



wwPDB NMR Structure Validation Summary Report ⓘ

Apr 26, 2016 – 04:05 PM BST

PDB ID : 1LC2
Title : Solution Structure Of Reduced Horse Heart Cytochrome c in 30% Acetonitrile
Solution, NMR 30 Structures
Authors : Sivakolundu, S.G.; Mabrouk, P.A.
Deposited on : 2002-04-04

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 ⓘ symbol.

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 : 1.7.1 (RC1), CSD as537be (2016)
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

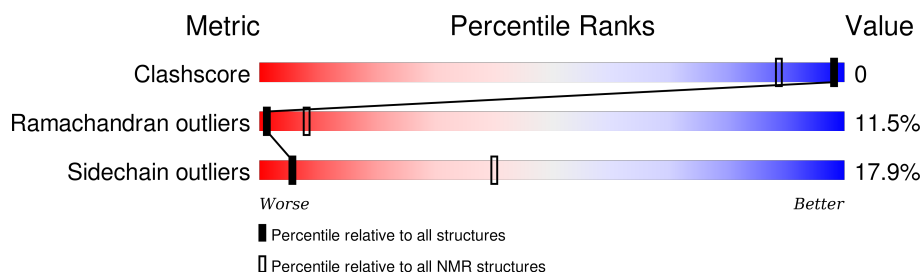
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

SOLUTION NMR

The overall completeness of chemical shifts assignment is 46%.

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 $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	A	104	 77% 17% • •

2 Ensemble composition and analysis

This entry contains 30 models. Model 14 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:1-A:23, A:28-A:104 (100)	0.57	14

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 2 single-model clusters were found.

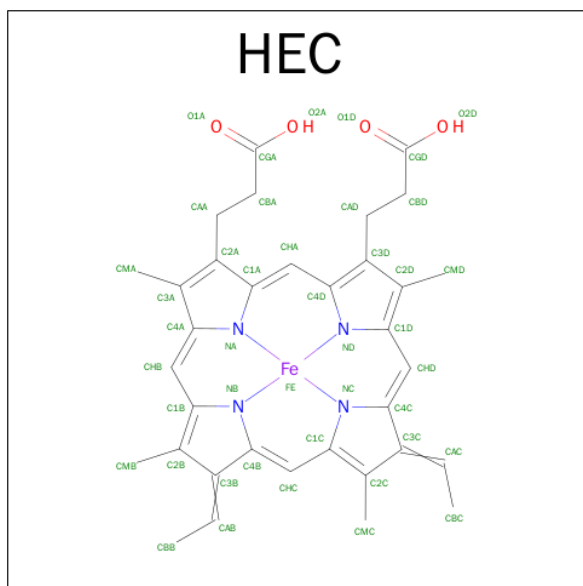
Cluster number	Models
1	1, 2, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 16, 17, 18, 19, 21, 23, 24, 25, 26, 27, 28, 29, 30
2	15, 22
Single-model clusters	3; 20

i

- Molecule 1 is a protein called CYTOCHROME C.

Mol	Chain	Residues	Atoms						Trace
1	A	104	Total	C	H	N	O	S	0
			1670	524	847	144	151	4	

- Molecule 2 is HEME C (three-letter code: HEC) (formula: $\text{C}_{34}\text{H}_{34}\text{FeN}_4\text{O}_4$).



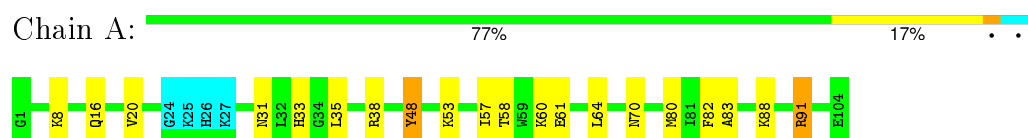
Mol	Chain	Residues	Atoms					
2	A	1	Total	C	Fe	H	N	O
			75	34	1	32	4	4

4 Residue-property plots [i](#)

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.

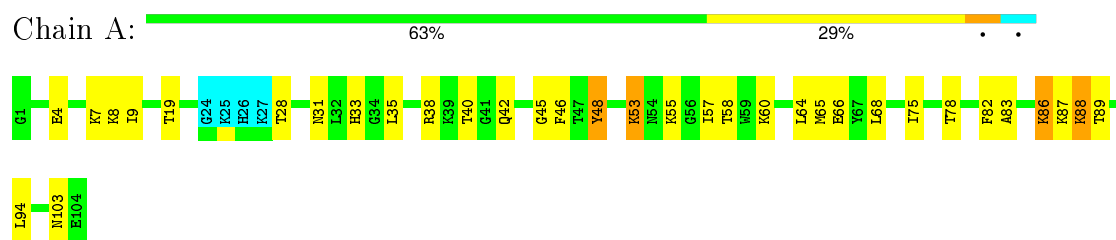
- Molecule 1: CYTOCHROME C



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

The representative model is number 14. Colouring as in section 4.1 above.

- Molecule 1: CYTOCHROME C



5 Refinement protocol and experimental data overview

The models were refined using the following method: *SIMULATED ANNEALING*, *TORSION ANGLE DYNAMICS*, *RESTRAINED ENERGY MINIMIZATION*.

Of the 30 calculated structures, 30 were deposited, based on the following criterion: *structures with the least restraint violations*.

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

Software name	Classification	Version
DYANA	structure solution	1.5
AMBER	refinement	5.0
AMBER	structure solution	5.0

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 5372
Number of chemical shift lists	1
Total number of shifts	677
Number of shifts mapped to atoms	677
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	46%

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

6 Model quality

6.1 Standard geometry

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

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	0.68±0.01	0±0/806 (0.0±0.0%)	1.09±0.03	2±1/1076 (0.2±0.1%)
All	All	0.68	0/24180 (0.0%)	1.09	53/32280 (0.2%)

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	2.4±1.2
All	All	0	72

There are no bond-length outliers.

5 of 10 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	91	ARG	NE-CZ-NH1	9.99	125.29	120.30	18	22
1	A	38	ARG	NE-CZ-NH1	9.23	124.92	120.30	24	13
1	A	91	ARG	NE-CZ-NH2	-6.98	116.81	120.30	28	9
1	A	48	TYR	CB-CG-CD2	-6.72	116.97	121.00	17	3
1	A	46	PHE	CB-CG-CD2	-5.79	116.75	120.80	5	1

There are no chirality outliers.

5 of 22 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	48	TYR	Sidechain	16

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group	Models (Total)
1	A	91	ARG	Sidechain	14
1	A	38	ARG	Sidechain	12
1	A	1	GLY	Peptide	4
1	A	103	ASN	Peptide	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	791	811	811	1±1
2	A	43	32	30	0±0
All	All	25020	25290	25230	23

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.

5 of 21 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:48:TYR:CG	1:A:53:LYS:HE2	0.62	2.30	23	2
1:A:35:LEU:CD2	2:A:105:HEC:HMA3	0.52	2.34	17	1
1:A:48:TYR:CD1	1:A:53:LYS:HE3	0.52	2.39	8	1
1:A:36:PHE:HB3	1:A:99:LYS:HE2	0.51	1.83	7	1
1:A:36:PHE:HB2	1:A:102:THR:HG21	0.50	1.83	1	2

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	98/104 (94%)	66±3 (68±3%)	20±4 (21±4%)	11±3 (12±3%)	1 8

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	2940/3120 (94%)	1991 (68%)	610 (21%)	339 (12%)	1 8

5 of 55 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	80	MET	20
1	A	82	PHE	19
1	A	58	THR	19
1	A	83	ALA	19
1	A	33	HIS	16

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	83/86 (97%)	68±3 (82±3%)	15±3 (18±3%)	5 40
All	All	2490/2580 (97%)	2044 (82%)	446 (18%)	5 40

5 of 62 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	HIS	27
1	A	48	TYR	24
1	A	16	GLN	21
1	A	91	ARG	20
1	A	53	LYS	20

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 carbohydrates in this entry.

6.6 Ligand geometry

1 ligand is modelled in this entry.

In the following table, the Counts columns list the number of bonds for which Mogul statistics could be retrieved, the number of bonds that are observed in the model and the number of bonds that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length is the number of standard deviations the observed value is removed from the expected value. A bond length with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond lengths.

Mol	Type	Chain	Res	Link	Bond lengths		
					Counts	RMSZ	#Z>2
2	HEC	A	105	1	24,50,50	1.42±0.03	0±0 (0±0%)

In the following table, the Counts columns list the number of angles for which Mogul statistics could be retrieved, the number of angles that are observed in the model and the number of angles that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond angle is the number of standard deviations the observed value is removed from the expected value. A bond angle with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond angles.

Mol	Type	Chain	Res	Link	Bond angles		
					Counts	RMSZ	#Z>2
2	HEC	A	105	1	19,82,82	2.74±0.11	2±0 (10±0%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	HEC	A	105	1	-	0±0,6,54,54	0±0,0,8,8

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
2	A	105	HEC	CBB-CAB-C3B	7.94	109.97	127.34	27	30
2	A	105	HEC	CBC-CAC-C3C	7.03	111.97	127.34	9	30
2	A	105	HEC	CBD-CAD-C3D	5.10	121.47	112.49	27	1

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

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

The completeness of assignment taking into account all chemical shift lists is 46% for the well-defined parts and 46% for the entire structure.

7.1 Chemical shift list 1

File name: BMRB entry 5372

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	677
Number of shifts mapped to atoms	677
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	31

7.1.2 Chemical shift referencing

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

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 46%, i.e. 579 atoms were assigned a chemical shift out of a possible 1252. 0 out of 9 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	¹ H	¹³ C	¹⁵ N
Backbone	196/492 (40%)	196/196 (100%)	0/200 (0%)	0/96 (0%)
Sidechain	339/666 (51%)	339/393 (86%)	0/242 (0%)	0/31 (0%)
Aromatic	44/94 (47%)	44/50 (88%)	0/41 (0%)	0/3 (0%)
Overall	579/1252 (46%)	579/639 (91%)	0/483 (0%)	0/130 (0%)

7.1.4 Statistically unusual chemical shifts ⓘ

The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

Mol	Chain	Res	Type	Atom	Shift, ppm	Expected range, ppm	Z-score
1	A	67	TYR	HE1	0.08	7.86 – 5.56	-28.8
1	A	67	TYR	HE2	0.44	7.86 – 5.56	-27.3
1	A	67	TYR	HD1	0.73	8.44 – 5.44	-20.7
1	A	67	TYR	HD2	0.92	8.44 – 5.44	-20.1
1	A	80	MET	HG3	-3.80	4.30 – 0.50	-16.3
1	A	18	HIS	HD2	0.08	9.28 – 4.78	-15.4
1	A	18	HIS	HE1	0.45	10.53 – 5.43	-14.8
1	A	80	MET	HB2	-2.61	3.73 – 0.33	-13.6
1	A	80	MET	HG2	-1.92	4.23 – 0.63	-12.1
1	A	29	GLY	HA3	-0.25	5.80 – 2.00	-10.9
1	A	80	MET	HE2	-3.33	4.28 – -0.52	-10.9
1	A	80	MET	HE3	-3.33	4.28 – -0.52	-10.9
1	A	80	MET	HE1	-3.33	4.28 – -0.52	-10.9
1	A	30	PRO	HD3	0.90	5.52 – 1.72	-7.2
1	A	46	PHE	HD2	4.97	8.56 – 5.56	-7.0
1	A	80	MET	HB3	-0.27	3.70 – 0.30	-6.7
1	A	82	PHE	HB2	0.54	4.85 – 1.15	-6.6
1	A	18	HIS	HB2	0.74	4.91 – 1.31	-6.6
1	A	30	PRO	HD2	1.33	5.45 – 1.85	-6.4
1	A	41	GLY	HA3	1.51	5.80 – 2.00	-6.3
1	A	71	PRO	HG2	-0.01	3.48 – 0.38	-6.3
1	A	32	LEU	HD23	-0.81	2.14 – -0.66	-5.5
1	A	32	LEU	HD22	-0.81	2.14 – -0.66	-5.5
1	A	32	LEU	HD21	-0.81	2.14 – -0.66	-5.5
1	A	97	TYR	HE1	5.50	7.86 – 5.56	-5.3
1	A	17	CYS	HB2	0.59	5.20 – 0.70	-5.2
1	A	46	PHE	HB2	1.06	4.85 – 1.15	-5.2
1	A	18	HIS	HB3	1.04	5.00 – 1.10	-5.2
1	A	32	LEU	HD11	-0.65	2.16 – -0.64	-5.0
1	A	32	LEU	HD12	-0.65	2.16 – -0.64	-5.0
1	A	32	LEU	HD13	-0.65	2.16 – -0.64	-5.0

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:

