG1	0.55	2.70
1:A:54:TRP:O	1:A:62:VAL:N	0.55	2.40
1:A:24:ILE:HD11	1:A:53:ILE:HG12	0.54	1.77
1:A:28:ASN:ND2	1:A:31:VAL:HG11	0.54	2.17
1:A:34:VAL:HG12	1:A:36:LEU:HD21	0.54	1.79
1:A:37:GLU:O	1:A:54:TRP:CZ3	0.54	2.61
1:A:47:ARG:HB2	1:A:49:ASN:ND2	0.53	2.18
1:A:54:TRP:O	1:A:62:VAL:HG23	0.53	2.04
1:A:40:THR:H	1:A:54:TRP:HZ3	0.52	1.47
1:A:9:TRP:CZ2	1:A:51:VAL:HG21	0.52	2.39
1:A:55:VAL:HG23	1:A:55:VAL:O	0.52	2.04
1:A:23:ILE:HG22	1:A:27:GLN:HG3	0.52	1.81

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)
1:A:5:GLY:CA	1:A:28:ASN:ND2	0.52	2.70
1:A:24:ILE:CD1	1:A:51:VAL:CG1	0.52	2.88
1:A:24:ILE:CD1	1:A:53:ILE:CD1	0.52	2.88
1:A:53:ILE:CG2	1:A:63:SER:CB	0.51	2.88
1:A:56:ASN:OD1	1:A:60:LEU:HB2	0.51	2.04
1:A:16:GLY:C	1:A:61:VAL:HG21	0.51	2.26
1:A:24:ILE:HD12	1:A:53:ILE:HD11	0.50	1.83
1:A:24:ILE:HB	1:A:33:ALA:HB2	0.50	1.83
1:A:24:ILE:CG1	1:A:51:VAL:HG11	0.50	2.36
1:A:13:VAL:HG22	1:A:62:VAL:HA	0.49	1.83
1:A:23:ILE:O	1:A:27:GLN:CB	0.48	2.61
1:A:12:LEU:HD22	1:A:23:ILE:HD12	0.48	1.85
1:A:37:GLU:O	1:A:39:GLY:N	0.48	2.46
1:A:55:VAL:CG2	1:A:55:VAL:O	0.48	2.62
1:A:24:ILE:CG2	1:A:31:VAL:HG23	0.48	2.38
1:A:36:LEU:CD1	1:A:42:VAL:HG11	0.48	2.39
1:A:47:ARG:CB	1:A:49:ASN:ND2	0.47	2.77
1:A:51:VAL:O	1:A:51:VAL:HG12	0.47	2.07
1:A:8:SER:HA	1:A:65:PRO:O	0.47	2.08
1:A:34:VAL:O	1:A:53:ILE:CG1	0.47	2.62
1:A:52:ARG:CG	1:A:54:TRP:HE1	0.47	2.22
1:A:8:SER:CA	1:A:65:PRO:O	0.47	2.62
1:A:21:LYS:HA	1:A:33:ALA:HB1	0.47	1.85
1:A:23:ILE:O	1:A:27:GLN:CG	0.47	2.62
1:A:23:ILE:O	1:A:27:GLN:HB2	0.47	2.09
1:A:1:SER:OG	1:A:2:SER:N	0.46	2.48
1:A:28:ASN:CG	1:A:31:VAL:HG13	0.46	2.30
1:A:9:TRP:HB2	1:A:65:PRO:HA	0.46	1.86
1:A:15:VAL:O	1:A:19:VAL:HB	0.46	2.10
1:A:57:LYS:C	1:A:59:GLY:N	0.46	2.69
1:A:54:TRP:C	1:A:62:VAL:HG23	0.46	2.30
1:A:24:ILE:CD1	1:A:51:VAL:HG11	0.46	2.41
1:A:56:ASN:ND2	1:A:58:ARG:CB	0.46	2.75
1:A:64:PRO:C	1:A:66:ARG:H	0.46	2.15
1:A:9:TRP:O	1:A:65:PRO:CB	0.45	2.64
1:A:35:ILE:HG12	1:A:53:ILE:HD12	0.44	1.88
1:A:5:GLY:HA3	1:A:28:ASN:CG	0.44	2.33
1:A:38:GLU:HA	1:A:54:TRP:CE3	0.44	2.47
1:A:54:TRP:O	1:A:62:VAL:CG2	0.43	2.66
1:A:3:CYS:CB	1:A:4:PRO:CD	0.43	2.92
1:A:36:LEU:HD12	1:A:54:TRP:CZ2	0.43	2.48
1:A:23:ILE:O	1:A:27:GLN:HG2	0.43	2.13

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)
1:A:49:ASN:ND2	1:A:49:ASN:C	0.43	2.70
1:A:21:LYS:HA	1:A:24:ILE:HD12	0.43	1.91
1:A:57:LYS:C	1:A:59:GLY:H	0.42	2.17
1:A:23:ILE:O	1:A:27:GLN:N	0.42	2.51
1:A:24:ILE:HD11	1:A:53:ILE:CG1	0.42	2.45
1:A:56:ASN:C	1:A:58:ARG:N	0.42	2.73
1:A:12:LEU:CD2	1:A:12:LEU:N	0.42	2.82
1:A:63:SER:O	1:A:65:PRO:CD	0.42	2.66
1:A:34:VAL:O	1:A:53:ILE:CD1	0.42	2.66
1:A:56:ASN:O	1:A:58:ARG:N	0.42	2.52
1:A:12:LEU:O	1:A:15:VAL:CG2	0.42	2.62
1:A:9:TRP:N	1:A:9:TRP:CE3	0.41	2.88
1:A:12:LEU:C	1:A:15:VAL:CG2	0.41	2.90
1:A:37:GLU:HG2	1:A:55:VAL:HG21	0.41	1.91
1:A:48:CYS:O	1:A:67:ILE:HD13	0.40	2.15
1:A:12:LEU:N	1:A:12:LEU:HD23	0.40	2.31
1:A:34:VAL:HG22	1:A:50:ARG:HH21	0.40	1.76
1:A:56:ASN:HD22	1:A:58:ARG:HG3	0.40	1.76
1:A:35:ILE:HA	1:A:53:ILE:O	0.40	2.16
1:A:37:GLU:O	1:A:54:TRP:HZ3	0.40	1.97
1:A:40:THR:HB	1:A:54:TRP:CH2	0.40	2.51

6.3 Torsion angles (i)

6.3.1 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 PDB entries followed by that with respect to all NMR entries. 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	67/69 (97%)	53 (79%)	10 (15%)	4 (6%)	3 20
All	All	67/69 (97%)	53 (79%)	10 (15%)	4 (6%)	3 20

All 4 Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type
1	A	10	PRO
1	A	38	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	41	PRO
1	A	65	PRO

6.3.2 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 PDB entries followed by that with respect to all NMR entries. 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	58/58 (100%)	41 (71%)	17 (29%)	1 17
All	All	58/58 (100%)	41 (71%)	17 (29%)	1 17

All 17 residues with a non-rotameric sidechain are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type
1	A	7	SER
1	A	12	LEU
1	A	18	SER
1	A	26	ARG
1	A	32	LYS
1	A	42	VAL
1	A	44	LYS
1	A	45	ASP
1	A	46	PHE
1	A	47	ARG
1	A	49	ASN
1	A	52	ARG
1	A	53	ILE
1	A	62	VAL
1	A	63	SER
1	A	66	ARG
1	A	67	ILE

6.3.3 RNA (i)

There are no RNA molecules in this entry.

6.4 Non-standard residues in protein, DNA, RNA chains (i)

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

6.5 Carbohydrates (i)

There are no monosaccharides in this entry.

6.6 Ligand geometry (i)

There are no ligands in this entry.

6.7 Other polymers (i)

There are no such molecules in this entry.

6.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

7 Chemical shift validation (i)

No chemical shift data were provided