

Full wwPDB NMR Structure Validation Report (i)

Apr 21, 2024 – 06:01 AM EDT

PDB ID : 2L0J

Title : Solid State NMR structure of the M2 proton channel from Influenza A Virus

in hydrated lipid bilayer

Authors: Sharma, M.; Yi, M.; Dong, H.; Qin, H.; Peterson, E.; Busath, D.D.; Zhou,

H.X.; Cross, T.A.

Deposited on : 2010-07-08

This is a Full wwPDB NMR Structure Validation 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)

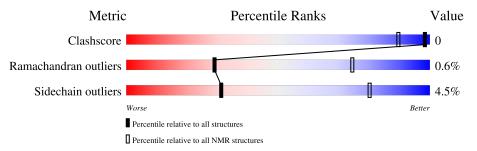
Validation Pipeline (wwPDB-VP) : 2.36.2

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure: $SOLID\text{-}STATE\ NMR$

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	NMR archive		
Metric	$(\# \mathrm{Entries})$	$(\# \mathrm{Entries})$		
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	44	84%	7%	•	7%
1	В	44	84%	9%		7%
1	С	44	84%	9%		7%
1	D	44	89%	Ī	5%	7%



2 Ensemble composition and analysis (i)

This entry contains 8 models. Model 6 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: fewest violations.

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:23-A:62, B:22-B:62, C:22-	0.81	6					
	C:62, D:22-D:62 (163)							

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	3, 4, 5, 6, 7, 8		
2	1, 2		



3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 2766 atoms, of which 1418 are hydrogens and 0 are deuteriums.

• Molecule 1 is a protein called Matrix protein 2.

Mol	Chain	Residues		Atoms				
1	A	41	Total	С	Н	N	О	0
1	Α	41	692	228	355	57	52	U
1	В	41	Total	С	Н	N	О	0
1	Б	41	691	228	354	57	52	U
1	С	41	Total	С	Н	N	О	0
1		41	692	228	355	57	52	U
1	D	41	Total	С	Н	N	О	0
1	D	41	691	228	354	57	52	U

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	19	SER	-	expression tag	UNP P63231
A	20	ASN	-	expression tag	UNP P63231
A	21	ALA	-	expression tag	UNP P63231
A	50	SER	CYS	engineered mutation	UNP P63231
В	19	SER	-	expression tag	UNP P63231
В	20	ASN	-	expression tag	UNP P63231
В	21	ALA	-	expression tag	UNP P63231
В	50	SER	CYS	engineered mutation	UNP P63231
С	19	SER	-	expression tag	UNP P63231
С	20	ASN	-	expression tag	UNP P63231
С	21	ALA	-	expression tag	UNP P63231
С	50	SER	CYS	engineered mutation	UNP P63231
D	19	SER	-	expression tag	UNP P63231
D	20	ASN	-	expression tag	UNP P63231
D	21	ALA	-	expression tag	UNP P63231
D	50	SER	CYS	engineered mutation	UNP P63231

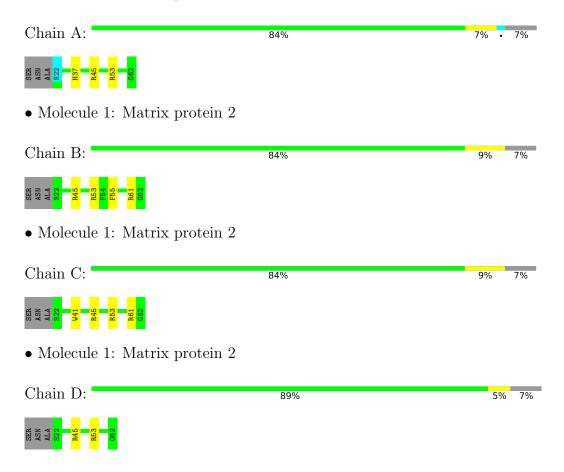


4 Residue-property plots (i)

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: Matrix protein 2



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: Matrix protein 2

Chain A: 84% 7% · 7%

SER ASN ALA R45 R45 R61

• Molecule 1: Matrix protein 2

Chain B: 77% 11% 5% 7%

• Molecule 1: Matrix protein 2

Chain C: 80% 9% 5% 7%

• Molecule 1: Matrix protein 2

Chain D: 89% 5% 7%



4.2.2 Score per residue for model 2

• Molecule 1: Matrix protein 2

Chain A: 80% 9% · · 7%



• Molecule 1: Matrix protein 2

Chain B: 86% 5% · 7%



• Molecule 1: Matrix protein 2

Chain C: 86% • 5% 7%





• Molecule 1: Matrix protein 2

Chain D: 84% 7% • 7%



4.2.3 Score per residue for model 3

 \bullet Molecule 1: Matrix protein 2

Chain A: 82% 7% · · 7%



• Molecule 1: Matrix protein 2

Chain B: 75% 18% 7%



• Molecule 1: Matrix protein 2

Chain C: 89% 5% 7%



• Molecule 1: Matrix protein 2

Chain D: 86% 7% 7%



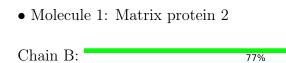
4.2.4 Score per residue for model 4

• Molecule 1: Matrix protein 2

Chain A: 80% 11% • 7%







• Molecule 1: Matrix protein 2

Chain C: 84% 9% 7%

SER ASN
ALA
S22
H37
R53
R61
G62

• Molecule 1: Matrix protein 2

Chain D: 84% 9% 7%



4.2.5 Score per residue for model 5

• Molecule 1: Matrix protein 2

Chain A: 84% 7% · 7%

SER A SN A LA S22 H37 H37 D44 R45

• Molecule 1: Matrix protein 2

Chain B: 82% 9% • 7%



• Molecule 1: Matrix protein 2

Chain C: 89% 5% 7%

• Molecule 1: Matrix protein 2

Chain D: 84% 5% 5% 7%





4.2.6 Score per residue for model 6 (medoid)

• Molecule 1: Matrix protein 2

Chain A: 82% 9% • 7%

SER ASN ALA ALA H37 H37 K45 R53 R53 R65 R65

• Molecule 1: Matrix protein 2

Chain B: 86% 7% 7%

SER ASN ALA SZ2 SZ2 F48 F48 R53 R53 G62

• Molecule 1: Matrix protein 2

Chain C: 86% 5% 7%

SER ASN ALA S22 S22 S22 R53 R61 R61 G62

• Molecule 1: Matrix protein 2

Chain D: 86% 7% 7%

SER ASN ALA S22 S22 N41 N41 R53 R61 G62

4.2.7 Score per residue for model 7

• Molecule 1: Matrix protein 2

Chain A: 82% 7% · · 7%



• Molecule 1: Matrix protein 2

Chain B: 89% ... 7%

SER
ASN
ALA
S22
R45
R45
R61
G62

• Molecule 1: Matrix protein 2

Chain C: 89% 5% 7%





• Molecule 1: Matrix protein 2

Chain D: 82% 11% 7%



4.2.8 Score per residue for model 8

• Molecule 1: Matrix protein 2

Chain A: 84% 5% · · 7%



• Molecule 1: Matrix protein 2

Chain B: 77% 7% 9% 7%



• Molecule 1: Matrix protein 2

Chain C: 73% 14% 7% 7%



• Molecule 1: Matrix protein 2

Chain D: 84% 9% 7%





5 Refinement protocol and experimental data overview (i)



The models were refined using the following method: simulated annealing and molecular dynamics, molecular dynamics.

Of the 8 calculated structures, 8 were deposited, based on the following criterion: structures with the least restraint violations.

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

Software name	Classification	Version
X-PLOR NIH	refinement	2.23
X-PLOR NIH	structure solution	2.23
NAMD	refinement	2.7b

No chemical shift data was provided. Note: This is a solid-state NMR structure, where hydrogen atoms are typically not assigned a chemical shift value, which may lead to lower completeness of assignment measure.



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	В	Sond lengths	Bond angles		
WIOI		RMSZ	#Z>5	RMSZ	#Z>5	
1	A	0.80 ± 0.04	$0\pm0/340~(~0.0\pm~0.0\%)$	1.16 ± 0.08	$2\pm1/458~(~0.5\pm~0.3\%)$	
1	В	0.81 ± 0.03	$0\pm0/346~(~0.0\pm~0.0\%)$	1.21 ± 0.07	$4\pm1/466~(~0.8\pm~0.3\%)$	
1	С	0.82 ± 0.03	$0\pm0/346~(~0.0\pm~0.0\%)$	1.22 ± 0.10	$3\pm2/466~(~0.7\pm~0.5\%)$	
1	D	0.84 ± 0.01	$0\pm0/346~(~0.0\pm~0.0\%)$	1.17 ± 0.05	$3\pm1/466~(~0.6\pm~0.2\%)$	
All	All	0.82	0/11024 ($0.0%$)	1.19	96/14848 (0.6%)	

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	С	0.0 ± 0.0	0.4 ± 0.5
1	D	0.0 ± 0.0	0.2 ± 0.4
1	A	0.0 ± 0.0	0.8 ± 0.7
1	В	0.1 ± 0.3	1.0 ± 1.1
All	All	1	19

There are no bond-length outliers.

All unique angle outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

Mol	Chain	Chain Res Type Atoms Z Observed($Observed(^o)$	$\mathrm{Ideal}(^{o})$	Mod	dels		
IVIOI	Chain	nes	Type	Atoms		Observed()	ideai()	Worst	Total
1	С	45	ARG	NE-CZ-NH1	8.60	124.60	120.30	1	4
1	С	47	PHE	CB-CG-CD1	8.36	126.65	120.80	8	1
1	В	45	ARG	NE-CZ-NH1	8.27	124.44	120.30	4	3
1	В	53	ARG	NE-CZ-NH1	8.22	124.41	120.30	6	5
1	С	47	PHE	CB-CG-CD2	-8.03	115.18	120.80	8	1
1	D	61	ARG	NE-CZ-NH2	-7.72	116.44	120.30	7	3
1	В	61	ARG	NE-CZ-NH1	7.61	124.10	120.30	6	5
1	D	53	ARG	NE-CZ-NH1	7.55	124.08	120.30	6	5
1	В	61	ARG	NE-CZ-NH2	-7.51	116.54	120.30	8	2

Continued on next page...



 $Continued\ from\ previous\ page...$

	nueu jron				77	01 1(0)	T1 1(a)	Mod	$\overline{\mathrm{dels}}$
Mol	Chain	Res	Type	Atoms	\mathbf{Z}	$Observed(^o)$	$\mathbf{Ideal}(^o)$	Worst	Total
1	D	61	ARG	NE-CZ-NH1	7.45	124.02	120.30	6	4
1	D	45	ARG	NE-CZ-NH1	7.38	123.99	120.30	7	4
1	A	61	ARG	NE-CZ-NH1	7.37	123.98	120.30	3	3
1	С	61	ARG	NE-CZ-NH1	7.17	123.89	120.30	5	6
1	В	52	TYR	CB-CG-CD2	-7.12	116.72	121.00	3	1
1	A	53	ARG	NE-CZ-NH1	7.04	123.82	120.30	7	3
1	A	53	ARG	NE-CZ-NH2	-7.02	116.79	120.30	3	2
1	A	45	ARG	NE-CZ-NH1	6.67	123.64	120.30	6	5
1	В	22	SER	N-CA-CB	6.65	120.48	110.50	4	1
1	С	55	PHE	CB-CG-CD1	6.58	125.41	120.80	1	2
1	С	55	PHE	CB-CG-CD2	-6.55	116.22	120.80	2	2
1	В	54	PHE	CB-CG-CD1	6.52	125.36	120.80	4	1
1	В	55	PHE	CB-CG-CD1	6.32	125.23	120.80	1	3
1	В	54	PHE	CB-CG-CD2	-6.29	116.40	120.80	4	1
1	A	52	TYR	CB-CG-CD2	-6.24	117.26	121.00	8	2
1	С	53	ARG	NE-CZ-NH1	6.23	123.41	120.30	8	4
1	A	54	PHE	CB-CG-CD1	6.02	125.01	120.80	4	1
1	В	52	TYR	CB-CG-CD1	5.91	124.55	121.00	3	1
1	D	55	PHE	CB-CG-CD1	5.84	124.89	120.80	2	1
1	D	48	PHE	CB-CG-CD2	-5.60	116.88	120.80	5	1
1	В	55	PHE	CB-CG-CD2	-5.57	116.90	120.80	1	3
1	D	48	PHE	CB-CG-CD1	5.52	124.67	120.80	5	1
1	D	52	TYR	CB-CG-CD1	-5.52	117.69	121.00	5	1
1	В	45	ARG	NE-CZ-NH2	-5.46	117.57	120.30	7	1
1	С	37	HIS	CA-CB-CG	5.46	122.87	113.60	8	1
1	С	45	ARG	NE-CZ-NH2	-5.45	117.57	120.30	8	1
1	D	47	PHE	CB-CG-CD1	5.45	124.61	120.80	4	1
1	D	47	PHE	CB-CG-CD2	-5.34	117.06	120.80	4	1
1	С	61	ARG	NE-CZ-NH2	-5.28	117.66	120.30	1	1
1	A	52	TYR	CB-CG-CD1	5.24	124.14	121.00	8	1
1	С	41	TRP	CB-CG-CD1	-5.23	120.20	127.00	6	1
1	D	45	ARG	NE-CZ-NH2	-5.19	117.71	120.30	7	1
1	С	41	TRP	CB-CG-CD2	5.14	133.28	126.60	6	1
1	В	53	ARG	NE-CZ-NH2	-5.12	117.74	120.30	6	1
1	A	54	PHE	CB-CG-CD2	-5.09	117.24	120.80	4	1
1	D	55	PHE	CB-CG-CD2	-5.05	117.27	120.80	2	1
1	В	22	SER	CB-CA-C	5.01	119.61	110.10	3	1

All unique chiral outliers are listed below.

\mathbf{Mol}	Chain	Res Type Atoms		Atoms	Models (Total)	
1	В	22	SER	CA	1	



All 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	53	ARG	Sidechain	3
1	В	53	ARG	Sidechain	3
1	D	45	ARG	Sidechain	2
1	В	45	ARG	Sidechain	2
1	С	45	ARG	Sidechain	2
1	A	45	ARG	Sidechain	2
1	В	61	ARG	Sidechain	2
1	С	53	ARG	Sidechain	1
1	A	52	TYR	Sidechain	1
1	В	52	TYR	Sidechain	1

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	331	351	350	0±0
1	В	337	354	354	0±0
1	С	337	355	354	0±1
1	D	337	354	354	0±0
All	All	10736	11312	11296	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	
Atom-1	Atom-2	Clash(A)	Distance(A)	Worst	Total
1:A:38:LEU:HD22	1:D:33:ILE:HG23	0.55	1.78	8	1
1:C:41:TRP:CD1	1:C:42:ILE:HG13	0.50	2.41	8	1
1:A:41:TRP:CZ2	1:A:45:ARG:NH1	0.46	2.84	2	1
1:C:41:TRP:CD1	1:C:42:ILE:N	0.46	2.84	8	1
1:B:33:ILE:HG21	1:C:38:LEU:CD2	0.45	2.41	8	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	39/44 (89%)	37±1 (96±2%)	2±1 (4±2%)	0±0 (0±0%)	100	100
1	В	39/44 (89%)	37±0 (95±1%)	2±1 (4±2%)	0±0 (1±1%)	20	68
1	С	39/44 (89%)	36±1 (93±2%)	3±1 (7±3%)	0±0 (0±1%)	44	80
1	D	39/44 (89%)	37±1 (95±2%)	1±1 (4±3%)	0±0 (1±1%)	16	63
All	All	1248/1408 (89%)	1180 (95%)	60 (5%)	8 (1%)	29	74

All 4 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	D	48	PHE	3
1	В	48	PHE	3
1	С	48	PHE	1
1	D	32	ILE	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	Perce	ntiles
1	A	35/38~(92%)	34±0 (96±1%)	1±0 (4±1%)	36	84
1	В	$36/38 \; (95\%)$	34±2 (93±5%)	2±2 (7±5%)	20	69
1	С	$36/38 \; (95\%)$	34±1 (95±2%)	2±1 (5±2%)	31	80
1	D	$36/38 \; (95\%)$	35±1 (97±2%)	1±1 (3±2%)	46	90
All	All	1144/1216 (94%)	1093 (96%)	51 (4%)	31	80

All 33 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	37	HIS	5
1	В	55	PHE	4
1	С	41	TRP	3
1	D	52	TYR	3
1	В	36	LEU	2
1	В	45	ARG	2
1	В	52	TYR	2
1	С	26	LEU	2
1	С	37	HIS	2
1	С	55	PHE	2
1	В	26	LEU	2
1	В	33	ILE	1
1	A	38	LEU	1
1	A	51	ILE	1
1	D	24	ASP	1
1	D	55	PHE	1
1	A	48	PHE	1
1	В	49	LYS	1
1	В	51	ILE	1
1	В	54	PHE	1
1	В	61	ARG	1
1	С	22	SER	1
1	A	44	ASP	1
1	С	51	ILE	1
1	D	43	LEU	1
1	D	41	TRP	1
1	A	47	PHE	1
1	С	33	ILE	1
1	D	26	LEU	1
1	A	52	TYR	1
1	В	40	LEU	1
1	В	47	PHE	1
1	С	47	PHE	1

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

