wwPDB NMR Structure Validation Summary Report (i)

Apr 26, 2016 – 02:30 PM BST

PDB ID	:	1GIW
Title	:	SOLUTION STRUCTURE OF REDUCED HORSE HEART CYTOCHROME C, NMR, MINIMIZED AVERAGE STRUCTURE
Authors	:	Banci, L.; Bertini, I.; Huber, J.G.; Spyroulias, G.A.; Turano, P.
Deposited on	:	1998-06-17

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:

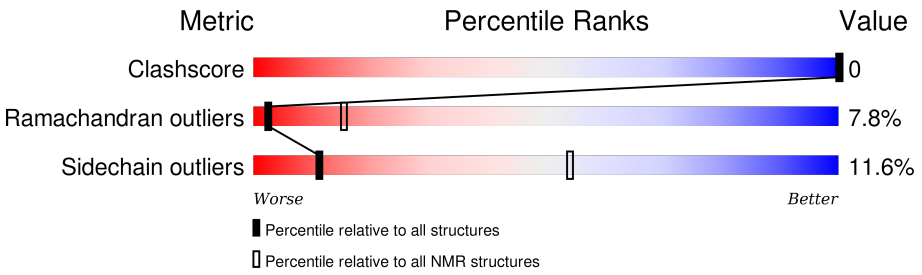
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Mogul	:	1.7.1 (RC1), CSD as537be (2016)
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	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 (i)

The following experimental techniques were used to determine the structure:
SOLUTION NMR

The overall completeness of chemical shifts assignment is 44%.

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	104	79% 20% •

2 Ensemble composition and analysis (i)

This entry contains 1 models. Identification of well-defined residues and clustering analysis are not possible.

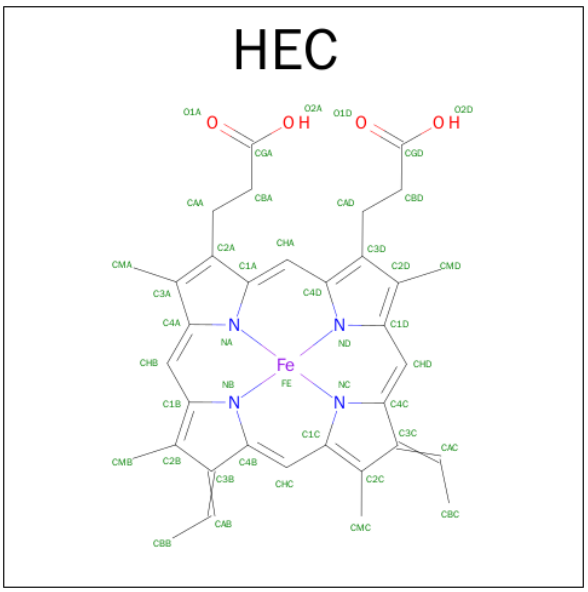
3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 1744 atoms, of which 878 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called CYTOCHROME C.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	104	1669	524	846	144	151	4	0

- Molecule 2 is HEME C (three-letter code: HEC) (formula: $C_{34}H_{34}FeN_4O_4$).



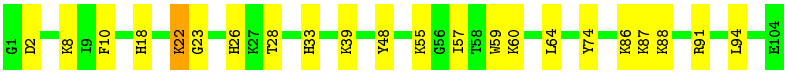
Mol	Chain	Residues	Atoms					
			Total	C	Fe	H	N	O
2	A	1	75	34	1	32	4	4

4 Residue-property plots (i)

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

Chain A:



5 Refinement protocol and experimental data overview (i)

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

Of the 40 calculated structures, 1 were deposited, based on the following criterion: *LEAST RESTRAINT VIOLATION.*

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

Software name	Classification	Version
AMBER	refinement	
DYANA	structure solution	
AMBER	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 4189
Number of chemical shift lists	2
Total number of shifts	709
Number of shifts mapped to atoms	692
Number of unparsed shifts	0
Number of shifts with mapping errors	17
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	44%

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

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: 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	0.56	0/839 (0.0%)	0.97	1/1118 (0.1%)
All	All	0.56	0/839 (0.0%)	0.97	1/1118 (0.1%)

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	4
All	All	0	4

There are no bond-length outliers.

All angle outliers are listed below.

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	33	HIS	CG-ND1-CE1	-6.34	97.45	105.70

There are no chirality outliers.

All planar outliers are listed below.

Mol	Chain	Res	Type	Group
1	A	18	HIS	Sidechain
1	A	91	ARG	Sidechain
1	A	10	PHE	Sidechain
1	A	74	TYR	Sidechain

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	866	878	877	-

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	102/104 (98%)	70 (69%)	24 (24%)	8 (8%)	2 15
All	All	102/104 (98%)	70 (69%)	24 (24%)	8 (8%)	2 15

5 of 8 Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type
1	A	26	HIS
1	A	87	LYS
1	A	23	GLY
1	A	22	LYS
1	A	59	TRP

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	86/86 (100%)	76 (88%)	10 (12%)	11 54
All	All	86/86 (100%)	76 (88%)	10 (12%)	11 54

5 of 10 residues with a non-rotameric sidechain are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type
1	A	64	LEU
1	A	55	LYS
1	A	8	LYS
1	A	94	LEU
1	A	39	LYS

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 carbohydrates in this entry.

6.6 Ligand geometry (i)

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	0.85	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	2.42	2 (10%)

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	-	0,6,54,54	0,0,8,8

There are no bond-length outliers.

All angle outliers are listed below. They are sorted according to the Z-score.

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	105	HEC	CBB-CAB-C3B	6.88	112.29	127.34
2	A	105	HEC	CBC-CAC-C3C	5.79	114.68	127.34

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 (i)

The completeness of assignment taking into account all chemical shift lists is 44% for the well-defined parts and 44% for the entire structure.

7.1 Chemical shift list 1

File name: BMRB entry 4189

Chemical shift list name: *assigned_ chem_ shift_ list_ 1*

7.1.1 Bookkeeping (i)

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	692
Number of shifts mapped to atoms	692
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	27

7.1.2 Chemical shift referencing (i)

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

7.1.3 Completeness of resonance assignments (i)

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 44%, i.e. 573 atoms were assigned a chemical shift out of a possible 1308. 0 out of 9 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}**H**	^{13}**C**	^{15}**N**
Backbone	201/512 (39%)	201/204 (99%)	0/208 (0%)	0/100 (0%)
Sidechain	321/695 (46%)	321/411 (78%)	0/251 (0%)	0/33 (0%)
Aromatic	51/101 (50%)	51/54 (94%)	0/43 (0%)	0/4 (0%)
Overall	573/1308 (44%)	573/669 (86%)	0/502 (0%)	0/137 (0%)

7.1.4 Statistically unusual chemical shifts (i)

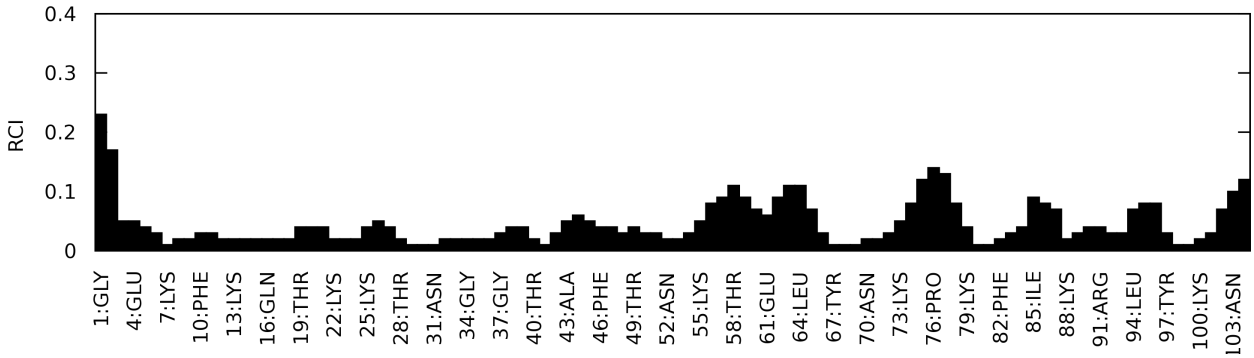
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

Mol	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	67	TYR	HE1	0.10	7.86 – 5.56	-28.7
1	A	67	TYR	HE2	0.43	7.86 – 5.56	-27.3
1	A	67	TYR	HD1	0.73	8.44 – 5.44	-20.7
1	A	67	TYR	HD2	0.94	8.44 – 5.44	-20.0
1	A	80	MET	HG3	-3.77	4.30 – 0.50	-16.2
1	A	18	HIS	HD2	0.12	9.28 – 4.78	-15.4
1	A	18	HIS	HE1	0.41	10.53 – 5.43	-14.8
1	A	80	MET	HB3	-2.64	3.70 – 0.30	-13.6
1	A	80	MET	HG2	-1.89	4.23 – 0.63	-12.0
1	A	29	GLY	HA2	-0.20	5.87 – 2.07	-11.0
1	A	80	MET	HE2	-3.30	4.28 – -0.52	-10.8
1	A	80	MET	HE3	-3.30	4.28 – -0.52	-10.8
1	A	80	MET	HE1	-3.30	4.28 – -0.52	-10.8
1	A	30	PRO	HD2	1.08	5.45 – 1.85	-7.1
1	A	46	PHE	HD1	5.01	8.56 – 5.56	-6.8
1	A	82	PHE	HB2	0.55	4.85 – 1.15	-6.6
1	A	80	MET	HB2	-0.22	3.73 – 0.33	-6.6
1	A	41	GLY	HA3	1.50	5.80 – 2.00	-6.3
1	A	18	HIS	HB3	0.78	5.00 – 1.10	-5.8
1	A	30	PRO	HD3	1.45	5.52 – 1.72	-5.7
1	A	18	HIS	HB2	1.07	4.91 – 1.31	-5.7
1	A	32	LEU	HD11	-0.76	2.16 – -0.64	-5.4
1	A	32	LEU	HD12	-0.76	2.16 – -0.64	-5.4
1	A	32	LEU	HD13	-0.76	2.16 – -0.64	-5.4
1	A	17	CYS	HB2	0.57	5.20 – 0.70	-5.3
1	A	97	TYR	HE1	5.52	7.86 – 5.56	-5.2
1	A	46	PHE	HB2	1.12	4.85 – 1.15	-5.1

7.1.5 Random Coil Index (RCI) plots (i)

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 4189

Chemical shift list name: *assigned_chem_shift_list_2*

7.2.1 Bookkeeping (i)

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	17
Number of shifts mapped to atoms	0
Number of unparsed shifts	0
Number of shifts with mapping errors	17
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.

- Chain not found in structure. First 5 (of 17) occurences are reported below.

Chain	Res	Type	Atom	Shift Data		
				Value	Uncertainty	Ambiguity
UNMAPPED	1	HEC	QM1	3.5	-1.0	1
UNMAPPED	1	HEC	QM8	2.16	-1.0	1
UNMAPPED	1	HEC	HP64	3.93	-1.0	1
UNMAPPED	1	HEC	HGM	9.59	-1.0	1
UNMAPPED	1	HEC	HP61	3.46	-1.0	1

7.2.2 Chemical shift referencing (i)

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

7.2.3 Completeness of resonance assignments (i)

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 0%, i.e. 0 atoms were assigned a chemical shift out of a possible 1308. 0 out of 9 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	0/512 (0%)	0/204 (0%)	0/208 (0%)	0/100 (0%)
Sidechain	0/695 (0%)	0/411 (0%)	0/251 (0%)	0/33 (0%)
Aromatic	0/101 (0%)	0/54 (0%)	0/43 (0%)	0/4 (0%)
Overall	0/1308 (0%)	0/669 (0%)	0/502 (0%)	0/137 (0%)

7.2.4 Statistically unusual chemical shifts (i)

There are no statistically unusual chemical shifts.

7.2.5 Random Coil Index (RCI) plots (i)

No *random coil index* (RCI) plot could be generated from the current chemical shift list (assigned_chem_shift_list_2). RCI is only applicable to proteins.