wwPDB NMR Structure Validation Summary Report (i)

Apr 27, 2016 – 02:30 AM BST

PDB ID	:	2MAW
Title	:	NMR structures of the alpha7 nAChR transmembrane domain.
Authors	:	Bondarenko, V.; Mowrey, D.; Xu, Y.; Tang, P.
Deposited on	:	2013-07-19

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 (i) symbol.

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

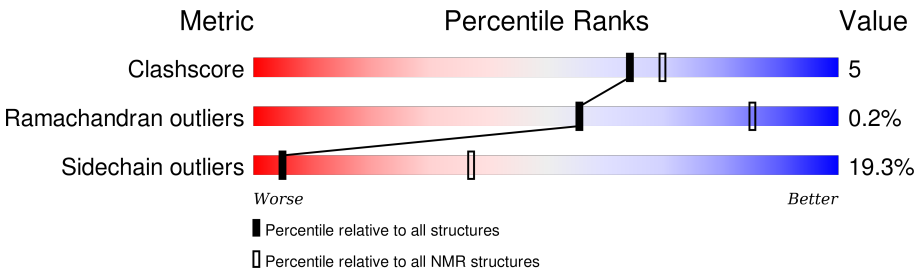
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Mogul	:	unknown
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	rb-20027457
Ideal 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 is 72%.

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	137	66% 15% 17% •

2 Ensemble composition and analysis (i)

This entry contains 20 models. Model 1 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:5-A:60, A:66-A:94, A:106-A:131 (111)	0.64	1

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. No single-model clusters were found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 2031 atoms, of which 1030 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Neuronal acetylcholine receptor subunit alpha-7.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	134	2031	655	1030	148	190	8	0

There are 17 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	SER	-	EXPRESSION TAG	UNP P36544
A	2	ASN	-	EXPRESSION TAG	UNP P36544
A	3	ALA	-	EXPRESSION TAG	UNP P36544
A	4	GLU	ARG	engineered mutation	UNP P36544
A	5	GLU	ARG	engineered mutation	UNP P36544
A	6	GLU	THR	engineered mutation	UNP P36544
A	61	SER	ALA	engineered mutation	UNP P36544
A	66	SER	VAL	engineered mutation	UNP P36544
A	68	SER	LEU	engineered mutation	UNP P36544
A	102	GLU	LYS	engineered mutation	UNP P36544
A	103	GLY	-	linker	UNP P36544
A	104	GLY	-	linker	UNP P36544
A	105	GLY	-	linker	UNP P36544
A	106	GLU	-	linker	UNP P36544
A	107	GLY	-	linker	UNP P36544
A	108	ILE	VAL	engineered mutation	UNP P36544
A	137	GLU	ALA	engineered mutation	UNP P36544

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: Neuronal acetylcholine receptor subunit alpha-7

Chain A: 66% 15% 17% •

SER ASN ALA E4 L14 I15 P16 C17 V18 L46 S47 L48 T49 V50 L54 S61 T62 S63 D64 S65 T76 V80 G81 L82 S83 V84 V85 V86 Y93 H94 H95 H96 D97 P98 D99 G100 G101 E102 G103 G104 G105 L111 C112 L113 M114 S117 V118 F119 P132 N133

F134 V135 E136 E137

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

The representative model is number 1. Colouring as in section 4.1 above.

- Molecule 1: Neuronal acetylcholine receptor subunit alpha-7

Chain A: 62% 18% • 17% •

SER ASN ALA E4 L7 L14 I15 P16 C17 V18 L19 L23 L26 V27 E36 K37 I38 S39 L40 L46 S47 L48 T49 V50 F51 M52 L53 L54 S61 T62 S63 D64 S65 T76 V80 V84 V85 V86 T87 V88 I89 V90 H95 H96 D97 P98 D99 G100 G101

E102 G103 G104 G105 S117 M129 P132 N133 F134 V135 E136 E137

5 Refinement protocol and experimental data overview (i)

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

Of the 25 calculated structures, 20 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
CYANA	structure solution	
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)	2maw_cs.str
Number of chemical shift lists	1
Total number of shifts	1313
Number of shifts mapped to atoms	1313
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	72%

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
1	A	836	905	905	9±2
All	All	16720	18100	18100	182

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 5.

5 of 52 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:80:VAL:O	1:A:84:VAL:HG23	0.62	1.94	5	5
1:A:9:TYR:O	1:A:13:LEU:HD12	0.61	1.95	7	2
1:A:84:VAL:O	1:A:88:VAL:HG23	0.61	1.95	13	6
1:A:14:LEU:O	1:A:18:VAL:HG12	0.60	1.97	3	19
1:A:76:THR:O	1:A:80:VAL:HG23	0.59	1.96	16	20

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	111/137 (81%)	106±2 (95±1%)	5±2 (5±1%)	0±0 (0±0%)	56 85
All	All	2220/2740 (81%)	2114 (95%)	102 (5%)	4 (0%)	56 85

All 2 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	94	HIS	2
1	A	35	GLY	2

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	97/117 (83%)	78±3 (81±3%)	19±3 (19±3%)	5 37
All	All	1940/2340 (83%)	1565 (81%)	375 (19%)	5 37

5 of 61 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	54	LEU	14
1	A	114	MET	13
1	A	119	PHE	13
1	A	48	LEU	13
1	A	117	SER	13

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 72% for the well-defined parts and 72% for the entire structure.

7.1 Chemical shift list 1

File name: 2maw_cs.str

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	1313
Number of shifts mapped to atoms	1313
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
$^{13}C_{\alpha}$	136	-0.85 ± 0.13	Should be applied
$^{13}C_{\beta}$	107	0.22 ± 0.02	None needed (< 0.5 ppm)
$^{13}C'$	103	-0.06 ± 0.09	None needed (< 0.5 ppm)
^{15}N	123	0.65 ± 0.29	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 72%, i.e. 942 atoms were assigned a chemical shift out of a possible 1304. 12 out of 33 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	501/547 (92%)	207/218 (95%)	192/222 (86%)	102/107 (95%)
Sidechain	409/673 (61%)	222/385 (58%)	184/281 (65%)	3/7 (43%)

Continued on next page...

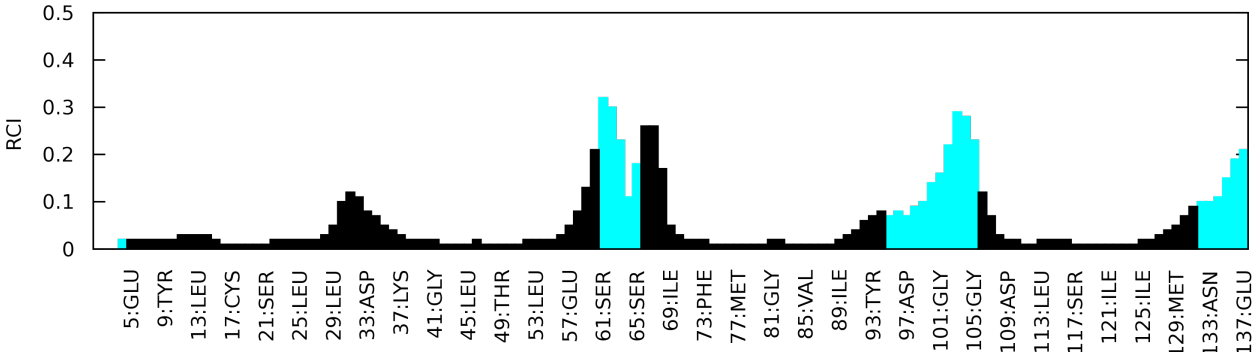
Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Aromatic	32/84 (38%)	17/45 (38%)	15/38 (39%)	0/1 (0%)
Overall	942/1304 (72%)	446/648 (69%)	391/541 (72%)	105/115 (91%)

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: