wwPDB NMR Structure Validation Summary Report (i)

Feb 20, 2022 – 02:36 AM EST

PDB ID	:	1U3N
Title	:	A SOD-like protein from B. subtilis, unstructured in solution, becomes ordered in the crystal: implications for function and for fibrillogenesis
Authors	:	Banci, L.; Bertini, I.; Calderone, V.; Cramaro, F.; Del Conte, R.; Fantoni, A.; Mangani, S.; Quattrone, A.; Viezzoli, M.S.
Deposited on	:	2004-07-22

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 following versions of software and data (see references (i)) were used in the production of this report:

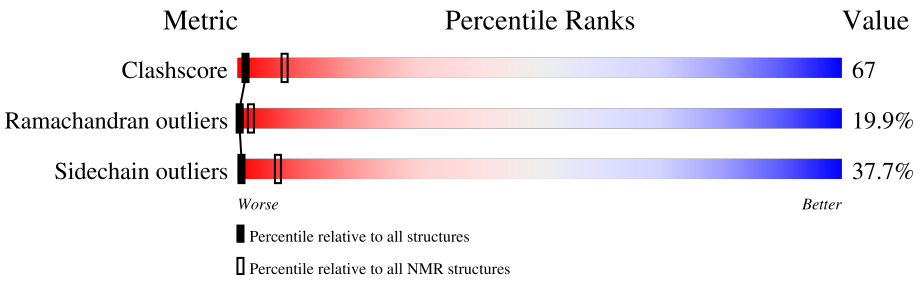
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.26
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.26

1 Overall quality at a glance (i)

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

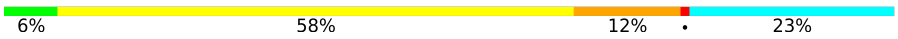
The overall completeness of chemical shifts assignment was not calculated.

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	162	

2 Ensemble composition and analysis (i)

This entry contains 30 models. Model 29 is the overall representative, medoid model (most similar to other models).

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:40-A:105, A:123-A:147, A:152-A:171, A:178-A:190 (124)	1.71	29

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 5 clusters and 5 single-model clusters were found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 2377 atoms, of which 1165 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Hypothetical superoxide dismutase-like protein yojM.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	162	2377	750	1165	215	243	4	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: Hypothetical superoxide dismutase-like protein yojM

Chain A: 6% 58% 12% • 23%

V35 V36 E37 T38 S39 A40 F41 G42 H43 H44 V45 Q46 L47 V48 N49 R50 E51 G52 K53 A54 V55 G56 F57 I58 E59 I60 K61 E62 S63 D64 D65 E66 G67 L68 D69 I70 H71 I72 S73 A74 N75 S76 L77 R78 P79 G80 A81 S82 L83 G84 F85 H86 I87 Y88 E89 K90 G91 S92 C93 V94

R95 P96 D97 F98 E99 S100 F105 N106 P107 L108 N109 K110 E111 H112 G113 F114 N115 N116 P117 M118 G119 H120 H121 A122 G123 D124 L125 L128 E129 V130 G131 A132 D133 G134 K135 V136 D137 V138 I139 M140 N141 A142 P143 D144 T145 S146 L147 K148 K149 G150 S151 K152 L153 N154 I155 L156 D157 E158

D159 G160 S161 A162 F163 I164 I165 H166 E167 Q168 A169 D170 D171 Y172 L173 T174 N175 P176 S177 G178 N179 S180 G181 A182 R183 I184 V185 C186 G187 A188 L189 L190 G191 N192 N193 E194 K195 Q196

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

The representative model is number 29. Colouring as in section 4.1 above.

• Molecule 1: Hypothetical superoxide dismutase-like protein yojM

Chain A: 22% 28% 24% • 23%

V35 V36 E37 T38 S39 A40 F41 G42 H43 H44 V45 Q46 L47 V48 N49 R50 K53 A54 V55 I58 E59 I60 K61 E62 S63 D64 D65 E66 G67 L68 D69 I70 H71 I72 S73 A74 N75 S76 L77 R78 P79 G80 A81 S82 L83 G84 F85 H86 I87 Y88 E89 K90 G91 S92 C93 V94 R95 P96

D97 F98 E99 N106 P107 L108 N109 K110 E111 H112 G113 F114 N115 N116 P117 M118 G119 H120 H121 A122 G123 D124 L125 P126 N127 L128 E129 V130 G131 A132 D133 G134 K135 V136 D137 V138 I139 M140 N141 A142 P143 D144 T145 K148 K149 G150 S151 K152 L153 N154 I155 L156 D157 E158 D159 G160 E167

Q168 A169 D170 D171 Y172 L173 T174 N175 P176 S177 G178 N179 S180 G181 A182 R183 I184 V185 C186 G187 A188 L189 L190 G191 N192 N193 E194 K195 Q196

5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *the structure of SOD-like from bacillus subtilis was determined using triple resonance three- and bi-dimensional and homonuclear bi-dimensional NMR technique.*

Of the ? calculated structures, 30 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
DYANA	refinement	

No chemical shift data was provided.

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	919	891	879	120±17
All	All	27570	26730	26370	3609

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 67.

5 of 2072 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:153:LEU:HD13	1:A:156:LEU:HD12	1.11	1.12	21	1
1:A:60:ILE:HG23	1:A:72:ILE:HD11	1.09	1.21	23	1
1:A:163:PHE:CE2	1:A:165:ILE:HD11	1.06	1.86	15	1
1:A:72:ILE:HG21	1:A:163:PHE:CE1	1.06	1.86	23	1
1:A:85:PHE:CE2	1:A:128:LEU:HD13	1.03	1.88	27	1

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	124/162 (77%)	58±5 (47±4%)	41±5 (33±4%)	25±5 (20±4%)	0 2
All	All	3720/4860 (77%)	1735 (47%)	1243 (33%)	742 (20%)	0 2

5 of 89 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	130	VAL	26
1	A	78	ARG	25
1	A	67	GLY	19
1	A	53	LYS	19
1	A	141	ASN	18

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	97/130 (75%)	60±5 (62±5%)	37±5 (38±5%)	1 7
All	All	2910/3900 (75%)	1813 (62%)	1097 (38%)	1 7

5 of 91 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	61	LYS	30
1	A	65	ASP	27
1	A	68	LEU	23
1	A	139	ILE	23
1	A	186	CYS	23

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)

No chemical shift data were provided