Full wwPDB NMR Structure Validation Report (i)

Mar 3, 2022 – 09:02 AM EST

PDB ID	:	2HNA
Title	:	Solution Structure of a bacterial apo-flavodoxin
Authors	:	Hu, Y.; Jin, C.
Deposited on	:	2006-07-12

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

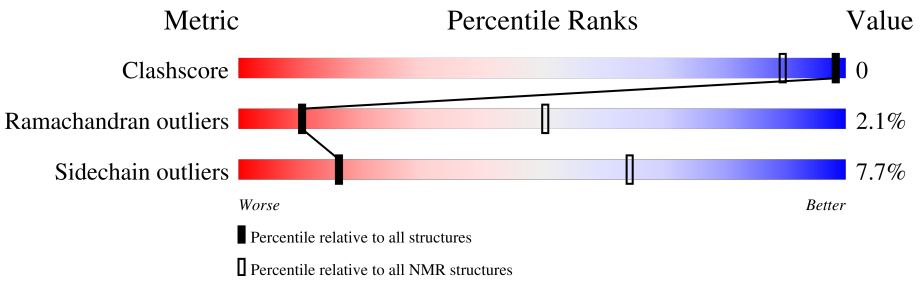
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.27
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.27

1 Overall quality at a glance (i)

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	147	73% 5% 22%

2 Ensemble composition and analysis (i)

This entry contains 21 models. 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:1-A:56, A:65-A:92, A:104-A:123, A:137-A:147 (115)	0.39	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 4 clusters and 8 single-model clusters were found.

Cluster number	Models
1	1, 10, 11, 12, 16, 19
2	5, 8, 18
3	7, 9
4	6, 13
Single-model clusters	2; 3; 4; 14; 15; 17; 20; 21

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 2196 atoms, of which 1081 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Protein mioC.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	147	2196	706	1081	178	229	2	0

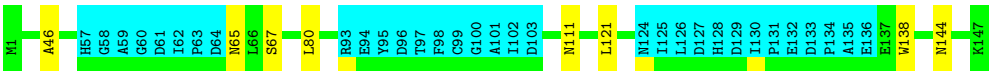
4 Residue-property plots (i)

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: Protein mioC

Chain A:



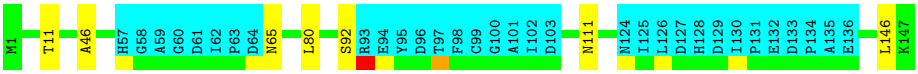
4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1 (medoid)

• Molecule 1: Protein mioC

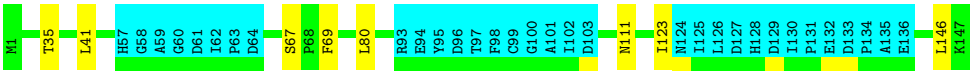
Chain A:



4.2.2 Score per residue for model 2

• Molecule 1: Protein mioC

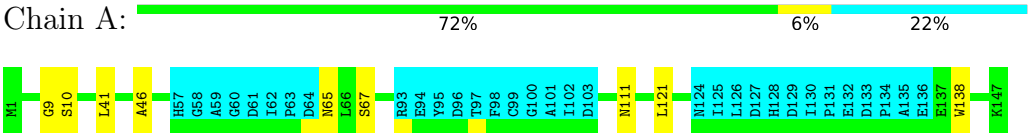
Chain A:



4.2.3 Score per residue for model 3

• Molecule 1: Protein mioC

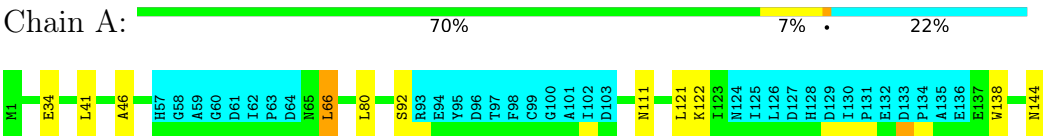
Chain A:



4.2.4 Score per residue for model 4

• Molecule 1: Protein mioC

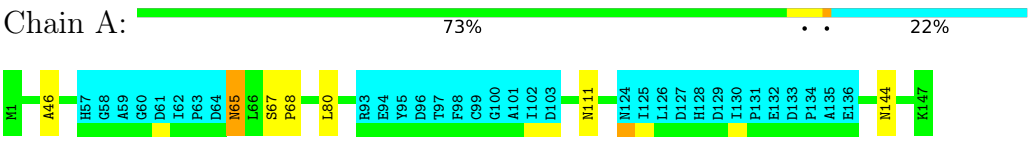
Chain A:



4.2.5 Score per residue for model 5

• Molecule 1: Protein mioC

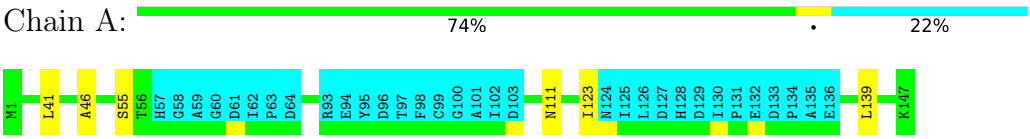
Chain A:



4.2.6 Score per residue for model 6

• Molecule 1: Protein mioC

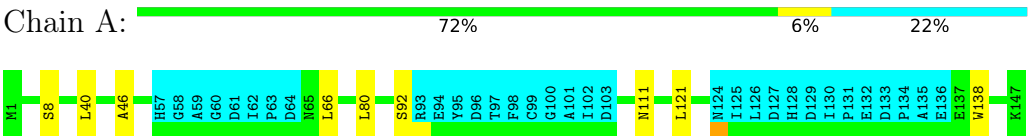
Chain A:



4.2.7 Score per residue for model 7

• Molecule 1: Protein mioC

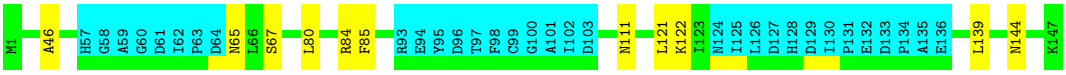
Chain A:



4.2.8 Score per residue for model 8

- Molecule 1: Protein mioC

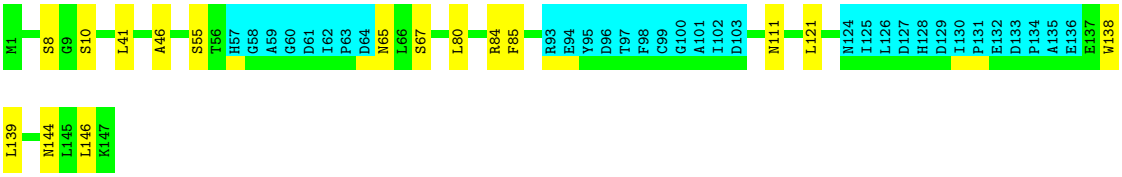
Chain A: 71% 7% 22%



4.2.9 Score per residue for model 9

- Molecule 1: Protein mioC

Chain A: 67% 11% 22%



4.2.10 Score per residue for model 10

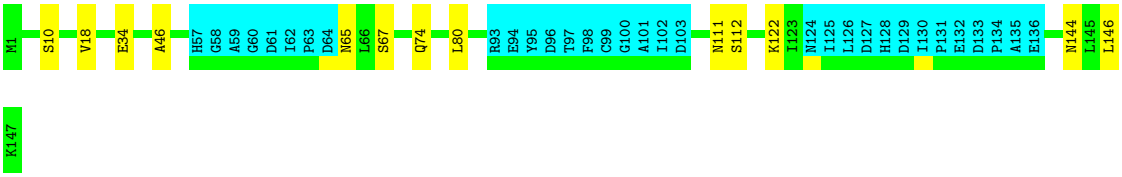
- Molecule 1: Protein mioC

Chain A: 73% 5% 22%

4.2.11 Score per residue for model 11

- Molecule 1: Protein mioC

Chain A: 69% 9% 22%



4.2.12 Score per residue for model 12

- Molecule 1: Protein mioC

Chain A: 69% 8% • 22%

M1 T35 A46 W50 H57 G58 A59 G60 D61 I62 P63 D64 N65 L66 S67 P68 V83 R84 R93 E94 Y95 D96 T97 F98 C99 G100 A101 I102 D103 N111 L121 K122 I123 N124 I125 L126 D127 H128 D129 I130 P131 E132 D133 P134 A135 E136 N144 L145 L146 K147

4.2.13 Score per residue for model 13

- Molecule 1: Protein mioC

Chain A: 72% 6% 22%

M1 S10 E34 A46 H57 G58 A59 G60 D61 I62 P63 D64 N65 L66 S67 L80 R93 E94 Y95 D96 T97 F98 C99 G100 A101 I102 D103 N111 L121 N124 I125 L126 D127 H128 D129 I130 P131 E132 D133 P134 A135 E136 E137 W138 K147

4.2.14 Score per residue for model 14

- Molecule 1: Protein mioC

Chain A: 71% 7% 22%

M1 S8 E34 A46 H57 G58 A59 G60 D61 I62 P63 D64 N65 Q74 R84 S92 R93 E94 Y95 D96 T97 F98 C99 G100 A101 I102 D103 N111 I123 N124 I125 L126 D127 H128 D129 I130 P131 E132 D133 P134 A135 E136 L139 N144 K147

4.2.15 Score per residue for model 15

- Molecule 1: Protein mioC

Chain A: 73% 5% 22%

M1 S8 G9 S10 A46 H57 G58 A59 G60 D61 I62 P63 D64 N65 L80 R84 R93 E94 Y95 D96 T97 F98 C99 G100 A101 I102 D103 N111 L121 N124 I125 L126 D127 H128 D129 I130 P131 E132 D133 P134 A135 E136 K147

4.2.16 Score per residue for model 16

- Molecule 1: Protein mioC

Chain A: 72% 6% 22%

M1 A46 H57 G58 A59 G60 D61 I62 P63 D64 F69 L80 R93 E94 Y95 D96 T97 F98 C99 G100 A101 I102 D103 N111 S112 L121 K122 I123 N124 I125 L126 D127 H128 D129 I130 P131 E132 D133 P134 A135 E136 E137 W138 L139 K147

4.2.17 Score per residue for model 17

- Molecule 1: Protein mioC

Chain A: 72% 6% 22%

4.2.18 Score per residue for model 18

- Molecule 1: Protein mioC

Chain A: 72% 6% 22%

4.2.19 Score per residue for model 19

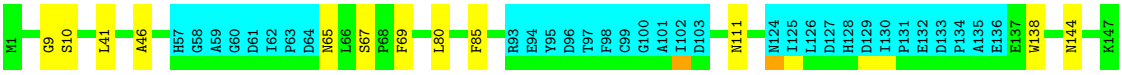
- Molecule 1: Protein mioC

Chain A: 69% 9% 22%

4.2.20 Score per residue for model 20

- Molecule 1: Protein mioC

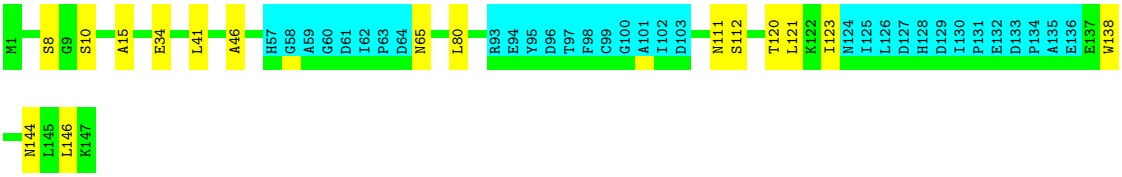
Chain A: 70% 8% 22%



4.2.21 Score per residue for model 21

- Molecule 1: Protein mioC

Chain A: 67% 11% 22%


M1
S8
G9
S10
A15
E34
L41
A46
H57
G58
A59
G60
D61
I62
P63
D64
N65
L80
R93
E94
Y95
D96
T97
F98
C99
G100
A101
I102
D103
N111
S112
T120
L121
K122
I123
N124
I125
L126
D127
H128
D129
I130
P131
E132
D133
P134
A135
E136
E137
W138
N144
L145
L146
K147

5 Refinement protocol and experimental data overview (i)

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

Of the 100 calculated structures, 21 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
CYANA	structure solution	2.0
Amber	refinement	7.0

No chemical shift data was provided.

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	0.60±0.00	0±0/885 (0.0± 0.0%)	0.95±0.02	0±0/1199 (0.0± 0.0%)
All	All	0.60	0/18585 (0.0%)	0.95	1/25179 (0.0%)

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	0.2±0.4
All	All	0	4

There are no bond-length outliers.

All unique angle outliers are listed below.

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	A	84	ARG	CD-NE-CZ	5.03	130.64	123.60	14	1

There are no chirality outliers.

All unique planar outliers are listed below.

Mol	Chain	Res	Type	Group	Models (Total)
1	A	84	ARG	Sidechain	4

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	869	872	872	0±0
All	All	18249	18312	18312	8

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:66:LEU:H	1:A:66:LEU:HD23	0.50	1.66	4	1
1:A:121:LEU:HD12	1:A:138:TRP:CE3	0.45	2.46	18	1
1:A:80:LEU:HD23	1:A:112:SER:HB3	0.44	1.88	16	3
1:A:121:LEU:HD23	1:A:122:LYS:H	0.42	1.75	8	1
1:A:50:TRP:CD1	1:A:83:VAL:HG21	0.41	2.50	12	1
1:A:50:TRP:CE3	1:A:85:PHE:CZ	0.40	3.10	19	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	113/147 (77%)	96±2 (85±2%)	15±2 (13±2%)	2±1 (2±1%)	10 50
All	All	2373/3087 (77%)	2014 (85%)	309 (13%)	50 (2%)	10 50

All 10 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	46	ALA	17
1	A	65	ASN	12
1	A	41	LEU	6
1	A	67	SER	4
1	A	55	SER	3
1	A	9	GLY	2
1	A	68	PRO	2
1	A	15	ALA	2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	92	SER	1
1	A	10	SER	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	92/118 (78%)	85±2 (92±2%)	7±2 (8±2%)	16 64
All	All	1932/2478 (78%)	1784 (92%)	148 (8%)	16 64

All 26 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	111	ASN	21
1	A	80	LEU	14
1	A	121	LEU	12
1	A	144	ASN	12
1	A	138	TRP	11
1	A	67	SER	10
1	A	146	LEU	8
1	A	123	ILE	8
1	A	10	SER	7
1	A	34	GLU	6
1	A	8	SER	6
1	A	139	LEU	5
1	A	69	PHE	4
1	A	92	SER	3
1	A	65	ASN	3
1	A	85	PHE	3
1	A	35	THR	2
1	A	41	LEU	2
1	A	66	LEU	2
1	A	122	LYS	2
1	A	74	GLN	2
1	A	11	THR	1
1	A	40	LEU	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	18	VAL	1
1	A	84	ARG	1
1	A	120	THR	1

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 monosaccharides 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)

No chemical shift data were provided