WORLDWIDE wwPDB PROTEIN DATA BANK

Full wwPDB NMR Structure Validation Report (i)

Feb 19, 2022 – 10:36 AM EST

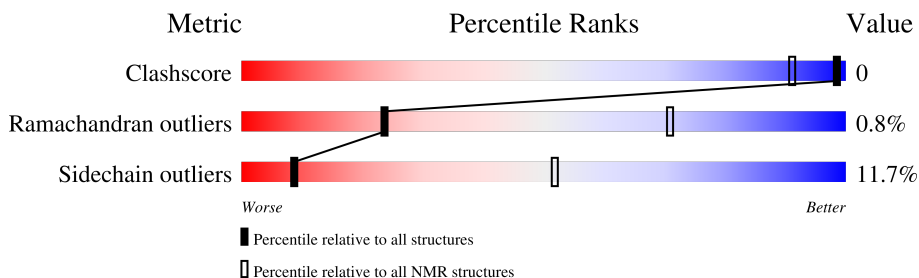|		
PDB ID	:	1QUZ
Title	:	Solution structure of the potassium channel scorpion toxin HSTX1
Authors	:	Savarin, P.; Romi-Lebrun, R.; Zinn-Justin, S.; Lebrun, B.; Nakajima, T.; Gilquin, B.; Menez, A.
Deposited on	:	1999-07-05

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 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	Percentile Ranks	Value
Clashscore		0
Ramachandran outliers		0.8%
Sidechain outliers		11.7%

Worse   *Better*

▮ Percentile relative to all structures

⧍ Percentile relative to all NMR structures

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	35	74% 17% • 6%

2 Ensemble composition and analysis ⓘ

This entry contains 20 models. Model 16 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:34 (33)	0.46	16

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 2 clusters and 3 single-model clusters were found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 511 atoms, of which 253 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called HSTX1 TOXIN.

Mol	Chain	Residues	Atoms						Trace
1	A	35	Total	C	H	N	O	S	1
			511	149	253	54	46	9	

4 Residue-property plots ⓘ

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.
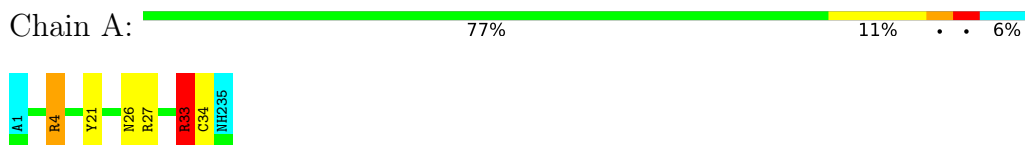
• Molecule 1: HSTX1 TOXIN

Chain A:

74%   17%   ·   6%
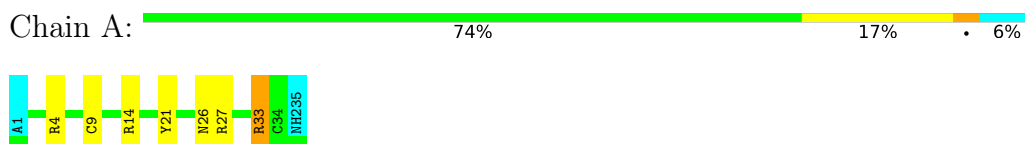
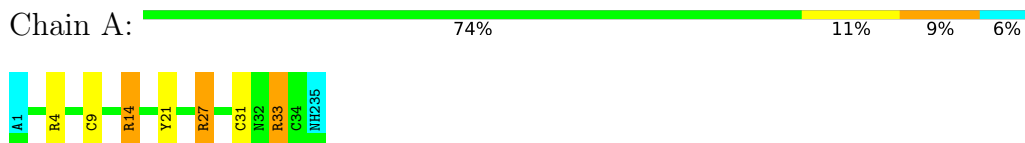## 4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1

• Molecule 1: HSTX1 TOXIN

Chain A:

77%   11%   ·   ·   6%
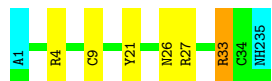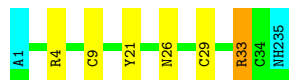
4.2.2 Score per residue for model 2

• Molecule 1: HSTX1 TOXIN

Chain A:

74%   11%   9%   6%

4.2.3 Score per residue for model 3

• Molecule 1: HSTX1 TOXIN

Chain A: 77%   14%   •   6%

A1 R4 C9 Y21 N26 R27 R33 C34 NH235

4.2.4 Score per residue for model 4

• Molecule 1: HSTX1 TOXIN

Chain A: 77%   14%   •   6%

A1 R4 C9 Y21 N26 C29 R33 C34 NH235

4.2.5 Score per residue for model 5

• Molecule 1: HSTX1 TOXIN

Chain A: 74%   17%   •   6%

A1 R4 C9 R14 Y21 N26 R27 R33 C34 NH235

4.2.6 Score per residue for model 6

• Molecule 1: HSTX1 TOXIN

Chain A: 86%   6%   •   6%

A1 R27 K28 R33 C34 NH235

4.2.7 Score per residue for model 7

• Molecule 1: HSTX1 TOXIN

Chain A: 77%   11%   6%   6%

A1 R4 R14 G18 N26 K30 R33 C34 NH235

4.2.8 Score per residue for model 8

• Molecule 1: HSTX1 TOXIN

Chain A: 83%   6%   6%   6%

A1 R4 Y21 R27 R33 C34 NH235

4.2.9 Score per residue for model 9

• Molecule 1: HSTX1 TOXIN

Chain A: 83%   6%   6%   6%

A1 C9 N26 R27 R33 C34 NH235
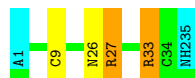
4.2.10 Score per residue for model 10

• Molecule 1: HSTX1 TOXIN

Chain A: 77%   11%   6%   6%

A1 C9 R14 G18 N26 R27 R33 C34 NH235

4.2.11 Score per residue for model 11

• Molecule 1: HSTX1 TOXIN

Chain A: 74%   14%   •   •   6%

A1 R4 R14 G18 Y21 R27 K28 C29 R33 C34 NH235

4.2.12 Score per residue for model 12

• Molecule 1: HSTX1 TOXIN

Chain A: 77%   17%   6%

A1 R4 C9 R14 C19 R27 R33 C34 NH235
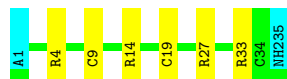
4.2.13 Score per residue for model 13

• Molecule 1: HSTX1 TOXIN

Chain A:

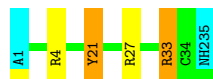#### 4.2.14 Score per residue for model 14

• Molecule 1: HSTX1 TOXIN

Chain A:

4.2.15 Score per residue for model 15

• Molecule 1: HSTX1 TOXIN

Chain A:

4.2.16 Score per residue for model 16 (medoid)

• Molecule 1: HSTX1 TOXIN
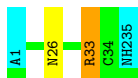
Chain A:

4.2.17 Score per residue for model 17

• Molecule 1: HSTX1 TOXIN

Chain A:

4.2.18 Score per residue for model 18

• Molecule 1: HSTX1 TOXIN

Chain A:

74%   17%   •   6%

A1 R4 K7 D8 C9 N26 C29 K30 C31 N32 R33 C34 NH235

4.2.19 Score per residue for model 19

• Molecule 1: HSTX1 TOXIN

Chain A:

69%   20%   6%   6%

A1 R4 K7 D8 C9 R14 Y21 R27 K30 C31 N32 R33 C34 NH235
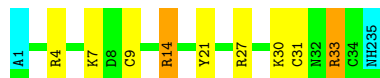
4.2.20 Score per residue for model 20

• Molecule 1: HSTX1 TOXIN

Chain A:

83%   9%   •   6%

A1 C9 Y21 R27 R33 C34 NH235

5 Refinement protocol and experimental data overview ⓘ

The models were refined using the following method: *A COMPLETLY EXTENDED CONFOR-MATION WAS USED AS THE STARTING STRUCTURE. 25 ITERATIONS WERE MADE. 500 STRUCTURES WERE CALCULATED. RESTRAINED MOLECULAR DYNAMICS AT 600K A SLOW COOLING. MINIMIZATION WERE MADE WITH AN ELECTROSTATIC TERM TOPALLH22X.PRO AND PARALLH22X.PRO.*

Of the 20 calculated structures, 20 were deposited, based on the following criterion: *all calculated structures submitted.*

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

Software name	Classification	Version
X-PLOR	structure solution	3.1
X-PLOR	refinement	3.1

No chemical shift data was provided.

6 Model quality ⓘ

6.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: NH2

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	0.95±0.03	0±0/255 (0.0± 0.0%)	1.61±0.14	5±2/337 (1.4± 0.5%)
All	All	0.95	0/5100 (0.0%)	1.62	92/6740 (1.4%)

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	0.0±0.0	0.7±0.7
All	All	0	13

There are no bond-length outliers.

All 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(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	A	27	ARG	NE-CZ-NH1	13.50	127.05	120.30	9	11
1	A	27	ARG	NE-CZ-NH2	-11.33	114.63	120.30	9	6
1	A	21	TYR	CB-CG-CD1	11.30	127.78	121.00	5	9
1	A	21	TYR	CB-CG-CD2	-11.29	114.22	121.00	3	8
1	A	4	ARG	NE-CZ-NH2	-11.02	114.79	120.30	3	7
1	A	4	ARG	NE-CZ-NH1	10.43	125.52	120.30	17	14
1	A	33	ARG	NE-CZ-NH1	10.27	125.43	120.30	5	19
1	A	33	ARG	NE-CZ-NH2	-9.21	115.69	120.30	14	9
1	A	14	ARG	NE-CZ-NH2	-8.86	115.87	120.30	19	4
1	A	14	ARG	NE-CZ-NH1	7.28	123.94	120.30	12	3
1	A	4	ARG	NH1-CZ-NH2	-5.83	112.99	119.40	15	1
1	A	29	CYS	CB-CA-C	-5.60	99.20	110.40	17	1

There are no chirality outliers.

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	33	ARG	Sidechain	5
1	A	14	ARG	Sidechain	5
1	A	27	ARG	Sidechain	3

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	252	244	244	0±0
All	All	5040	4880	4880	1

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 0.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models Worst	Total
1:A:33:ARG:HE	1:A:33:ARG:HA	0.43	1.73	14	1

6.3   Torsion angles ⓘ

6.3.1   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 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	33/35 (94%)	28±2 (86±5%)	4±1 (13±5%)	0±0 (1±1%)	24	71
All	All	660/700 (94%)	566 (86%)	89 (13%)	5 (1%)	24	71

All 2 unique Ramachandran outliers are listed below. They are sorted by the frequency of occur-

rence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	18	GLY	3
1	A	26	ASN	2

6.3.2 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 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	30/30 (100%)	26±1 (88±4%)	4±1 (12±4%)	9	52
All	All	600/600 (100%)	530 (88%)	70 (12%)	9	52

All 16 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	33	ARG	19
1	A	9	CYS	12
1	A	26	ASN	11
1	A	31	CYS	4
1	A	27	ARG	3
1	A	29	CYS	3
1	A	19	CYS	3
1	A	34	CYS	2
1	A	28	LYS	2
1	A	30	LYS	2
1	A	21	TYR	2
1	A	13	CYS	2
1	A	7	LYS	2
1	A	4	ARG	1
1	A	11	ASP	1
1	A	14	ARG	1

6.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

6.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.6 Ligand geometry ⓘ

There are no ligands in this entry.

6.7 Other polymers ⓘ

There are no such molecules in this entry.

6.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

7 Chemical shift validation ⓘ

No chemical shift data were provided