wwPDB NMR Structure Validation Summary Report (i)

Apr 26, 2016 – 02:40 PM BST

PDB ID	:	1E91
Title	:	STRUCTURE OF THE COMPLEX OF THE MAD1-SIN3B INTERACTION DOMAINS
Authors	:	Spronk, C.A.E.M.; Tessari, M.; Kaan, A.M.; Jansen, J.F.A.; Vermeulen, M.; Stunnenberg, H.G.; Vuister, G.W.
Deposited on	:	2000-10-04

This is a wwPDB NMR Structure Validation Summary Report for a publicly released PDB entry.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at
http://wwpdb.org/validation/2016/NMRValidationReportHelp
with specific help available everywhere you see the (i) symbol.

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

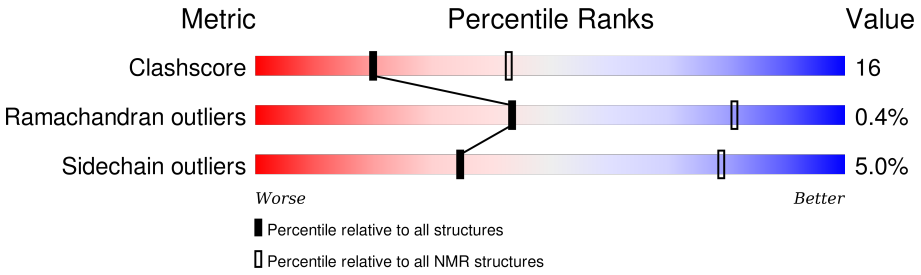
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Mogul	:	unknown
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	rb-20027457
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20027457

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 80%.

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	114402	11133
Ramachandran outliers	111179	9975
Sidechain outliers	111093	9958

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	85	33% 24% • 42%
2	B	13	54% 38% 8%

2 Ensemble composition and analysis

This entry contains 20 models. Model 19 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative.

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:10-A:37, A:57-A:77, B:2-B:13 (61)	0.21	19

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 1 single-model cluster was found.

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

3 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 1615 atoms, of which 795 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called PAIRED AMPHIPATHIC HELIX PROTEIN SIN3B.

Mol	Chain	Residues	Atoms						Trace
1	A	85	Total	C	H	N	O	S	0
			1405	454	691	122	137	1	

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	83	ALA	GLY	CONFLICT	UNP Q62141

- Molecule 2 is a protein called MAD PROTEIN (MAX DIMERIZER).

Mol	Chain	Residues	Atoms						Trace
2	B	13	Total	C	H	N	O	S	0
			210	67	104	15	23	1	

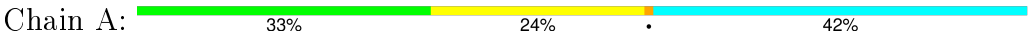
4 Residue-property plots

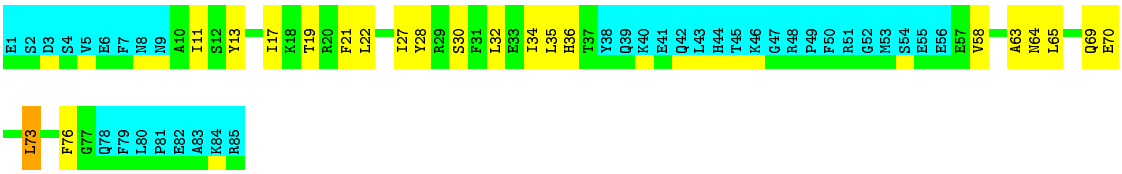
4.1 Average score per residue in the NMR ensemble

These plots are provided for all protein, RNA and DNA 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: PAIRED AMPHIPATHIC HELIX PROTEIN SIN3B

Chain A:





- Molecule 2: MAD PROTEIN (MAX DIMERIZER)

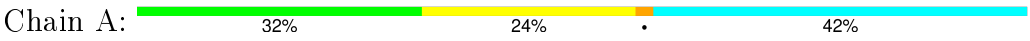
Chain B:

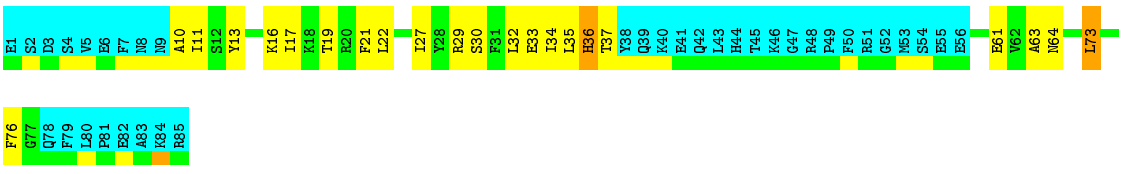
4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 19. Colouring as in section 4.1 above.

- Molecule 1: PAIRED AMPHIPATHIC HELIX PROTEIN SIN3B

Chain A:





- Molecule 2: MAD PROTEIN (MAX DIMERIZER)

Chain B:


N1
I2
L5
L6
E7
A8
L12
E13

5 Refinement protocol and experimental data overview

The models were refined using the following method: *SIMULATED ANNEALING, ITERATIVE NOE-ASSIGNMENT.*

Of the 100 calculated structures, 20 were deposited, based on the following criterion: *30 STRUCTURES WITH NO RESTRAINT VIOLATIONS, 20 LOWEST ENERGY STRUCTURES FINALLY SELECTED.*

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

Software name	Classification	Version
CNS 1.0	refinement	
CNS; ARIA	structure solution	

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)	BMRB entry 5457, BMRB entry 4841
Number of chemical shift lists	3
Total number of shifts	1538
Number of shifts mapped to atoms	1257
Number of unparsed shifts	0
Number of shifts with mapping errors	281
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	80%

No validations of the models with respect to experimental NMR restraints is performed at this time.

6 Model quality

6.1 Standard geometry

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

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	410	404	402	16±3
2	B	98	96	96	4±2
All	All	10160	10000	9960	326

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 16.

5 of 67 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:63:ALA:HA	1:A:73:LEU:HD21	0.81	1.53	3	16
1:A:11:ILE:HG23	2:B:12:LEU:HD11	0.70	1.63	19	1
1:A:35:LEU:HB2	2:B:6:LEU:HD13	0.65	1.68	10	9
1:A:35:LEU:HD22	2:B:2:ILE:HG23	0.63	1.68	4	15
1:A:66:PHE:HB2	1:A:73:LEU:HD23	0.63	1.70	7	3

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	49/85 (58%)	44±1 (90±2%)	5±1 (10±2%)	0±0 (0±1%)	56 85
2	B	11/13 (85%)	11±1 (97±5%)	0±0 (1±3%)	0±0 (1±3%)	19 64
All	All	1200/1960 (61%)	1094 (91%)	101 (8%)	5 (0%)	43 81

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)
2	B	2	ILE	3
1	A	70	GLU	1
1	A	66	PHE	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	45/78 (58%)	43±1 (95±2%)	2±1 (5±2%)	38 81
2	B	10/11 (91%)	9±1 (93±6%)	1±1 (7±6%)	23 69
All	All	1100/1780 (62%)	1045 (95%)	55 (5%)	35 79

5 of 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	73	LEU	16
2	B	3	GLN	12
1	A	69	GLN	9
1	A	36	HIS	7
1	A	27	ILE	3

6.3.3 RNA (i)

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 carbohydrates 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

The completeness of assignment taking into account all chemical shift lists is 80% for the well-defined parts and 82% for the entire structure.

7.1 Chemical shift list 1

File name: BMRB entry 4841

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	1271
Number of shifts mapped to atoms	1061
Number of unparsed shifts	0
Number of shifts with mapping errors	210
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	0

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

- Residue not found in structure. First 5 (of 210) occurences are reported below.

Chain	Res	Type	Atom	Shift Data		
				Value	Uncertainty	Ambiguity
A	101	LYS	HE3	2.981	0.03	1
A	100	GLN	HB2	1.951	0.03	2
A	102	ASN	N	120.559	0.1	1
A	92	GLY	CA	45.444	0.1	1
A	96	MET	C	176.199	0.1	1

7.1.2 Chemical shift referencing

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	105	-0.51 ± 0.09	Should be applied
$^{13}C_{\beta}$	98	0.42 ± 0.10	None needed (< 0.5 ppm)
$^{13}C'$	101	-0.46 ± 0.10	None needed (< 0.5 ppm)

Continued on next page...

Continued from previous page...

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
^{15}N	102	-0.36 ± 0.37	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 71%, i.e. 557 atoms were assigned a chemical shift out of a possible 790. 9 out of 12 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	242/303 (80%)	97/121 (80%)	97/122 (80%)	48/60 (80%)
Sidechain	272/402 (68%)	166/233 (71%)	103/154 (67%)	3/15 (20%)
Aromatic	43/85 (51%)	28/45 (62%)	15/36 (42%)	0/4 (0%)
Overall	557/790 (71%)	291/399 (73%)	215/312 (69%)	51/79 (65%)

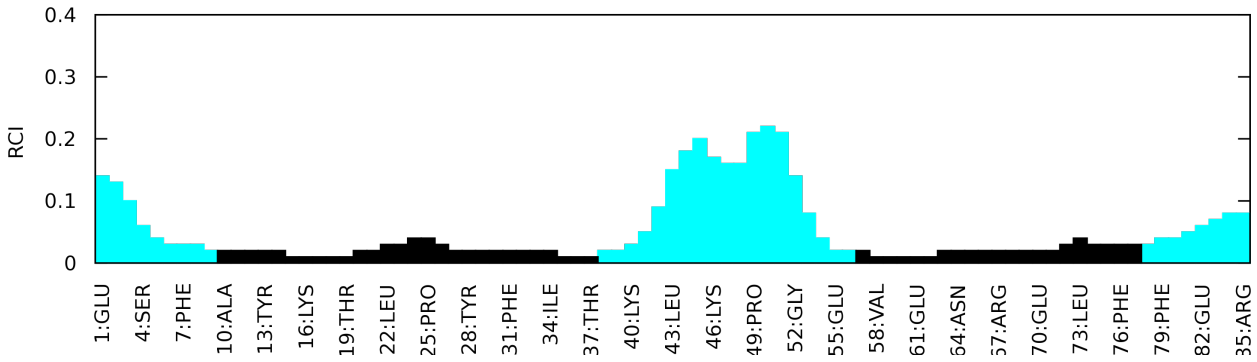
7.1.4 Statistically unusual chemical shifts

There are no statistically unusual chemical shifts.

7.1.5 Random Coil Index (RCI) plots

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.

Random coil index (RCI) for chain A:



7.2 Chemical shift list 2

File name: BMRB entry 5457

Chemical shift list name: *assigned_chem_shift_list_1*

7.2.1 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	169
Number of shifts mapped to atoms	98
Number of unparsed shifts	0
Number of shifts with mapping errors	71
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	0

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

- Residue not found in structure. First 5 (of 71) occurences are reported below.

Chain	Res	Type	Atom	Shift Data		
				Value	Uncertainty	Ambiguity
B	19	GLU	HB3	2.082	0.01	1
B	18	ARG	HG2	1.448	0.01	2
B	22	ALA	HB1	1.404	0.01	1
B	18	ARG	HD3	3.033	0.01	2
B	2	ARG	HE	7.329	0.01	1

7.2.2 Chemical shift referencing

No chemical shift referencing corrections were calculated (not enough data).

7.2.3 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 9%, i.e. 72 atoms were assigned a chemical shift out of a possible 790. 0 out of 12 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	24/303 (8%)	24/121 (20%)	0/122 (0%)	0/60 (0%)

Continued on next page...

Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Sidechain	44/402 (11%)	44/233 (19%)	0/154 (0%)	0/15 (0%)
Aromatic	4/85 (5%)	4/45 (9%)	0/36 (0%)	0/4 (0%)
Overall	72/790 (9%)	72/399 (18%)	0/312 (0%)	0/79 (0%)

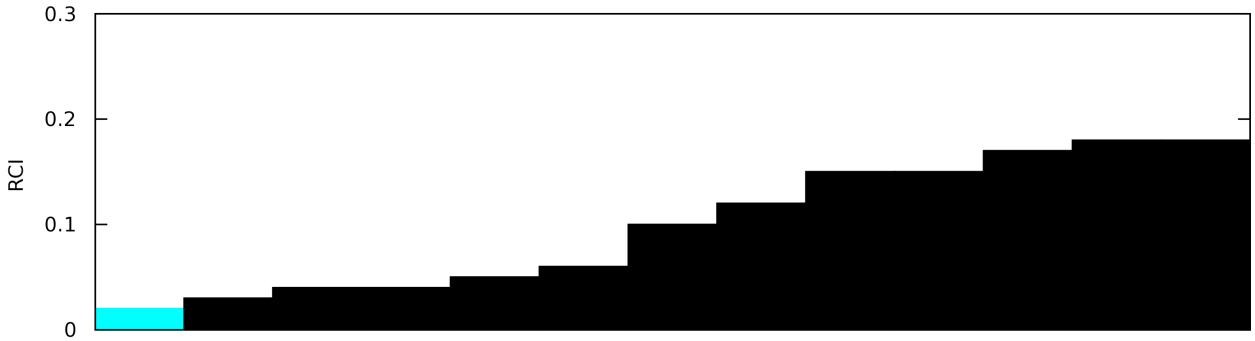
7.2.4 Statistically unusual chemical shifts

There are no statistically unusual chemical shifts.

7.2.5 Random Coil Index (RCI) plots

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.

Random coil index (RCI) for chain B:



7.3 Chemical shift list 3

File name: BMRB entry 4841

Chemical shift list name: *assigned_chem_shift_list_2*

7.3.1 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	98
Number of shifts mapped to atoms	98

Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	0

7.3.2 Chemical shift referencing

No chemical shift referencing corrections were calculated (not enough data).

7.3.3 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 9%, i.e. 72 atoms were assigned a chemical shift out of a possible 790. 0 out of 12 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	24/303 (8%)	24/121 (20%)	0/122 (0%)	0/60 (0%)
Sidechain	44/402 (11%)	44/233 (19%)	0/154 (0%)	0/15 (0%)
Aromatic	4/85 (5%)	4/45 (9%)	0/36 (0%)	0/4 (0%)
Overall	72/790 (9%)	72/399 (18%)	0/312 (0%)	0/79 (0%)

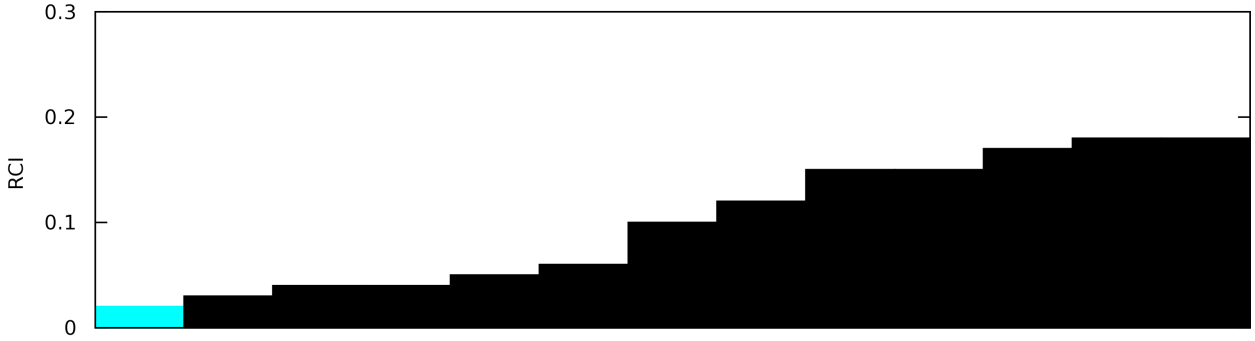
7.3.4 Statistically unusual chemical shifts

There are no statistically unusual chemical shifts.

7.3.5 Random Coil Index (RCI) plots

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.

Random coil index (RCI) for chain B:


0.3
0.2
0.1
0
RCI