



wwPDB NMR Structure Validation Summary Report ⓘ

Apr 26, 2016 – 11:38 PM BST

PDB ID : 2K3V
Title : Solution Structure of a Tetrahaem Cytochrome from Shewanella Frigidimarina
Authors : Paixao, V.B.; Turner, D.L.; Salgueiro, C.A.; Brennan, L.; Reid, G.A.; Chapman, S.K.
Deposited on : 2008-05-19

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

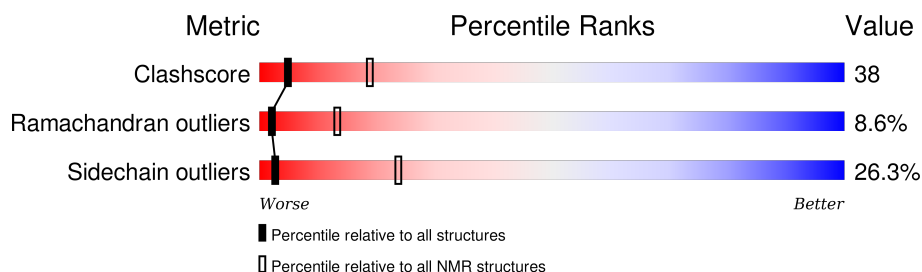
1 Overall quality at a glance

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	86	

2 Ensemble composition and analysis ⓘ

This entry contains 20 models. Model 3 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *lowest energy*.

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:79, A:83-A:86 (81)	0.15	3

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, 7, 8, 9, 10, 11, 12, 13, 14, 18, 20
2	6, 15, 16, 17, 19
Single-model clusters	5

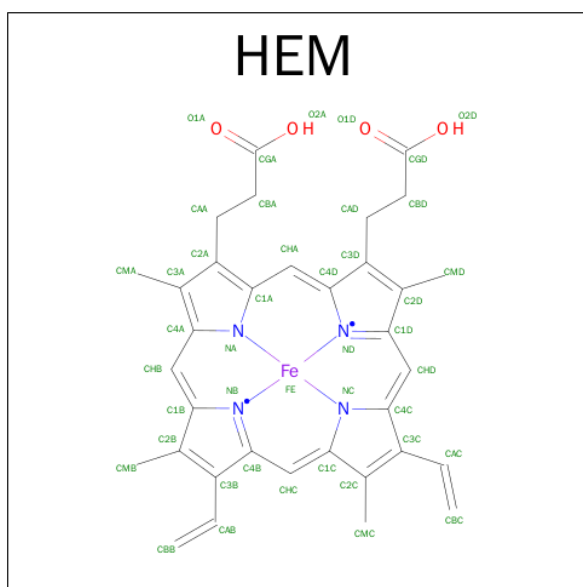
3 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 1376 atoms, of which 561 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Tetraheme cytochrome c-type.

Mol	Chain	Residues	Atoms						Trace
1	A	86	Total	C	H	N	O	S	0
			1204	380	561	114	138	11	

- Molecule 2 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (formula: $C_{34}H_{32}FeN_4O_4$).



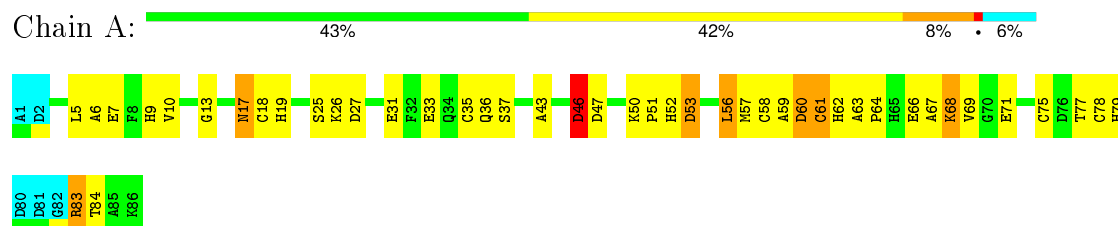
Mol	Chain	Residues	Atoms				
2	A	1	Total	C	Fe	N	O
			43	34	1	4	4
2	A	1	Total	C	Fe	N	O
			43	34	1	4	4
2	A	1	Total	C	Fe	N	O
			43	34	1	4	4
2	A	1	Total	C	Fe	N	O
			43	34	1	4	4

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.

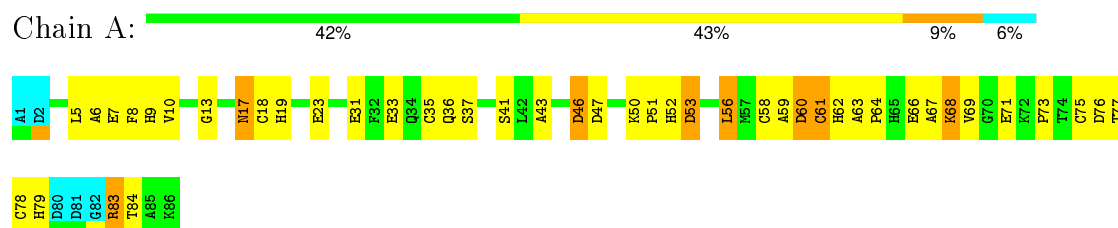
- Molecule 1: Tetraheme cytochrome c-type



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

The representative model is number 3. Colouring as in section 4.1 above.

- Molecule 1: Tetraheme cytochrome c-type



5 Refinement protocol and experimental data overview

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

Of the 600 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
PARADYANA	structure solution	
PARADYANA	refinement	

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 15765
Number of chemical shift lists	1
Total number of shifts	551
Number of shifts mapped to atoms	474
Number of unparsed shifts	57
Number of shifts with mapping errors	20
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: HEM

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	610	538	530	41±3
2	A	172	0	120	24±3
All	All	15640	10760	13000	1095

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 38.

5 of 191 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:43:ALA:HB2	1:A:53:ASP:OD1	0.92	1.63	20	8
1:A:43:ALA:HB2	1:A:53:ASP:OD2	0.90	1.66	8	7
2:A:218:HEM:CGD	2:A:218:HEM:HMD1	0.88	1.99	11	6
1:A:56:LEU:HD12	1:A:60:ASP:HB3	0.75	1.57	8	4
2:A:238:HEM:HMA2	2:A:238:HEM:HBA2	0.71	1.61	8	5

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	80/86 (93%)	54±1 (68±2%)	19±2 (24±2%)	7±1 (9±1%)	2	13
All	All	1600/1720 (93%)	1084 (68%)	379 (24%)	137 (9%)	2	13

5 of 12 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	60	ASP	20
1	A	46	ASP	20
1	A	66	GLU	20
1	A	53	ASP	20
1	A	13	GLY	20

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	66/69 (96%)	49±2 (74±3%)	17±2 (26±3%)	3	24
All	All	1320/1380 (96%)	973 (74%)	347 (26%)	3	24

5 of 34 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	46	ASP	20
1	A	17	ASN	20
1	A	5	LEU	20
1	A	83	ARG	20
1	A	61	CYS	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 ⓘ

4 ligands are 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	HEM	A	218	1	24,50,50	1.70±0.00	0±0 (0±0%)
2	HEM	A	238	1	24,50,50	1.70±0.01	0±0 (0±0%)
2	HEM	A	261	1	24,50,50	1.70±0.00	0±0 (0±0%)
2	HEM	A	278	1	24,50,50	1.70±0.00	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	HEM	A	218	1	16,82,82	2.45±0.00	2±0 (12±0%)
2	HEM	A	238	1	16,82,82	2.45±0.00	2±0 (12±0%)

Mol	Type	Chain	Res	Link	Counts	Bond angles	
						RMSZ	#Z>2
2	HEM	A	261	1	16,82,82	2.45±0.00	2±0 (12±0%)
2	HEM	A	278	1	16,82,82	2.45±0.00	2±0 (12±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	HEM	A	218	1	-	0±0,6,54,54	0±0,0,8,8
2	HEM	A	238	1	-	0±0,6,54,54	0±0,0,8,8
2	HEM	A	261	1	-	0±0,6,54,54	0±0,0,8,8
2	HEM	A	278	1	-	0±0,6,54,54	0±0,0,8,8

There are no bond-length outliers.

5 of 8 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	218	HEM	C3C-CAC-CBC	6.19	113.95	126.40	13	20
2	A	261	HEM	C3B-CAB-CBB	6.19	113.95	126.40	3	20
2	A	238	HEM	C3B-CAB-CBB	6.19	113.95	126.40	11	20
2	A	238	HEM	C3C-CAC-CBC	6.19	113.95	126.40	12	20
2	A	261	HEM	C3C-CAC-CBC	6.19	113.95	126.40	8	20

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 15765

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	551
Number of shifts mapped to atoms	474
Number of unparsed shifts	57
Number of shifts with mapping errors	20
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	40

The following errors were found when reading this chemical shift list.

- Chemical shift has been reported more than once. First 5 (of 57) occurrences are reported below.

Shift ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
494	UNMAPPED	1	HEM_red	HA62	3.444	0.005	2
495	UNMAPPED	1	HEM_red	HA63	4.557	0.006	2
496	UNMAPPED	1	HEM_red	HA72	2.711	0.004	2
497	UNMAPPED	1	HEM_red	HA73	3.710	0.005	2
498	UNMAPPED	1	HEM_red	HAM	9.447	0.002	1

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

- Chain not found in structure. First 5 (of 20) occurrences are reported below.

Chain	Res	Type	Atom	Shift Data		
				Value	Uncertainty	Ambiguity
UNMAPPED	1	HEM_red	HA63	4.461	0.003	2
UNMAPPED	1	HEM_red	QM1	3.583	0.002	1

Continued on next page...

Continued from previous page...

Chain	Res	Type	Atom	Shift Data		
				Value	Uncertainty	Ambiguity
UNMAPPED	1	HEM_red	HGM	9.669	0.001	1
UNMAPPED	1	HEM_red	HT4A	5.755	0.005	1
UNMAPPED	1	HEM_red	HAM	8.987	0.002	1

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 46%, i.e. 418 atoms were assigned a chemical shift out of a possible 900. 0 out of 6 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	¹ H	¹³ C	¹⁵ N
Backbone	158/397 (40%)	158/158 (100%)	0/162 (0%)	0/77 (0%)
Sidechain	222/421 (53%)	222/250 (89%)	0/159 (0%)	0/12 (0%)
Aromatic	38/82 (46%)	38/46 (83%)	0/28 (0%)	0/8 (0%)
Overall	418/900 (46%)	418/454 (92%)	0/349 (0%)	0/97 (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	19	HIS	HD2	0.14	9.28 – 4.78	-15.3
1	A	49	HIS	HD2	0.15	9.28 – 4.78	-15.3
1	A	19	HIS	HE1	0.87	10.53 – 5.43	-13.9
1	A	79	HIS	HD2	0.77	9.28 – 4.78	-13.9
1	A	65	HIS	HD2	0.84	9.28 – 4.78	-13.8
1	A	62	HIS	HD2	0.85	9.28 – 4.78	-13.7
1	A	52	HIS	HD2	0.90	9.28 – 4.78	-13.6
1	A	65	HIS	HE1	1.16	10.53 – 5.43	-13.4
1	A	79	HIS	HE1	1.22	10.53 – 5.43	-13.3
1	A	49	HIS	HE1	1.23	10.53 – 5.43	-13.2
1	A	9	HIS	HD2	1.10	9.28 – 4.78	-13.2
1	A	9	HIS	HE1	1.31	10.53 – 5.43	-13.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atom	Shift, ppm	Expected range, ppm	Z-score
1	A	62	HIS	HE1	1.32	10.53 – 5.43	-13.1
1	A	52	HIS	HE1	1.32	10.53 – 5.43	-13.1
1	A	39	HIS	HD2	1.27	9.28 – 4.78	-12.8
1	A	39	HIS	HE1	1.55	10.53 – 5.43	-12.6
1	A	6	ALA	HB3	-1.58	2.61 – 0.11	-11.8
1	A	6	ALA	HB1	-1.58	2.61 – 0.11	-11.8
1	A	6	ALA	HB2	-1.58	2.61 – 0.11	-11.8
1	A	73	PRO	HG2	-1.12	3.48 – 0.38	-9.9
1	A	51	PRO	HB3	-0.90	3.81 – 0.21	-8.1
1	A	24	PRO	HG3	-0.67	3.56 – 0.26	-7.8
1	A	24	PRO	HB3	-0.72	3.81 – 0.21	-7.6
1	A	64	PRO	HB2	-0.45	3.82 – 0.32	-7.2
1	A	83	ARG	HG2	-0.34	2.92 – 0.22	-7.1
1	A	19	HIS	HB2	0.57	4.91 – 1.31	-7.0
1	A	49	HIS	HA	1.56	6.81 – 2.41	-6.9
1	A	45	MET	HB3	-0.22	3.70 – 0.30	-6.5
1	A	24	PRO	HA	2.26	6.05 – 2.75	-6.5
1	A	19	HIS	HB3	0.80	5.00 – 1.10	-5.8
1	A	83	ARG	HD2	1.80	4.27 – 1.97	-5.7
1	A	83	ARG	HD3	1.69	4.36 – 1.86	-5.7
1	A	49	HIS	HB2	1.12	4.91 – 1.31	-5.5
1	A	49	HIS	HB3	0.93	5.00 – 1.10	-5.4
1	A	65	HIS	HB2	1.15	4.91 – 1.31	-5.4
1	A	73	PRO	HG3	0.12	3.56 – 0.26	-5.4
1	A	79	HIS	HB2	1.18	4.91 – 1.31	-5.4
1	A	51	PRO	HD2	1.76	5.45 – 1.85	-5.2
1	A	18	CYS	HB2	0.63	5.20 – 0.70	-5.2
1	A	5	LEU	HG	-0.15	3.16 – -0.14	-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:

