



# wwPDB NMR Structure Validation Summary Report ⓘ

Mar 23, 2026 – 08:50 AM UTC

PDB ID : 2CA7 / pdb\_00002ca7  
BMRB ID : 6506  
Title : Konkunitzin-S1 Is The First Member Of A New Kunitz-Type Neurotoxin Family- Structural and Functional Characterization  
Authors : Bayrhuber, M.; Vijayan, V.; Ferber, M.; Graf, R.; Korukottu, J.; Imperial, J.; Garrett, J.E.; Olivera, B.M.; Terlau, H.; Zweckstetter, M.; Becker, S.  
Deposited on : 2005-12-19

This is a wwPDB NMR 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/NMRValidationReportHelp>

with specific help available everywhere you see the ⓘ 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 ⓘ](#)) were used in the production of this report:

MolProbity : 4-5-2 with Phenix2.0  
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)  
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) : 2.49

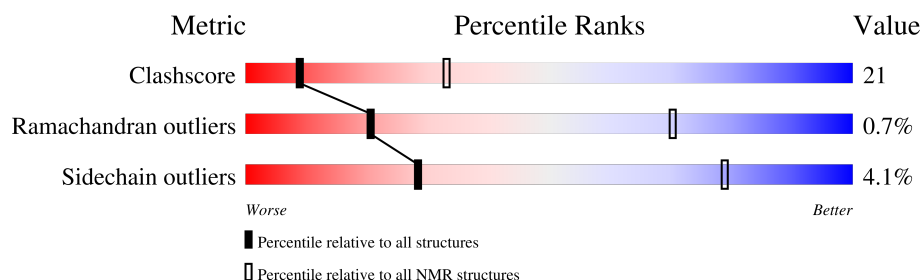
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*SOLUTION NMR*

The overall completeness of chemical shifts assignment is 72%.

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	229148	14424
Ramachandran outliers	224038	12848
Sidechain outliers	223484	12823

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  $\geq 3$ , 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  $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	60	

## 2 Ensemble composition and analysis

This entry contains 20 models. The atoms present in the NMR models are not consistent. Some calculations may have failed as a result. All residues are included in the validation scores. Model 1 is the overall representative, medoid model (most similar to other models).

The following residues are included in the computation of the global validation metrics.

Well-defined (core) protein residues			
Well-defined core	Residue range (total)	Backbone RMSD (Å)	Medoid model
1	A:4-A:60 (57)	0.62	1

Ill-defined regions of proteins are excluded from the global statistics.

Ligands and non-protein polymers are included in the analysis.

The models can be grouped into 3 clusters. No single-model clusters were found.

Cluster number	Models
1	1, 2, 3, 5, 6, 7, 9, 11, 12
2	4, 14, 15, 16, 17, 18, 19
3	8, 10, 13, 20

### 3 Entry composition [i](#)

There is only 1 type of molecule in this entry. The entry contains 785 atoms, of which 300 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called CONKUNITZIN-S1.

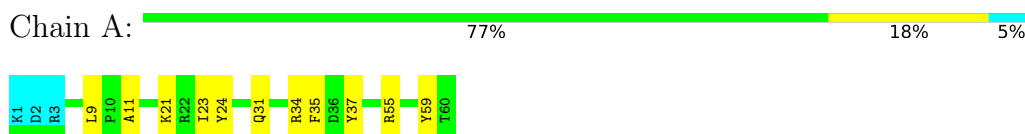
Mol	Chain	Residues	Atoms						Trace
1	A	60	Total	C	H	N	O	S	0
			785	293	300	92	96	4	

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

### 4.1 Average score per residue in the NMR ensemble

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.

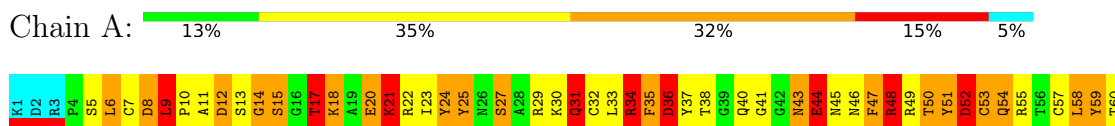
- Molecule 1: CONKUNITZIN-S1



### 4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 1. Colouring as in section 4.1 above.

- Molecule 1: CONKUNITZIN-S1



## 5 Refinement protocol and experimental data overview

The models were refined using the following method: *MOLECULAR DYNAMICS, SIMULATED ANNEALING, TORSION ANGLE DYNAMICS*.

Of the 20 calculated structures, 20 were deposited, based on the following criterion: *structures with the lowest energy*.

The following table shows the software used for structure solution, optimisation and refinement.

Software name	Classification	Version
Xplor-NIH	refinement	2.9.7
XwinNMR	structure solution	3.5
NMRPipe	structure solution	027.12.56
NMRDraw	structure solution	2004.245.17.31
Sparky	structure solution	110
ARIA2alpha	structure solution	
CYANA	structure solution	32
MARS	structure solution	1.0

The following table shows chemical shift validation statistics as aggregates over all chemical shift files. Detailed validation can be found in section 7 of this report.

Chemical shift file(s)	working_cs.cif
Number of chemical shift lists	1
Total number of shifts	566
Number of shifts mapped to atoms	498
Number of unparsed shifts	0
Number of shifts with mapping errors	68
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	72%

## 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	1.95±3.22	9±34/465 ( 1.9± 7.4%)	1.87±2.47	11±37/622 ( 1.8± 5.9%)
All	All	3.76	173/9300 ( 1.9%)	3.10	224/12440 ( 1.8%)

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	0.1±0.4	0.2±0.4
All	All	2	5

5 of 161 unique bond outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)	Models	
								Worst	Total
1	A	48	ARG	CZ-NH2	-72.43	0.39	1.33	1	1
1	A	48	ARG	CD-NE	-69.48	0.48	1.46	1	1
1	A	49	ARG	NE-CZ	-63.90	0.62	1.33	1	1
1	A	49	ARG	CZ-NH1	-56.61	0.53	1.32	1	1
1	A	29	ARG	CZ-NH1	-54.91	0.55	1.32	1	1

5 of 195 unique angle outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)	Models	
								Worst	Total
1	A	54	GLN	OE1-CD-NE2	-85.30	37.30	122.60	1	1
1	A	25	TYR	CD1-CG-CD2	-64.63	21.16	118.10	1	1
1	A	51	TYR	CD1-CG-CD2	-64.63	21.16	118.10	1	1
1	A	35	PHE	CD1-CG-CD2	-64.50	21.86	118.60	1	1
1	A	35	PHE	CE1-CZ-CE2	-54.10	22.63	120.00	1	1

All unique chiral outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Atoms	Models (Total)
1	A	17	THR	CB	1
1	A	50	THR	CB	1

All unique planar outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Group	Models (Total)
1	A	48	ARG	Sidechain	3
1	A	29	ARG	Sidechain	1
1	A	49	ARG	Sidechain	1

## 6.2 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 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	457	426	426	18±48
All	All	9140	8390	8516	367

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 21.

5 of 311 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:59:TYR:CZ	1:A:59:TYR:CD1	1.56	1.93	1	1
1:A:44:GLU:CD	1:A:44:GLU:CB	1.51	1.79	1	1
1:A:59:TYR:CG	1:A:59:TYR:CE2	1.50	1.96	1	1
1:A:34:ARG:NH1	1:A:34:ARG:NE	1.49	1.61	1	1
1:A:24:TYR:CZ	1:A:24:TYR:CD1	1.48	1.99	1	1



## 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	56/60 (93%)	54±1 (96±2%)	2±1 (3±2%)	0±1 (1±1%)	20	70
All	All	1120/1200 (93%)	1078 (96%)	34 (3%)	8 (1%)	20	70

All 3 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	11	ALA	5
1	A	10	PRO	2
1	A	6	LEU	1

### 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	49/52 (94%)	47±3 (96±5%)	2±3 (4±5%)	28	79
All	All	980/1040 (94%)	940 (96%)	40 (4%)	28	79

5 of 20 unique residues with a non-rotameric sidechain are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	21	LYS	4
1	A	23	ILE	4
1	A	9	LEU	3
1	A	31	GLN	3
1	A	34	ARG	3

### 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 oligosaccharides 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](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	1-A	4

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	10:PRO	C	11:ALA	N	1.19
1	A	38:THR	C	39:GLY	N	1.17
1	A	2:ASP	C	3:ARG	N	0.72
1	A	1:LYS	C	2:ASP	N	0.37

## 7 Chemical shift validation

The completeness of assignment taking into account all chemical shift lists is 72% for the well-defined parts and 70% for the entire structure.

### 7.1 Chemical shift list 1

File name: working\_cs.cif

Chemical shift list name: *assigned\_chem\_shift\_list\_1*

#### 7.1.1 Bookkeeping

The following table shows the results of parsing the chemical shift list and reports the number of nuclei with statistically unusual chemical shifts.

Total number of shifts	566
Number of shifts mapped to atoms	498
Number of unparsed shifts	0
Number of shifts with mapping errors	68
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	5

The following assigned chemical shifts were not mapped to the molecules present in the coordinate file.

- No matching atom found in the structure. First 5 (of 68) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	1	LYS	HA	4.009	0.003	1
1	A	1	LYS	HB2	1.901	0.003	2
1	A	1	LYS	HD2	1.689	0.003	2
1	A	1	LYS	HE2	2.991	0.003	2
1	A	1	LYS	HG2	1.455	0.003	2
1	A	2	ASP	H	8.69	0.001	1
1	A	2	ASP	HA	4.651	0.003	1
1	A	2	ASP	HB2	2.515	0.003	2
1	A	2	ASP	HB3	2.694	0.003	2
1	A	6	LEU	HG	1.558	0.003	1
1	A	13	SER	HB2	4.186	0.003	2
1	A	15	SER	HB2	3.627	0.003	2
1	A	15	SER	HB3	4.111	0.003	2
1	A	17	THR	HB	4.492	0.005	4

*Continued on next page...*

*Continued from previous page...*

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	17	THR	HG21	1.121	0.003	4
1	A	17	THR	HG22	1.121	0.003	4
1	A	17	THR	HG23	1.121	0.003	1
1	A	18	LYS	HD2	1.621	0.003	2
1	A	18	LYS	HE2	2.953	0.003	2
1	A	18	LYS	HG2	1.314	0.003	2
1	A	21	LYS	HD2	1.701	0.003	2
1	A	21	LYS	HD3	1.452	0.003	2
1	A	21	LYS	HE2	2.91	0.003	2
1	A	24	TYR	HD1	6.959	0.003	3
1	A	24	TYR	HE1	6.653	0.003	3
1	A	27	SER	HB2	4.082	0.003	2
1	A	29	ARG	HD2	3.033	0.003	2
1	A	30	LYS	HD2	1.667	0.003	2
1	A	30	LYS	HD3	1.701	0.003	2
1	A	30	LYS	HE2	3.018	0.003	2
1	A	31	GLN	HB2	1.378	0.001	2
1	A	31	GLN	HB3	1.753	0.001	2
1	A	31	GLN	HG2	2.246	0.001	2
1	A	33	LEU	HB2	1.385	0.003	2
1	A	33	LEU	HB3	1.667	0.003	2
1	A	33	LEU	HD11	0.778	0.003	2
1	A	33	LEU	HD12	0.778	0.003	2
1	A	33	LEU	HD13	0.778	0.003	2
1	A	33	LEU	HD21	0.821	0.003	2
1	A	33	LEU	HD22	0.821	0.003	2
1	A	33	LEU	HD23	0.821	0.003	2
1	A	33	LEU	HG	1.167	0.003	1
1	A	34	ARG	HD2	3.162	0.003	2
1	A	35	PHE	HD1	7.104	0.003	3
1	A	35	PHE	HE1	7.157	0.003	3
1	A	43	ASN	HB2	3.185	0.003	2
1	A	43	ASN	HB3	2.91	0.003	2
1	A	47	PHE	HD1	7.265	0.003	3
1	A	47	PHE	HE1	7.563	0.003	3
1	A	48	ARG	HD2	3.224	0.003	2
1	A	48	ARG	HG2	1.813	0.005	2
1	A	48	ARG	HG3	1.59	0.004	2
1	A	49	ARG	HD2	3.307	0.003	2
1	A	49	ARG	HG2	1.786	0.003	2
1	A	49	ARG	HG3	1.6	0.003	2

*Continued on next page...*

*Continued from previous page...*

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	51	TYR	HD1	7.134	0.003	3
1	A	51	TYR	HE1	6.831	0.003	3
1	A	54	GLN	HG2	2.042	0.003	2
1	A	54	GLN	HG3	2.501	0.002	2
1	A	58	LEU	HD11	0.867	0.003	2
1	A	58	LEU	HD12	0.867	0.003	2
1	A	58	LEU	HD13	0.867	0.003	2
1	A	58	LEU	HD21	0.739	0.003	2
1	A	58	LEU	HD22	0.739	0.003	2
1	A	58	LEU	HD23	0.739	0.003	2
1	A	58	LEU	HG	1.176	0.003	1
1	A	59	TYR	HD1	6.982	0.003	3
1	A	59	TYR	HE1	6.745	0.003	3

### 7.1.2 Chemical shift referencing ⓘ

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction $\pm$ precision, ppm	Suggested action
$^{13}\text{C}_\alpha$	58	$0.52 \pm 0.22$	Should be applied
$^{13}\text{C}_\beta$	54	$-0.13 \pm 0.45$	None needed ( $< 0.5$ ppm)
$^{13}\text{C}'$	55	$0.41 \pm 0.16$	None needed ( $< 0.5$ ppm)
$^{15}\text{N}$	55	$-0.08 \pm 0.94$	None needed ( $< 0.5$ ppm)

### 7.1.3 Completeness of resonance assignments ⓘ

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 72%, i.e. 544 atoms were assigned a chemical shift out of a possible 756. 0 out of 4 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	$^1\text{H}$	$^{13}\text{C}$	$^{15}\text{N}$
Backbone	274/286 (96%)	112/117 (96%)	109/114 (96%)	53/55 (96%)
Sidechain	259/405 (64%)	170/256 (66%)	89/121 (74%)	0/28 (0%)
Aromatic	11/65 (17%)	10/30 (33%)	1/35 (3%)	0/0 (—%)
Overall	544/756 (72%)	292/403 (72%)	199/270 (74%)	53/83 (64%)

### 7.1.4 Statistically unusual chemical shifts [i](#)

The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

List Id	Chain	Res	Type	Atom	Shift, ppm	Expected range, ppm	Z-score
1	A	24	TYR	CD1	54.99	125.84 – 139.60	-56.5
1	A	11	ALA	HB1	-0.40	0.14 – 2.58	-7.2
1	A	11	ALA	HB2	-0.40	0.14 – 2.58	-7.2
1	A	11	ALA	HB3	-0.40	0.14 – 2.58	-7.2
1	A	53	CYS	HA	1.89	1.97 – 7.35	-5.2

### 7.1.5 Random Coil Index (RCI) plots [i](#)

The image below reports *random coil index* values for the protein chains in the structure. The height of each bar gives a probability of a given residue to be disordered, as predicted from the available chemical shifts and the amino acid sequence. A value above 0.2 is an indication of significant predicted disorder. The colour of the bar shows whether the residue is in the well-defined core (black) or in the ill-defined residue ranges (cyan), as described in section 2 on ensemble composition. If well-defined core and ill-defined regions are not identified then it is shown as gray bars.

Random coil index (RCI) for chain A:

