wwPDB NMR Structure Validation Summary Report (i)

Apr 26, 2016 – 02:37 PM BST

PDB ID	:	1G90
Title	:	NMR Solution Structure of Outer Membrane Protein A Transmembrane Domain: 10 conformers
Authors	:	Arora, A.; Abildgaard, F.; Bushweller, J.H.; Tamm, L.K.
Deposited on	:	2000-11-21

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:

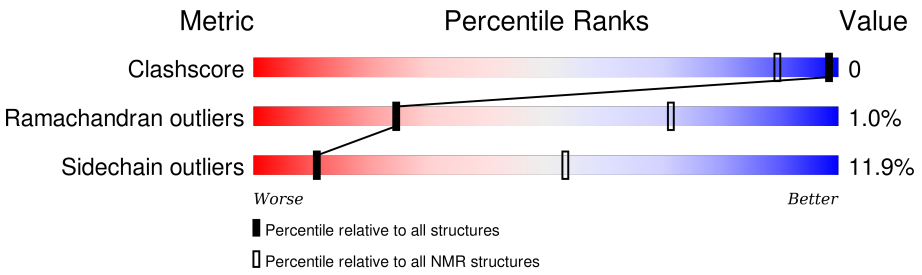
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Mogul	:	unknown
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	rb-20027457
Ideal 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 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	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	176	

2 Ensemble composition and analysis (i)

This entry contains 10 models. Model 4 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:5-A:16, A:35-A:45, A:49-A:55, A:76-A:87, A:91-A:103, A:120-A:143, A:160-A:170 (90)	0.62	4

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	2, 3, 5, 7, 9, 10
2	1, 4, 8
Single-model clusters	6

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 2610 atoms, of which 1263 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called OUTER MEMBRANE PROTEIN A.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	176	2610	851	1263	229	263	4	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	15	PHE	TRP	engineered	UNP P0A910
A	57	PHE	TRP	engineered	UNP P0A910
A	102	PHE	TRP	engineered	UNP P0A910
A	143	PHE	TRP	engineered	UNP P0A910

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: OUTER MEMBRANE PROTEIN A

Chain A: 48% • 49%

A1 P2 K3 D4 N5 Q17 Y18 H19 D20 T21 G22 F23 I24 N25 N26 N27 G28 P29 T30 H31 E32 N33 Q34 N46 P47 Y48 D56 F57 L58 G59 R60 M61 P62 Y63 K64 G65 S66 V67 E68 N69 G70 A71 Y72 K73 A74 Q75 T88 D89 D90 R96 R103 A104 D105 T106

K107 S108 N109 V110 Y111 G112 K113 N114 H115 D116 T117 G118 V119 Y129 R138 T144 N145 N146 I147 G148 D149 A150 H151 T152 I153 G154 T155 R156 P157 D158 N159 R169 F170 G171 Q172 G173 E174 A175 A176

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

The representative model is number 4. Colouring as in section 4.1 above.

- Molecule 1: OUTER MEMBRANE PROTEIN A

Chain A: 45% 6% • 49%

A1 P2 K3 D4 N5 Q17 Y18 H19 D20 T21 G22 F23 I24 N25 N26 N27 G28 P29 T30 H31 E32 N33 Q34 N46 P47 Y48 D56 F57 L58 G59 R60 M61 P62 Y63 K64 G65 S66 V67 E68 N69 G70 A71 Y72 K73 A74 Q75 L79 K82 Y85 T88 D89 D90 L91

D92 R96 F102 R103 A104 D105 T106 K107 S108 N109 V110 Y111 G112 K113 N114 H115 D116 T117 G118 V119 V122 Y129 R138 L139 E140 T144 N145 N146 I147 G148 D149 A150 H151 T152 I153 G154 T155 R156 P157 D158 N159 G171 Q172 G173 E174 A175 A176

5 Refinement protocol and experimental data overview (i)

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

Of the 10 calculated structures, 10 were deposited, based on the following criterion: *all calculated structures submitted*.

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

Software name	Classification	Version
DYANA	structure solution	1.5
OPAL	refinement	2.6

No chemical shift data was provided. 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)

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	0.75±0.01	0±0/711 (0.0±0.0%)	1.03±0.04	1±1/963 (0.1±0.1%)
All	All	0.75	0/7110 (0.0%)	1.03	13/9630 (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	0.0±0.0	1.9±1.2
All	All	0	19

There are no bond-length outliers.

5 of 7 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(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	A	96	ARG	NE-CZ-NH2	-10.84	114.88	120.30	9	3
1	A	138	ARG	NE-CZ-NH1	7.44	124.02	120.30	9	1
1	A	138	ARG	NE-CZ-NH2	-7.00	116.80	120.30	2	3
1	A	96	ARG	NE-CZ-NH1	6.34	123.47	120.30	2	2
1	A	169	ARG	NE-CZ-NH2	-5.95	117.32	120.30	9	2

There are no chirality outliers.

5 of 7 unique planar outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Group	Models (Total)
1	A	96	ARG	Sidechain	5
1	A	129	TYR	Sidechain	5
1	A	168	TYR	Sidechain	3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group	Models (Total)
1	A	85	TYR	Sidechain	2
1	A	103	ARG	Sidechain	2

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	691	680	680	0±1
All	All	6910	6800	6800	5

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 0.

All unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:40:PHE:CD2	1:A:54:GLY:HA3	0.47	2.45	1	1
1:A:12:LYS:HE2	1:A:40:PHE:CE1	0.44	2.47	1	1
1:A:82:LYS:HE2	1:A:94:TYR:CE1	0.43	2.48	7	2
1:A:8:TYR:CZ	1:A:42:GLY:HA3	0.41	2.50	6	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	90/176 (51%)	85±1 (95±1%)	4±2 (4±2%)	1±2 (1±2%)	24	71
All	All	900/1760 (51%)	853 (95%)	38 (4%)	9 (1%)	24	71

5 of 7 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	87	ILE	2
1	A	5	ASN	2
1	A	76	GLY	1
1	A	45	VAL	1
1	A	122	VAL	1

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	67/135 (50%)	59±3 (88±4%)	8±3 (12±4%)	11	53
All	All	670/1350 (50%)	590 (88%)	80 (12%)	11	53

5 of 38 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	5	ASN	9
1	A	129	TYR	4
1	A	80	THR	4
1	A	12	LYS	3
1	A	79	LEU	3

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)

No chemical shift data were provided