Full wwPDB NMR Structure Validation Report (i)

Sep 29, 2024 – 01:35 PM EDT

PDB ID	:	2N1N
BMRB ID	:	25568
Title	:	Solution structure of VSTx1
Authors	:	Lau, H.Y.; King, G.F.; Mobli, M.
Deposited on	:	2015-04-12

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:

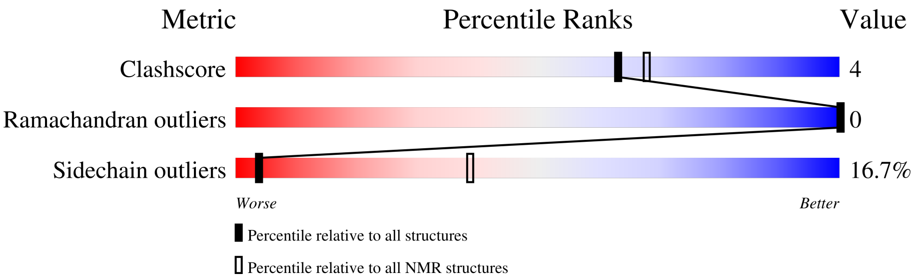
MolProbity	:	4.02b-467
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

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 95%.

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	210492	14027
Ramachandran outliers	207382	12486
Sidechain outliers	206894	12463

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	35	

2 Ensemble composition and analysis (i)

This entry contains 20 models. Model 3 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *lowest energy*.

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:2-A:33 (32)	0.07	3

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 and 1 single-model cluster was found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 546 atoms, of which 265 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Kappa-theraphotoxin-Gr3a.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	35	546	178	265	48	48	7	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	SER	-	expression tag	UNP P60980

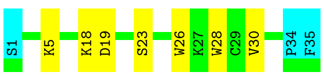
4 Residue-property plots (i)

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: Kappa-theraphotoxin-Gr3a

Chain A:



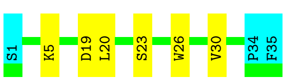
4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1

- Molecule 1: Kappa-theraphotoxin-Gr3a

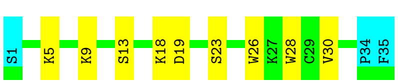
Chain A:



4.2.2 Score per residue for model 2

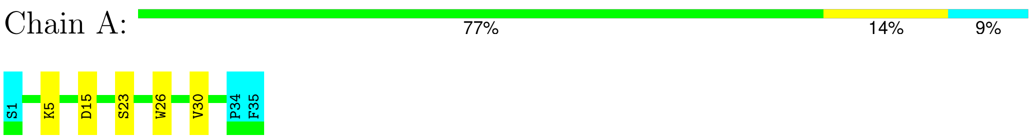
- Molecule 1: Kappa-theraphotoxin-Gr3a

Chain A:



4.2.3 Score per residue for model 3 (medoid)

- Molecule 1: Kappa-theraphotoxin-Gr3a

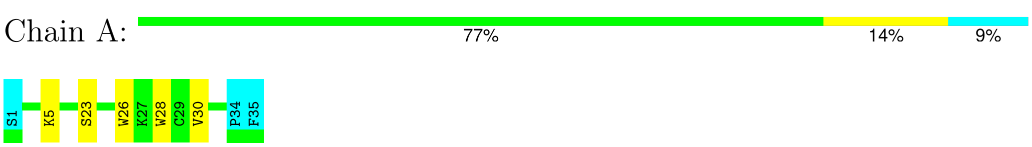


4.2.4 Score per residue for model 4

- Molecule 1: Kappa-theraphotoxin-Gr3a

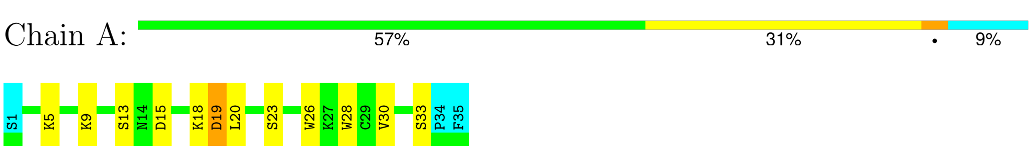
4.2.5 Score per residue for model 5

- Molecule 1: Kappa-theraphotoxin-Gr3a



4.2.6 Score per residue for model 6

- Molecule 1: Kappa-theraphotoxin-Gr3a



4.2.7 Score per residue for model 7

- Molecule 1: Kappa-theraphotoxin-Gr3a

4.2.8 Score per residue for model 8

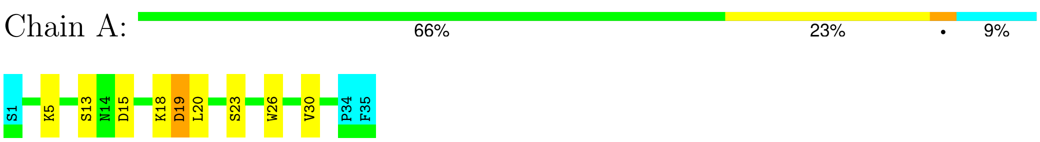
• Molecule 1: Kappa-theraphotoxin-Gr3a

4.2.9 Score per residue for model 9

• Molecule 1: Kappa-theraphotoxin-Gr3a

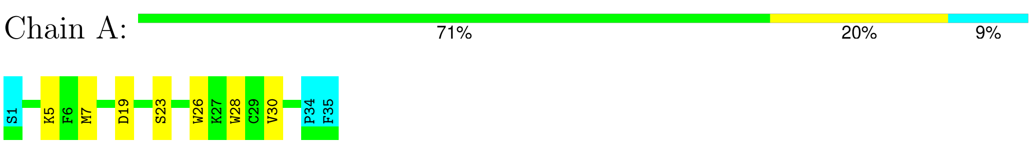
4.2.10 Score per residue for model 10

• Molecule 1: Kappa-theraphotoxin-Gr3a



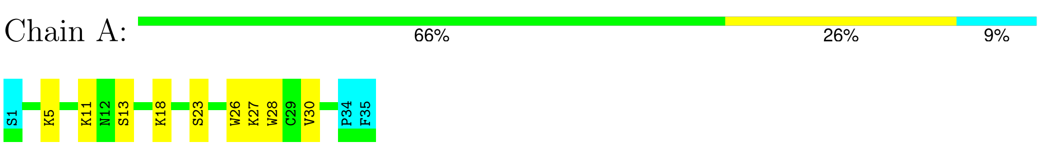
4.2.11 Score per residue for model 11

• Molecule 1: Kappa-theraphotoxin-Gr3a



4.2.12 Score per residue for model 12

• Molecule 1: Kappa-theraphotoxin-Gr3a



4.2.13 Score per residue for model 13

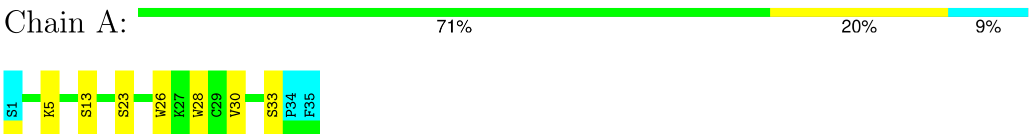
- Molecule 1: Kappa-theraphotoxin-Gr3a

Chain A:

4.2.14 Score per residue for model 14

- Molecule 1: Kappa-theraphotoxin-Gr3a

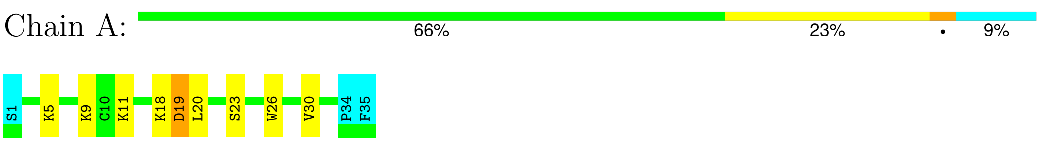
Chain A:



4.2.15 Score per residue for model 15

- Molecule 1: Kappa-theraphotoxin-Gr3a

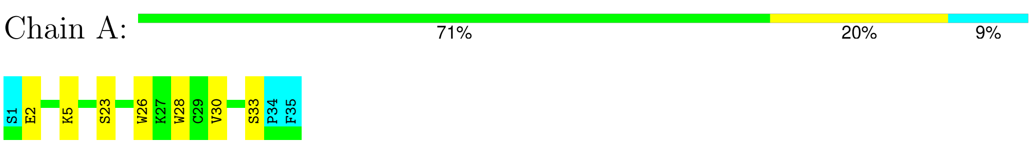
Chain A:



4.2.16 Score per residue for model 16

- Molecule 1: Kappa-theraphotoxin-Gr3a

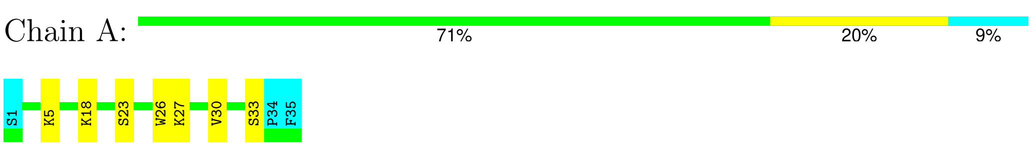
Chain A:



4.2.17 Score per residue for model 17

- Molecule 1: Kappa-theraphotoxin-Gr3a

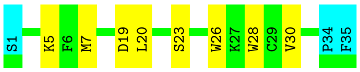
Chain A:



4.2.18 Score per residue for model 18

- Molecule 1: Kappa-theraphotoxin-Gr3a

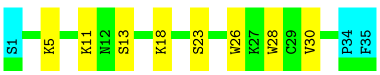
Chain A:

4.2.19 Score per residue for model 19

- Molecule 1: Kappa-theraphotoxin-Gr3a

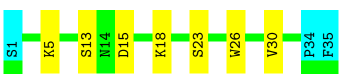
Chain A:

4.2.20 Score per residue for model 20

- Molecule 1: Kappa-theraphotoxin-Gr3a

Chain A:

5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *torsion angle dynamics*.

Of the 200 calculated structures, 20 were deposited, based on the following criterion: *target function*.

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

Software name	Classification	Version
CYANA	structure solution	
CYANA	refinement	

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	433
Number of shifts mapped to atoms	433
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	95%

6 Model quality (i)

6.1 Standard geometry (i)

There are no covalent bond-length or bond-angle outliers.

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	257	244	244	2±1
All	All	5140	4880	4880	40

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 4.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:26:TRP:CD1	1:A:30:VAL:HG11	0.55	2.36	19	20
1:A:19:ASP:O	1:A:20:LEU:HD12	0.41	2.15	18	8
1:A:26:TRP:HB3	1:A:28:TRP:CE2	0.40	2.51	11	12

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	32/35 (91%)	29±1 (90±2%)	3±1 (10±2%)	0±0 (0±0%)	100	100
All	All	640/700 (91%)	573 (90%)	67 (10%)	0 (0%)	100	100

There are no Ramachandran outliers.

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	30/33 (91%)	25±2 (83±5%)	5±2 (17±5%)	4	39
All	All	600/660 (91%)	500 (83%)	100 (17%)	4	39

All 12 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	5	LYS	20
1	A	23	SER	20
1	A	18	LYS	12
1	A	13	SER	9
1	A	33	SER	8
1	A	9	LYS	6
1	A	19	ASP	6
1	A	7	MET	5
1	A	15	ASP	4
1	A	2	GLU	4
1	A	11	LYS	4
1	A	27	LYS	2

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)

There are no chain breaks in this entry.

7 Chemical shift validation (i)

The completeness of assignment taking into account all chemical shift lists is 95% for the well-defined parts and 94% 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 (i)

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	433
Number of shifts mapped to atoms	433
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	6

7.1.2 Chemical shift referencing (i)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	35	0.36 ± 0.49	None needed (< 0.5 ppm)
$^{13}C_{\beta}$	34	0.28 ± 0.25	None needed (< 0.5 ppm)
$^{13}C'$	32	2.19 ± 0.35	Should be applied
^{15}N	33	1.49 ± 0.71	Should be applied

7.1.3 Completeness of resonance assignments (i)

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 95%, i.e. 400 atoms were assigned a chemical shift out of a possible 421. 0 out of 4 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	159/161 (99%)	65/65 (100%)	62/64 (97%)	32/32 (100%)
Sidechain	197/214 (92%)	133/137 (97%)	61/67 (91%)	3/10 (30%)

Continued on next page...

Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Aromatic	44/46 (96%)	23/23 (100%)	18/20 (90%)	3/3 (100%)
Overall	400/421 (95%)	221/225 (98%)	141/151 (93%)	38/45 (84%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 94%, i.e. 433 atoms were assigned a chemical shift out of a possible 459. 0 out of 4 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	169/174 (97%)	69/70 (99%)	67/70 (96%)	33/34 (97%)
Sidechain	212/229 (93%)	143/147 (97%)	66/72 (92%)	3/10 (30%)
Aromatic	52/56 (93%)	28/28 (100%)	21/25 (84%)	3/3 (100%)
Overall	433/459 (94%)	240/245 (98%)	154/167 (92%)	39/47 (83%)

7.1.4 Statistically unusual chemical shifts (i)

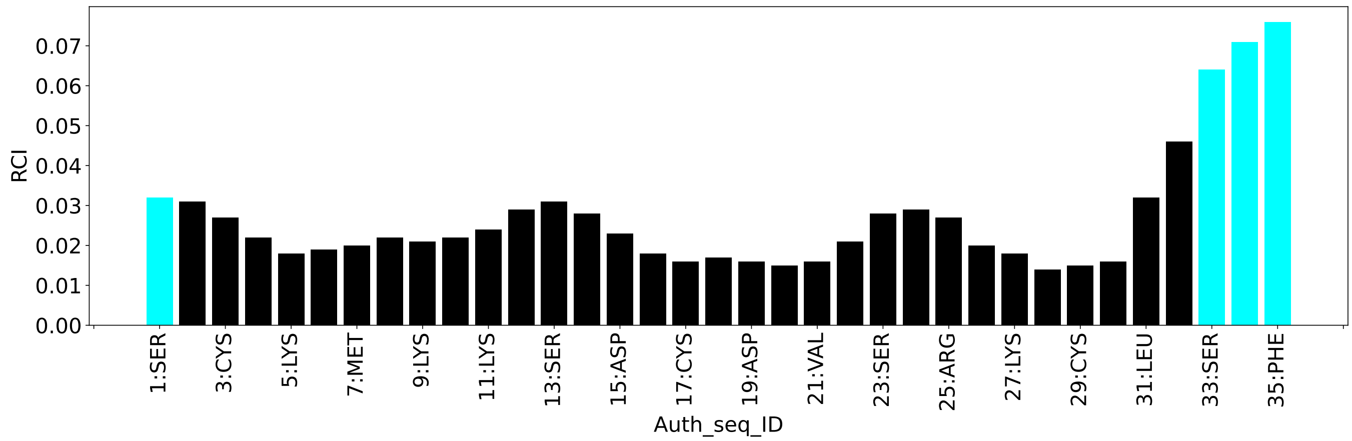
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

List Id	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	25	ARG	NE	110.88	76.53 – 92.65	16.3
1	A	7	MET	HG2	0.29	0.65 – 4.19	-6.0
1	A	26	TRP	HB3	1.02	1.31 – 4.93	-5.8
1	A	26	TRP	HB2	1.31	1.51 – 4.87	-5.6
1	A	5	LYS	HB2	0.51	0.58 – 2.97	-5.3
1	A	25	ARG	HG2	0.21	0.26 – 2.87	-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:


RCI
0.00
0.01
0.02
0.03
0.04
0.05
0.06
0.07
1:SER
3:CYS
5:LYS
7:MET
9:LYS
11:LYS
13:SER
15:ASP
17:CYS
19:ASP
21:VAL
23:SER
25:ARG
27:LYS
29:CYS
31:LEU
33:SER
35:PHE
Auth_seq_ID