wwPDB NMR Structure Validation Summary Report (i)

Jun 16, 2024 – 11:02 PM EDT

PDB ID	:	5OWI
BMRB ID	:	27242
Title	:	The dynamic dimer structure of the chaperone Trigger Factor (conformer 1)
Authors	:	Morgado, L.; Burmann, B.M.; Sharpe, T.; Mazur, A.; Hiller, S.
Deposited on	:	2017-09-01

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
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:

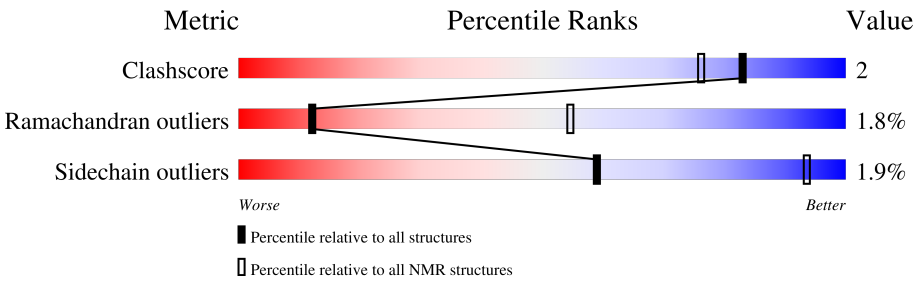
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 5%.

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	432	93% 6% •
1	B	432	93% 5% •

2 Ensemble composition and analysis (i)

This entry contains 10 models. Model 10 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:1-A:252, A:259-A:432, B:1-B:252, B:259-B:432 (852)	1.83	10

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, 4, 5, 8, 10
2	2, 6, 7, 9
Single-model clusters	3

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 13582 atoms, of which 6810 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Trigger factor.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	432	6791	2119	3405	582	674	11	0
1	B	432	6791	2119	3405	582	674	11	0

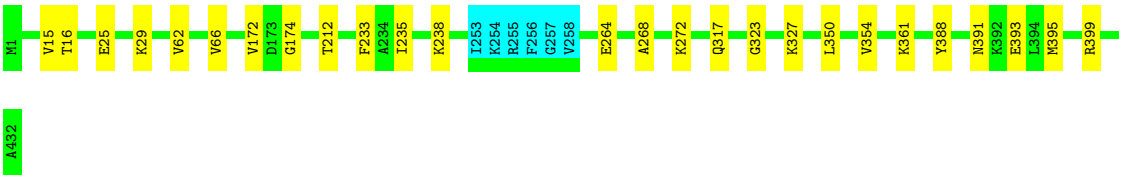
4 Residue-property plots (i)

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.

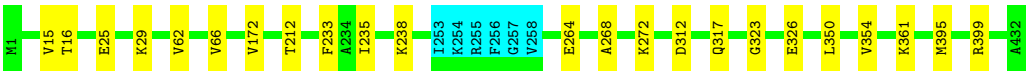
• Molecule 1: Trigger factor

Chain A: 93% 6% •



• Molecule 1: Trigger factor

Chain B: 93% 5% •

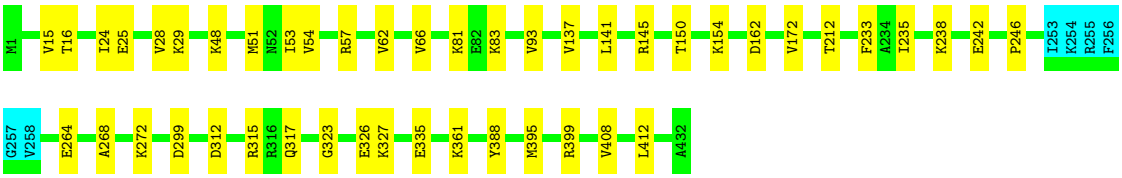


4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 10. Colouring as in section 4.1 above.

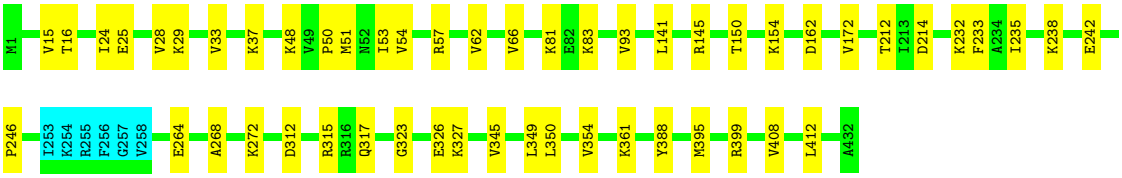
• Molecule 1: Trigger factor

Chain A: 88% 11% •



• Molecule 1: Trigger factor

Chain B: 87% 12% •


M1
V15
T16
I24
E25
V28
K29
V33
K37
K48
V49
P50
M51
N52
I53
V54
R57
V62
V66
K81
E82
K83
V93
L141
R145
T150
K154
D162
V172
T212
I213
D214
K232
F233
A234
I235
K238
E242
P246
I253
K254
R255
F256
G257
V258
E264
A268
K272
D312
R315
R316
Q317
G323
E326
K327
V345
L349
L350
V354
K361
Y388
M395
R399
V408
L412
A432

5 Refinement protocol and experimental data overview (i)

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

Of the 100 calculated structures, 10 were deposited, based on the following criterion: *structures with the lowest energy*.

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

Software name	**Classification**	**Version**
HADDOCK	structure calculation	
Xplor-NIH	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)	working_cs.cif
Number of chemical shift lists	1
Total number of shifts	564
Number of shifts mapped to atoms	564
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	5%

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
1	A	3336	3347	3345	17±3
1	B	3336	3347	3345	17±4
All	All	66720	66940	66900	316

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 2.

5 of 157 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:397:ASN:ND2	1:B:38:LYS:HA	0.72	2.00	8	2
1:A:57:ARG:CZ	1:B:193:ARG:HG2	0.66	2.21	2	1
1:A:38:LYS:HA	1:B:397:ASN:ND2	0.65	2.06	8	1
1:B:25:GLU:O	1:B:29:LYS:HG3	0.64	1.93	10	7
1:A:25:GLU:O	1:A:29:LYS:HG3	0.64	1.93	10	7

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	424/432 (98%)	397±3 (94±1%)	19±2 (5±1%)	8±2 (2±0%)	12	54
1	B	424/432 (98%)	397±2 (94±0%)	20±2 (5±0%)	7±1 (2±0%)	13	56
All	All	8480/8640 (98%)	7945 (94%)	386 (5%)	149 (2%)	12	54

5 of 39 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	16	THR	10
1	A	235	ILE	10
1	B	16	THR	10
1	B	235	ILE	10
1	A	172	VAL	9

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	354/359 (99%)	347±2 (98±1%)	7±2 (2±1%)	57	93
1	B	354/359 (99%)	347±2 (98±1%)	7±2 (2±1%)	59	93
All	All	7080/7180 (99%)	6943 (98%)	137 (2%)	59	93

5 of 46 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	317	GLN	10
1	B	317	GLN	10
1	A	264	GLU	8
1	B	264	GLU	8
1	A	361	LYS	8

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)

The completeness of assignment taking into account all chemical shift lists is 5% for the well-defined parts and 5% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *TFconv.str*

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	564
Number of shifts mapped to atoms	564
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 ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	108	0.72 ± 0.08	Should be checked
$^{13}C_{\beta}$	93	1.13 ± 0.14	Should be checked
$^{13}C'$	0	—	None (insufficient data)
^{15}N	181	-0.50 ± 0.31	None needed (< 0.5 ppm)

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 5%, i.e. 562 atoms were assigned a chemical shift out of a possible 11636. 0 out of 144 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	468/4274 (11%)	180/1742 (10%)	107/1704 (6%)	181/828 (22%)
Sidechain	92/6842 (1%)	0/4410 (0%)	92/2150 (4%)	0/282 (0%)

Continued on next page...

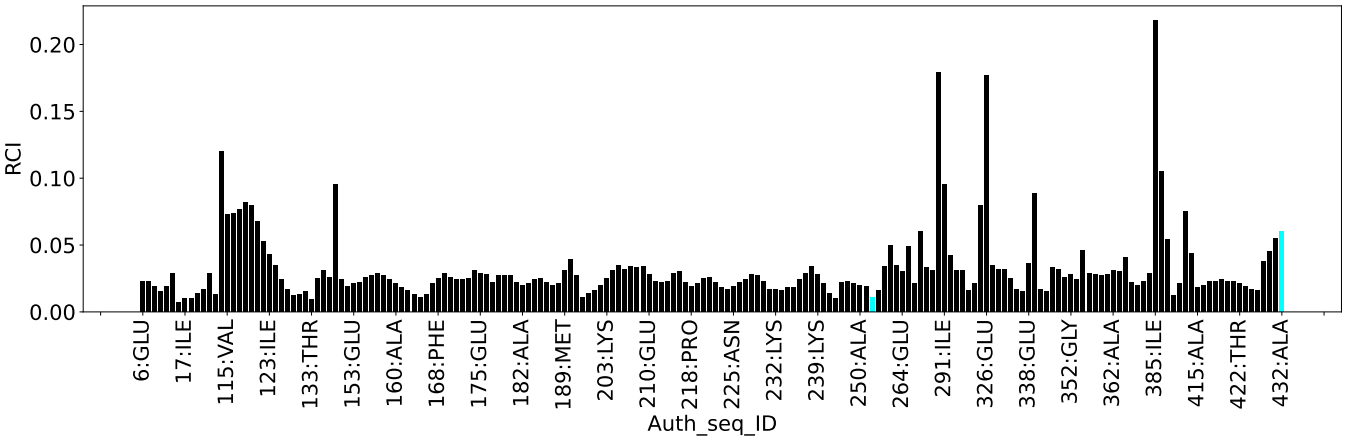
Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Aromatic	2/520 (0%)	1/252 (0%)	0/258 (0%)	1/10 (10%)
Overall	562/11636 (5%)	181/6404 (3%)	199/4112 (5%)	182/1120 (16%)

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.If well-defined core and ill-defined regions are not identified then it is shown as gray bars.

Random coil index (RCI) for chain A: