



# Full wwPDB NMR Structure Validation Report ⓘ

May 28, 2020 – 11:43 pm BST

PDB ID : 2MHL  
Title : NMR solution Structure of the E.coli Outer Membrane Protein W  
Authors : Horst, R.; Stanczak, P.; Wuthrich, K.  
Deposited on : 2013-11-26

This is a Full wwPDB NMR Structure Validation Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto: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 ⓘ 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  
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 : 2.11  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.11



## 2 Ensemble composition and analysis i

This entry contains 20 models. Model 2 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 target function*.

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:2-A:18, A:35-A:64, A:75-A:115, A:124-A:192 (157)	1.69	2

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 3 clusters and 1 single-model cluster was found.

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

### 3 Entry composition

There is only 1 type of molecule in this entry. The entry contains 2697 atoms, of which 1215 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Outer membrane protein W.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	192	2697	945	1215	252	278	7	0

There is a discrepancy between the modelled and reference sequences:

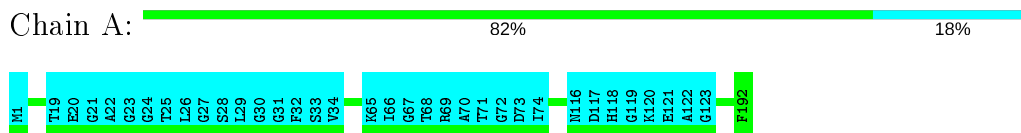
Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	EXPRESSION TAG	UNP B1XBK5

## 4 Residue-property plots

### 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: Outer membrane protein W

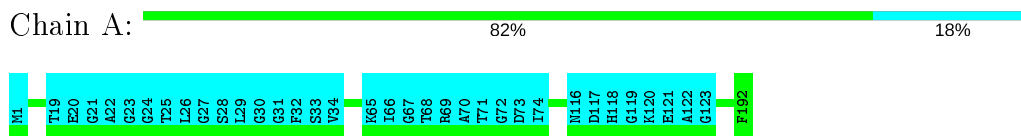


### 4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

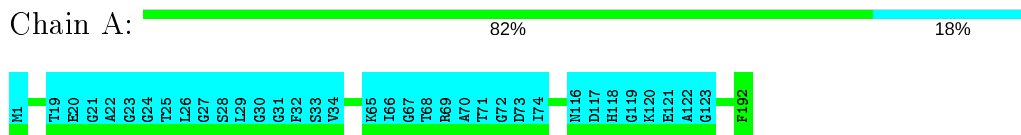
#### 4.2.1 Score per residue for model 1

- Molecule 1: Outer membrane protein W



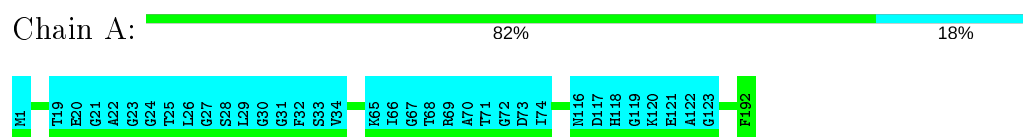
#### 4.2.2 Score per residue for model 2 (medoid)

- Molecule 1: Outer membrane protein W



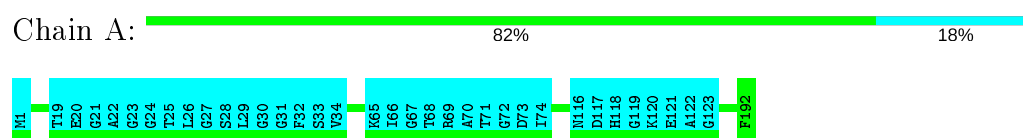
### 4.2.3 Score per residue for model 3

- Molecule 1: Outer membrane protein W



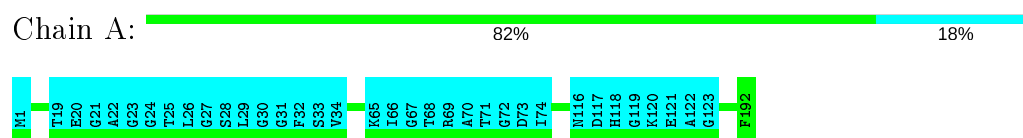
### 4.2.4 Score per residue for model 4

- Molecule 1: Outer membrane protein W



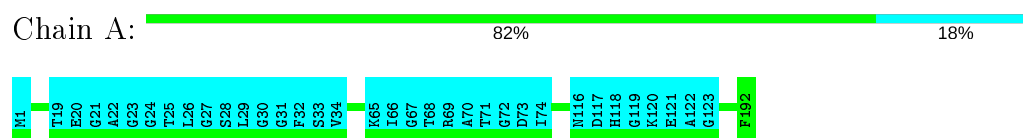
### 4.2.5 Score per residue for model 5

- Molecule 1: Outer membrane protein W



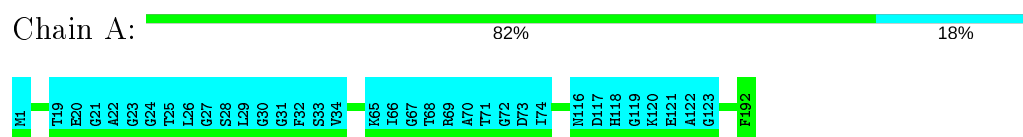
### 4.2.6 Score per residue for model 6

- Molecule 1: Outer membrane protein W



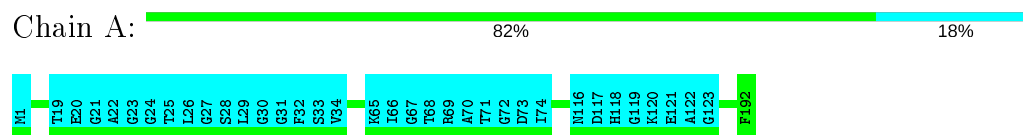
### 4.2.7 Score per residue for model 7

- Molecule 1: Outer membrane protein W



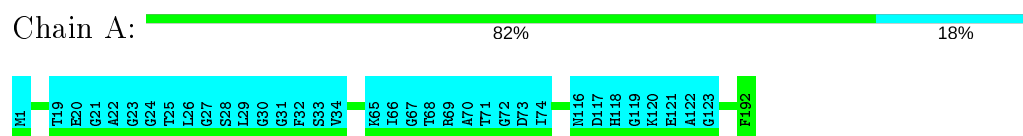
#### 4.2.8 Score per residue for model 8

- Molecule 1: Outer membrane protein W



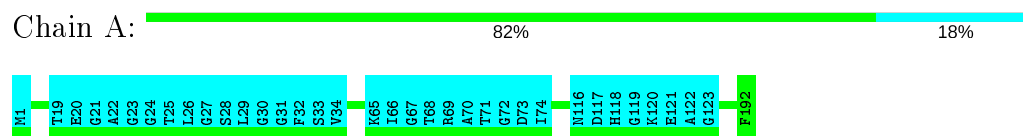
#### 4.2.9 Score per residue for model 9

- Molecule 1: Outer membrane protein W



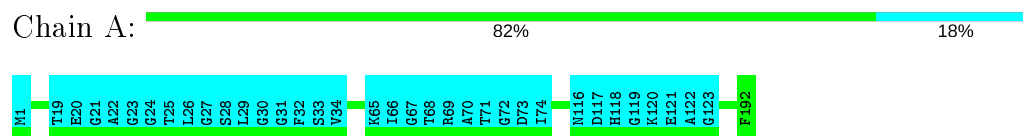
#### 4.2.10 Score per residue for model 10

- Molecule 1: Outer membrane protein W



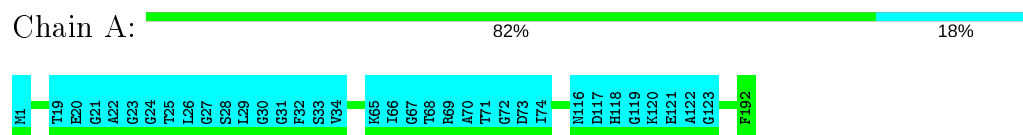
#### 4.2.11 Score per residue for model 11

- Molecule 1: Outer membrane protein W



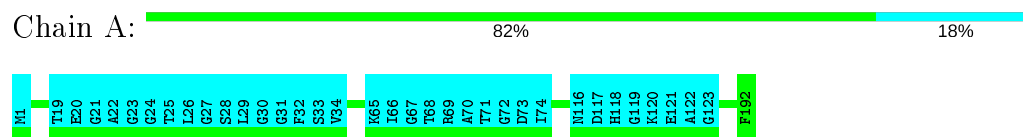
#### 4.2.12 Score per residue for model 12

- Molecule 1: Outer membrane protein W



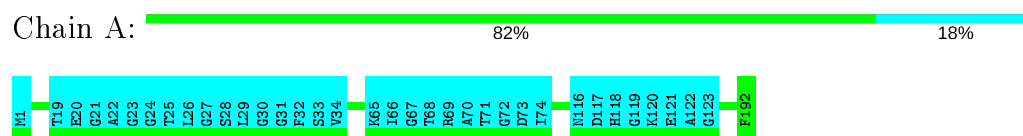
#### 4.2.13 Score per residue for model 13

- Molecule 1: Outer membrane protein W



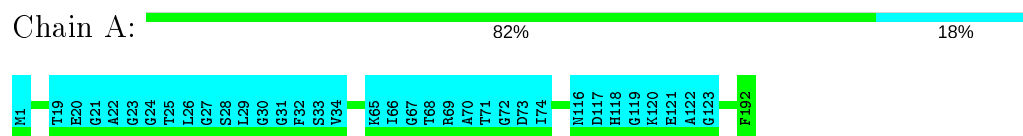
#### 4.2.14 Score per residue for model 14

- Molecule 1: Outer membrane protein W



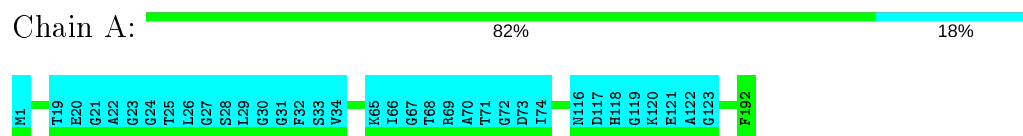
#### 4.2.15 Score per residue for model 15

- Molecule 1: Outer membrane protein W



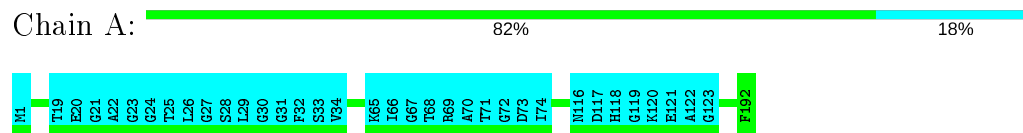
#### 4.2.16 Score per residue for model 16

- Molecule 1: Outer membrane protein W



#### 4.2.17 Score per residue for model 17

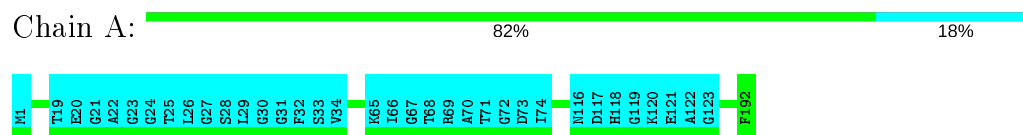
- Molecule 1: Outer membrane protein W





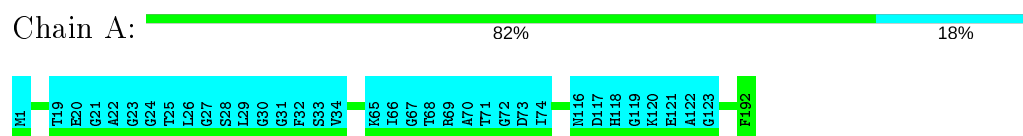
#### 4.2.18 Score per residue for model 18

- Molecule 1: Outer membrane protein W



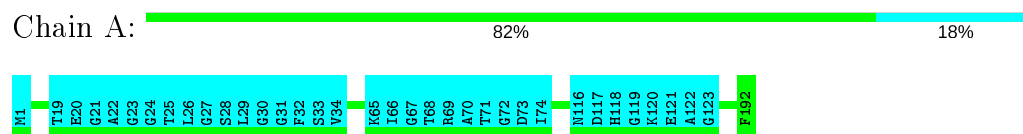
#### 4.2.19 Score per residue for model 19

- Molecule 1: Outer membrane protein W



#### 4.2.20 Score per residue for model 20

- Molecule 1: Outer membrane protein W



## 5 Refinement protocol and experimental data overview

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

Of the 200 calculated structures, 20 were deposited, based on the following criterion: *20 structures for lowest target function*.

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

Software name	Classification	Version
CYANA	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)	input_cs.cif
Number of chemical shift lists	1
Total number of shifts	824
Number of shifts mapped to atoms	824
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	36%

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](#)

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 [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	24960	20340	0	-

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	0	-	-	-	-
All	All	0	-	-	-	-

There are no Ramachandran outliers.

### 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	0	-	-	-
All	All	0	-	-	-

There are no protein residues with a non-rotameric sidechain to report.

### 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](#)

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](#)

The completeness of assignment taking into account all chemical shift lists is 36% for the well-defined parts and 37% for the entire structure.

### 7.1 Chemical shift list 1

File name: input\_cs.cif

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	824
Number of shifts mapped to atoms	824
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	0

#### 7.1.2 Chemical shift referencing [i](#)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction $\pm$ precision, ppm	Suggested action
$^{13}\text{C}_\alpha$	174	1.23 $\pm$ 0.13	Should be applied
$^{13}\text{C}_\beta$	144	0.71 $\pm$ 0.16	Should be applied
$^{13}\text{C}'$	152	0.76 $\pm$ 0.13	Should be applied
$^{15}\text{N}$	174	0.83 $\pm$ 0.30	Should be applied

#### 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 36%, i.e. 672 atoms were assigned a chemical shift out of a possible 1890. 0 out of 23 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	$^1\text{H}$	$^{13}\text{C}$	$^{15}\text{N}$
Backbone	545/773 (71%)	141/308 (46%)	263/314 (84%)	141/151 (93%)
Sidechain	121/850 (14%)	0/497 (0%)	121/316 (38%)	0/37 (0%)

*Continued on next page...*

Continued from previous page...

	Total	<sup>1</sup> H	<sup>13</sup> C	<sup>15</sup> N
Aromatic	6/267 (2%)	3/142 (2%)	0/115 (0%)	3/10 (30%)
Overall	672/1890 (36%)	144/947 (15%)	384/745 (52%)	144/198 (73%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 37%, i.e. 824 atoms were assigned a chemical shift out of a possible 2238. 0 out of 26 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	<sup>1</sup> H	<sup>13</sup> C	<sup>15</sup> N
Backbone	674/948 (71%)	174/378 (46%)	326/384 (85%)	174/186 (94%)
Sidechain	144/1007 (14%)	0/587 (0%)	144/377 (38%)	0/43 (0%)
Aromatic	6/283 (2%)	3/151 (2%)	0/121 (0%)	3/11 (27%)
Overall	824/2238 (37%)	177/1116 (16%)	470/882 (53%)	177/240 (74%)

#### 7.1.4 Statistically unusual chemical shifts [i](#)

There are no statistically unusual chemical shifts.

#### 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:

