Full wwPDB NMR Structure Validation Report (i)

Apr 21, 2024 – 08:13 AM EDT

PDB ID	:	2L68
BMRB ID	:	17303
Title	:	Solution Structure of Human Holo L-FABP
Authors	:	Cai, J.; Luecke, C.; Chen, Z.; Qiao, Y.; Klimtchuk, E.S.; Hamilton, J.A.
Deposited on	:	2010-11-17

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:

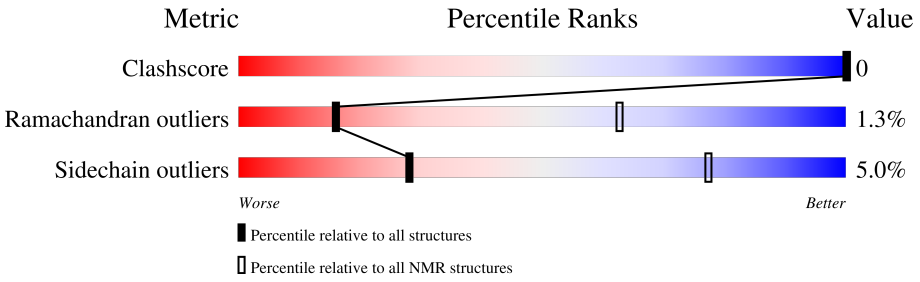
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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	126	86% 5% 10%

2 Ensemble composition and analysis (i)

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

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:3-A:53, A:59-A:71, A:78-A:127 (114)	0.67	20

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 5 single-model clusters were found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 2004 atoms, of which 1017 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Fatty acid-binding protein, liver.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	126	2004	626	1017	163	194	4	0

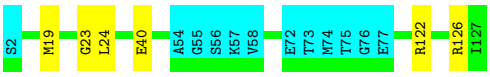
4 Residue-property plots (i)

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: Fatty acid-binding protein, liver

Chain A: 86% 5% 10%



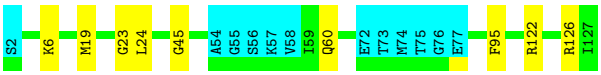
4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1

• Molecule 1: Fatty acid-binding protein, liver

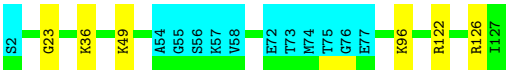
Chain A: 83% 7% 10%



4.2.2 Score per residue for model 2

• Molecule 1: Fatty acid-binding protein, liver

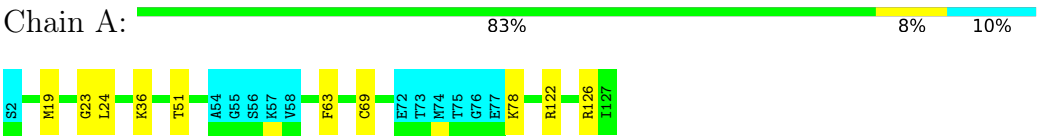
Chain A: 86% 5% 10%



4.2.3 Score per residue for model 3

• Molecule 1: Fatty acid-binding protein, liver

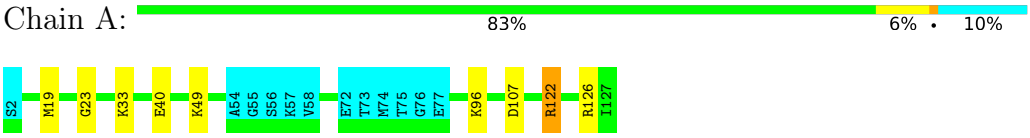
Chain A:



4.2.4 Score per residue for model 4

• Molecule 1: Fatty acid-binding protein, liver

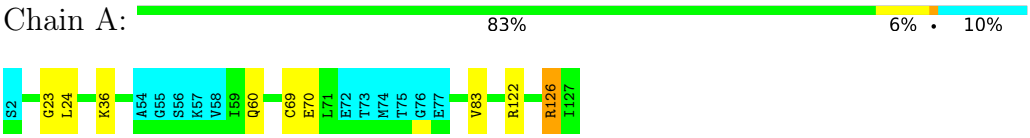
Chain A:



4.2.5 Score per residue for model 5

• Molecule 1: Fatty acid-binding protein, liver

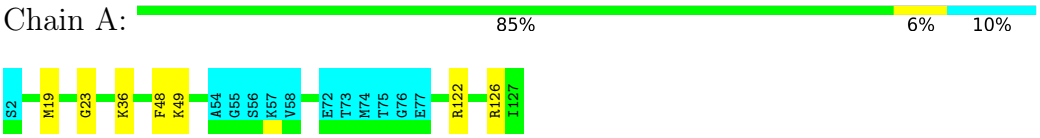
Chain A:



4.2.6 Score per residue for model 6

• Molecule 1: Fatty acid-binding protein, liver

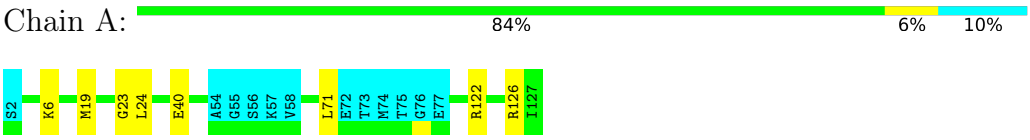
Chain A:



4.2.7 Score per residue for model 7

• Molecule 1: Fatty acid-binding protein, liver

Chain A:



4.2.8 Score per residue for model 8

• Molecule 1: Fatty acid-binding protein, liver

Chain A: 84% 6% 10%

S2 M19 G23 E40 A54 G55 S56 K57 V58 E62 E72 T73 M74 T75 G76 E77 F95 N111 R122 R126 I127

4.2.9 Score per residue for model 9

• Molecule 1: Fatty acid-binding protein, liver

Chain A: 83% 8% 10%

S2 G23 L24 E40 A54 G55 S56 K57 V58 E62 E70 L71 E72 T73 M74 T75 G76 E77 F95 E103 G106 R122 R126 I127

4.2.10 Score per residue for model 10

• Molecule 1: Fatty acid-binding protein, liver

Chain A: 82% 9% 10%

S2 M19 G23 L24 K31 E40 A54 G55 S56 K57 V58 E70 L71 E72 T73 M74 T75 G76 E77 K99 E103 L104 R122 R126 I127

4.2.11 Score per residue for model 11

• Molecule 1: Fatty acid-binding protein, liver

Chain A: 84% 6% 10%

S2 M19 G23 E40 A54 G55 S56 K57 V58 E70 L71 E72 T73 M74 T75 G76 E77 K78 I108 R122 R126 I127

4.2.12 Score per residue for model 12

• Molecule 1: Fatty acid-binding protein, liver

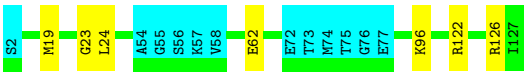
Chain A: 84% 6% 10%

S2 M19 G23 E40 K46 A54 G55 S56 K57 V58 I59 Q60 E72 T73 M74 T75 G76 E77 N89 R122 R126 I127

4.2.13 Score per residue for model 13

• Molecule 1: Fatty acid-binding protein, liver

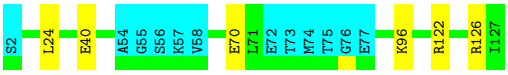
Chain A: 85% 6% 10%



4.2.14 Score per residue for model 14

• Molecule 1: Fatty acid-binding protein, liver

Chain A: 86% 5% 10%



4.2.15 Score per residue for model 15

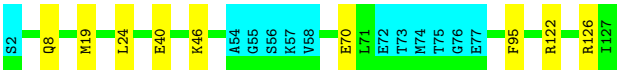
• Molecule 1: Fatty acid-binding protein, liver

Chain A: 85% 5% • 10%

4.2.16 Score per residue for model 16

• Molecule 1: Fatty acid-binding protein, liver

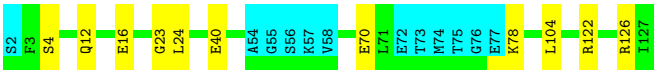
Chain A: 83% 7% 10%



4.2.17 Score per residue for model 17

• Molecule 1: Fatty acid-binding protein, liver

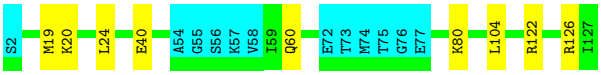
Chain A: 82% 9% 10%



4.2.18 Score per residue for model 18

• Molecule 1: Fatty acid-binding protein, liver

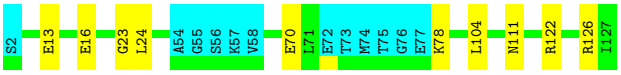
Chain A:



4.2.19 Score per residue for model 19

• Molecule 1: Fatty acid-binding protein, liver

Chain A:



4.2.20 Score per residue for model 20 (medoid)

• Molecule 1: Fatty acid-binding protein, liver

Chain A:

5 Refinement protocol and experimental data overview (i)

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

Of the 100 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	2.1
Discover	refinement	2000

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	1.02±0.00	0±0/916 (0.0± 0.0%)	1.05±0.02	2±0/1226 (0.2± 0.0%)
All	All	1.02	0/18320 (0.0%)	1.05	44/24520 (0.2%)

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	122	ARG	NE-CZ-NH1	7.91	124.26	120.30	8	20
1	A	126	ARG	NE-CZ-NH1	7.79	124.20	120.30	17	20
1	A	115	LEU	CB-CA-C	5.41	120.47	110.20	15	1
1	A	122	ARG	NE-CZ-NH2	-5.35	117.62	120.30	12	1
1	A	126	ARG	NE-CZ-NH2	-5.12	117.74	120.30	20	1
1	A	95	PHE	CB-CA-C	5.02	120.43	110.40	1	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
1	A	906	939	939	0±0
All	All	18120	18780	18780	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:63:PHE:HB2	1:A:69:CYS:SG	0.40	2.56	3	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	113/126 (90%)	99±2 (88±2%)	12±2 (11±2%)	2±1 (1±1%)	16	63
All	All	2260/2520 (90%)	1985 (88%)	245 (11%)	30 (1%)	16	63

All 9 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	23	GLY	17
1	A	95	PHE	4
1	A	96	LYS	3
1	A	45	GLY	1
1	A	83	VAL	1
1	A	106	GLY	1
1	A	89	ASN	1
1	A	115	LEU	1
1	A	4	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	103/112 (92%)	98±2 (95±2%)	5±2 (5±2%)	28	77

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	2060/2240 (92%)	1958 (95%)	102 (5%)	28 77

All 32 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	19	MET	14
1	A	24	LEU	13
1	A	40	GLU	11
1	A	70	GLU	9
1	A	60	GLN	5
1	A	36	LYS	5
1	A	78	LYS	4
1	A	62	GLU	4
1	A	104	LEU	4
1	A	49	LYS	3
1	A	46	LYS	3
1	A	6	LYS	2
1	A	111	ASN	2
1	A	103	GLU	2
1	A	31	LYS	2
1	A	12	GLN	2
1	A	16	GLU	2
1	A	96	LYS	1
1	A	51	THR	1
1	A	33	LYS	1
1	A	107	ASP	1
1	A	122	ARG	1
1	A	69	CYS	1
1	A	126	ARG	1
1	A	48	PHE	1
1	A	71	LEU	1
1	A	99	LYS	1
1	A	108	ILE	1
1	A	8	GLN	1
1	A	20	LYS	1
1	A	80	LYS	1
1	A	13	GLU	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