Full wwPDB NMR Structure Validation Report ⓘ

Nov 6, 2023 – 08:20 PM EST

PDB ID	:	2KBU
Title	:	NMR solution structure of Pin1 WW domain mutant with beta turn mimic at position 12
Authors	:	Fuller, A.A.; Bhabha, G.; Case, D.A.
Deposited on	:	2008-12-08

The following versions of software and data (see references ⓘ) were used in the production of this report:

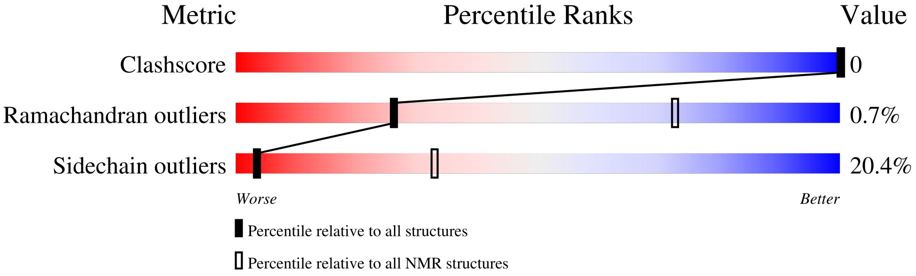
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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	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	31	

2 Ensemble composition and analysis ⓘ

This entry contains 15 models. Model 1 is the overall representative, medoid model (most similar to other models). The authors have identified model 001 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:6-A:11, A:13-A:26 (20)	0.51	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	2, 5, 6, 8, 9, 15
2	1, 10, 11, 13, 14
3	3, 4, 7, 12

3 Entry composition ⓘ

There is only 1 type of molecule in this entry. The entry contains 516 atoms, of which 251 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Peptidyl-prolyl cis-trans isomerase NIMA-interacting 1.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	31	516	171	251	48	45	1	0

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	ARG	SEE REMARK 999	UNP Q13526
A	?	-	SER	SEE REMARK 999	UNP Q13526
A	?	-	SER	SEE REMARK 999	UNP Q13526
A	12	CFD	GLY	SEE REMARK 999	UNP Q13526
A	26	PHE	TRP	engineered mutation	UNP Q13526

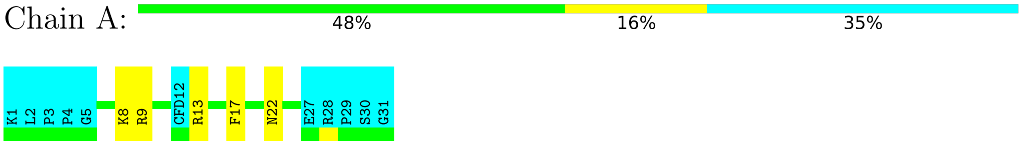
4 Residue-property plots

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: Peptidyl-prolyl cis-trans isomerase NIMA-interacting 1

Chain A:



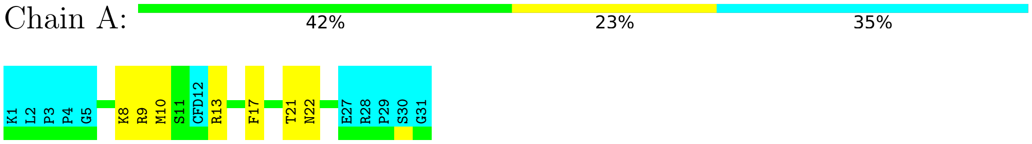
4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1 (medoid)

- Molecule 1: Peptidyl-prolyl cis-trans isomerase NIMA-interacting 1

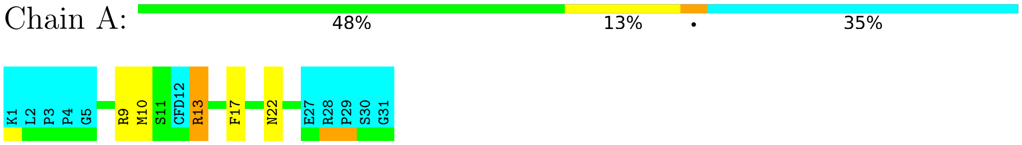
Chain A:



4.2.2 Score per residue for model 2

- Molecule 1: Peptidyl-prolyl cis-trans isomerase NIMA-interacting 1

Chain A:

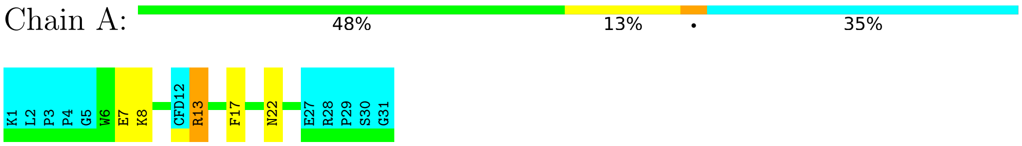


4.2.3 Score per residue for model 3

- Molecule 1: Peptidyl-prolyl cis-trans isomerase NIMA-interacting 1

4.2.4 Score per residue for model 4

- Molecule 1: Peptidyl-prolyl cis-trans isomerase NIMA-interacting 1

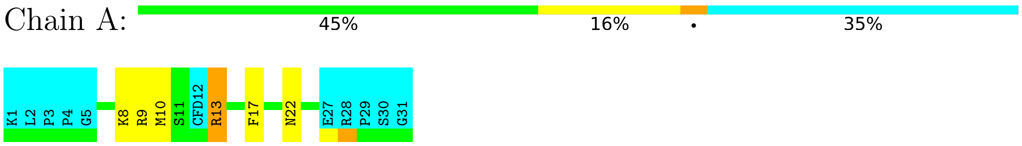


4.2.5 Score per residue for model 5

- Molecule 1: Peptidyl-prolyl cis-trans isomerase NIMA-interacting 1

4.2.6 Score per residue for model 6

- Molecule 1: Peptidyl-prolyl cis-trans isomerase NIMA-interacting 1



4.2.7 Score per residue for model 7

- Molecule 1: Peptidyl-prolyl cis-trans isomerase NIMA-interacting 1

4.2.8 Score per residue for model 8

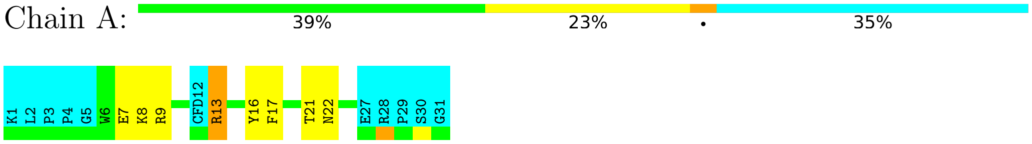
- Molecule 1: Peptidyl-prolyl cis-trans isomerase NIMA-interacting 1

Chain A:

4.2.9 Score per residue for model 9

- Molecule 1: Peptidyl-prolyl cis-trans isomerase NIMA-interacting 1

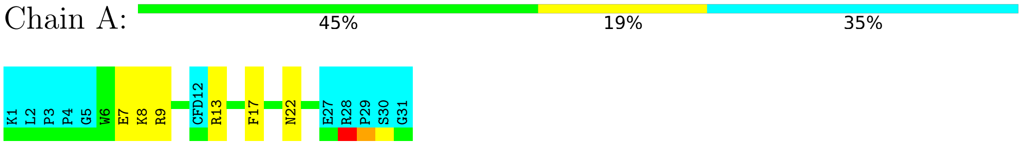
Chain A:



4.2.10 Score per residue for model 10

- Molecule 1: Peptidyl-prolyl cis-trans isomerase NIMA-interacting 1

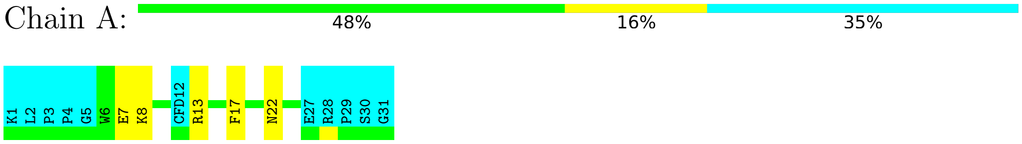
Chain A:



4.2.11 Score per residue for model 11

- Molecule 1: Peptidyl-prolyl cis-trans isomerase NIMA-interacting 1

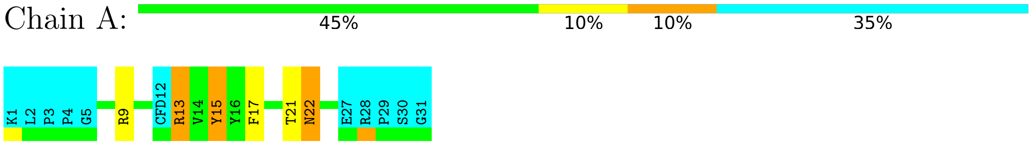
Chain A:



4.2.12 Score per residue for model 12

- Molecule 1: Peptidyl-prolyl cis-trans isomerase NIMA-interacting 1

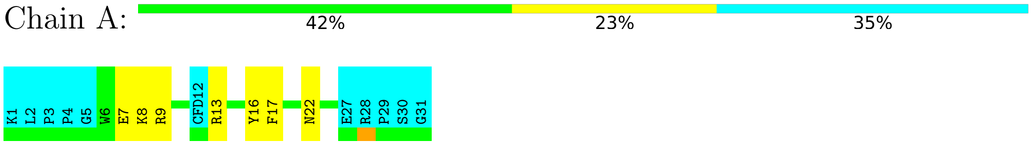
Chain A:



4.2.13 Score per residue for model 13

- Molecule 1: Peptidyl-prolyl cis-trans isomerase NIMA-interacting 1

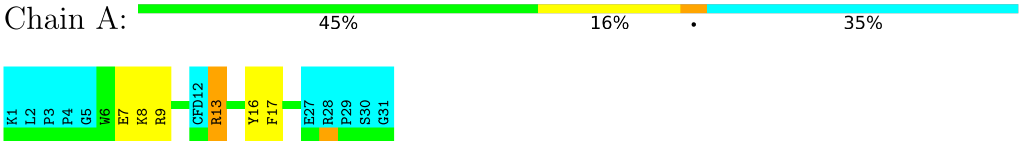
Chain A:



4.2.14 Score per residue for model 14

- Molecule 1: Peptidyl-prolyl cis-trans isomerase NIMA-interacting 1

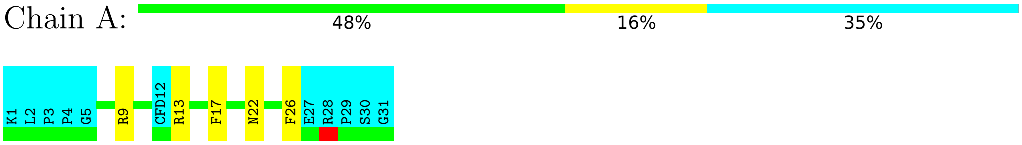
Chain A:



4.2.15 Score per residue for model 15

- Molecule 1: Peptidyl-prolyl cis-trans isomerase NIMA-interacting 1

Chain A:



5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *simulated annealing*.

Of the 27 calculated structures, 15 were deposited, based on the following criterion: *structures with acceptable covalent geometry*.

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

Software name	Classification	Version
Amber	refinement	
Amber	structure solution	

No chemical shift data was provided.

6 Model quality (i)

6.1 Standard geometry (i)

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

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	0.80±0.01	0±0/188 (0.0± 0.0%)	1.29±0.08	2±1/252 (1.0± 0.5%)
All	All	0.80	0/2820 (0.0%)	1.29	36/3780 (1.0%)

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(°)	Ideal(°)	Models	
								Worst	Total
1	A	13	ARG	NE-CZ-NH1	10.40	125.50	120.30	8	13
1	A	9	ARG	NE-CZ-NH1	9.37	124.98	120.30	3	12
1	A	9	ARG	NE-CZ-NH2	-7.39	116.61	120.30	12	5
1	A	16	TYR	CB-CG-CD2	-5.80	117.52	121.00	9	1
1	A	13	ARG	NE-CZ-NH2	-5.58	117.51	120.30	10	2
1	A	9	ARG	NH1-CZ-NH2	-5.28	113.60	119.40	13	2
1	A	15	TYR	CB-CG-CD1	-5.11	117.94	121.00	12	1

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
All	All	2730	2550	2535	-

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 -.

There are no clashes.

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	20/31 (65%)	18±1 (91±6%)	2±1 (8±5%)	0±0 (1±2%)	26	73
All	All	300/465 (65%)	274 (91%)	24 (8%)	2 (1%)	26	73

All 1 unique Ramachandran outliers are listed below.

Mol	Chain	Res	Type	Models (Total)
1	A	22	ASN	2

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	19/27 (70%)	15±1 (80±7%)	4±1 (20±7%)	3	33
All	All	285/405 (70%)	227 (80%)	58 (20%)	3	33

All 11 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	17	PHE	15
1	A	22	ASN	10
1	A	8	LYS	9
1	A	7	GLU	7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	13	ARG	6
1	A	10	MET	3
1	A	21	THR	3
1	A	16	TYR	2
1	A	25	GLN	1
1	A	15	TYR	1
1	A	26	PHE	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