wwPDB NMR Structure Validation Summary Report

May 4, 2024 – 10:03 AM EDT

PDB ID	:	6W0O
EMDB ID	:	EMD-21501
BMRB ID	:	30731
Title	:	Amyloid-beta(1-40) fibril derived from Alzheimer's disease cortical tissue
Authors	:	Ghosh, U.; Thurber, K.R.; Tycko, R.
Deposited on	:	2020-03-02

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:

EMDB validation analysis	:	NOT EXECUTED
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	NOT EXECUTED
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

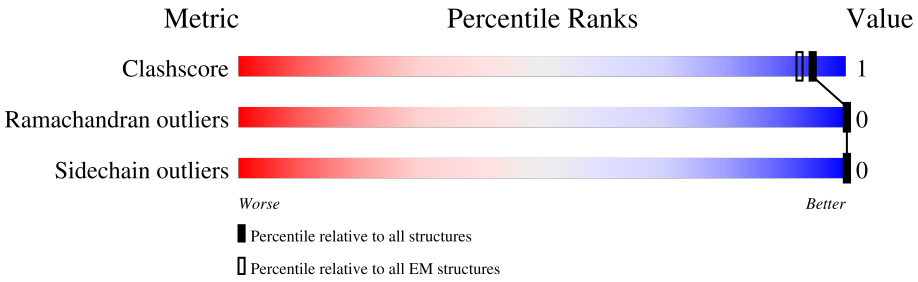
1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY, SOLID-STATE NMR

The reported resolution of this entry is 2.77 Å.

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	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826

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	1	40	60% 18% 22%
1	2	40	60% 18% 22%
1	3	40	60% 18% 22%
1	4	40	65% 12% 22%
1	5	40	65% 12% 22%
1	6	40	68% 10% 22%

2 Ensemble composition and analysis

This entry contains 10 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: *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	1:16-1:39, 2:16-2:39, 3:16-3:39, 4:14-4:39, 5:14-5:39, 6:14-6:40 (151)	0.18	6

Ill-defined regions of proteins are excluded from the global statistics.

Ligands and non-protein polymers are included in the analysis.

NmrClust was unable to cluster the ensemble.

Error message: NMCparsrange - Unexpected character.

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 2808 atoms, of which 1404 are hydrogens and 0 are deuteriums.

In the tables below, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Amyloid-beta precursor protein.

Mol	Chain	Residues	Atoms						AltConf	Trace
1	1	31	Total	C	H	N	O	S	0	
			468	152	234	39	42	1		
1	2	31	Total	C	H	N	O	S	0	
			468	152	234	39	42	1		
1	3	31	Total	C	H	N	O	S	0	
			468	152	234	39	42	1		
1	4	31	Total	C	H	N	O	S	0	
			468	152	234	39	42	1		
1	5	31	Total	C	H	N	O	S	0	
			468	152	234	39	42	1		
1	6	31	Total	C	H	N	O	S	0	
			468	152	234	39	42	1		

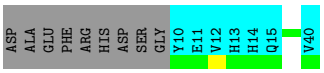
4 Residue-property plots

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.

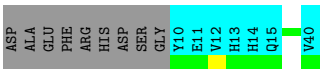
• Molecule 1: Amyloid-beta precursor protein

Chain 1:



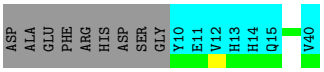
• Molecule 1: Amyloid-beta precursor protein

Chain 2:



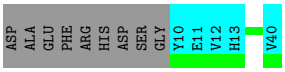
• Molecule 1: Amyloid-beta precursor protein

Chain 3:



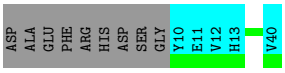
• Molecule 1: Amyloid-beta precursor protein

Chain 4:



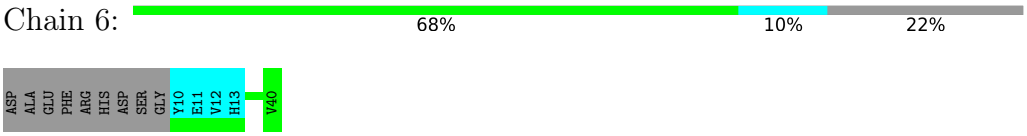
• Molecule 1: Amyloid-beta precursor protein

Chain 5:



• Molecule 1: Amyloid-beta precursor protein

Chain 6:



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

The representative model is number 6. Colouring as in section 4.1 above.

• Molecule 1: Amyloid-beta precursor protein

Chain 1:

• Molecule 1: Amyloid-beta precursor protein

Chain 2:

• Molecule 1: Amyloid-beta precursor protein

Chain 3:

• Molecule 1: Amyloid-beta precursor protein

Chain 4:

• Molecule 1: Amyloid-beta precursor protein

Chain 5:

• Molecule 1: Amyloid-beta precursor protein

Chain 6:


ASP
ALA
GLU
PHE
ARG
HIS
ASP
SER
GLY
Y10
E11
V12
H13
V24
V40

5 Refinement protocol and experimental data overview

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

Of the 60 calculated structures, 10 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	2.53

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	98
Number of shifts mapped to atoms	98
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%

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

6.1 Standard geometry

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

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	1	169	179	179	0±0
1	2	169	179	179	0±1
1	3	169	179	179	0±0
1	4	188	194	194	1±1
1	5	188	194	194	1±1
1	6	196	203	203	1±1
All	All	10790	11280	11280	24

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 1.

5 of 21 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:5:36:VAL:HG12	1:6:36:VAL:HG22	0.56	1.78	2	1
1:4:36:VAL:HG12	1:5:36:VAL:HG22	0.55	1.77	2	1
1:4:24:VAL:HG12	1:5:24:VAL:HG22	0.49	1.84	6	1
1:5:24:VAL:HG12	1:6:24:VAL:HG22	0.49	1.84	6	1
1:1:34:LEU:CD2	1:2:34:LEU:HD12	0.46	2.41	7	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	1	24/40 (60%)	23±0 (94±2%)	1±0 (6±2%)	0±0 (0±0%)	100	100
1	2	24/40 (60%)	23±0 (94±2%)	1±0 (6±2%)	0±0 (0±0%)	100	100
1	3	24/40 (60%)	23±0 (94±2%)	1±0 (6±2%)	0±0 (0±0%)	100	100
1	4	26/40 (65%)	25±0 (96±0%)	1±0 (4±0%)	0±0 (0±0%)	100	100
1	5	26/40 (65%)	25±0 (96±0%)	1±0 (4±0%)	0±0 (0±0%)	100	100
1	6	26/40 (65%)	25±0 (96±0%)	1±0 (4±0%)	0±0 (0±0%)	100	100
All	All	1500/2400 (62%)	1428 (95%)	72 (5%)	0 (0%)	100	100

There are no Ramachandran outliers.

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	1	17/31 (55%)	17±0 (100±0%)	0±0 (0±0%)	100	100
1	2	17/31 (55%)	17±0 (100±0%)	0±0 (0±0%)	100	100
1	3	17/31 (55%)	17±0 (100±0%)	0±0 (0±0%)	100	100
1	4	19/31 (61%)	19±0 (100±0%)	0±0 (0±0%)	100	100
1	5	19/31 (61%)	19±0 (100±0%)	0±0 (0±0%)	100	100
1	6	20/31 (65%)	20±0 (100±0%)	0±0 (0±0%)	100	100
All	All	1090/1860 (59%)	1090 (100%)	0 (0%)	100	100

There are no protein residues with a non-rotameric sidechain to report.

6.3.3 RNA

There are no RNA molecules in this entry.

6.4 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

6.5 Carbohydrates

There are no monosaccharides in this entry.

6.6 Ligand geometry

There are no ligands in this entry.

6.7 Other polymers

There are no such molecules in this entry.

6.8 Polymer linkage issues

There are no chain breaks in this entry.

7 Chemical shift validation

The completeness of assignment taking into account all chemical shift lists is 5% for the well-defined parts and 4% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *brain_shifts2_STAR.txt*

7.1.1 Bookkeeping

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	98
Number of shifts mapped to atoms	98
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

No chemical shift referencing corrections were calculated (not enough data).

7.1.3 Completeness of resonance assignments

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. 98 atoms were assigned a chemical shift out of a possible 1953. 0 out of 37 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	62/785 (8%)	0/332 (0%)	42/302 (14%)	20/151 (13%)
Sidechain	36/1027 (4%)	0/685 (0%)	36/321 (11%)	0/21 (0%)
Aromatic	0/141 (0%)	0/72 (0%)	0/66 (0%)	0/3 (0%)
Overall	98/1953 (5%)	0/1089 (0%)	78/689 (11%)	20/175 (11%)

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.

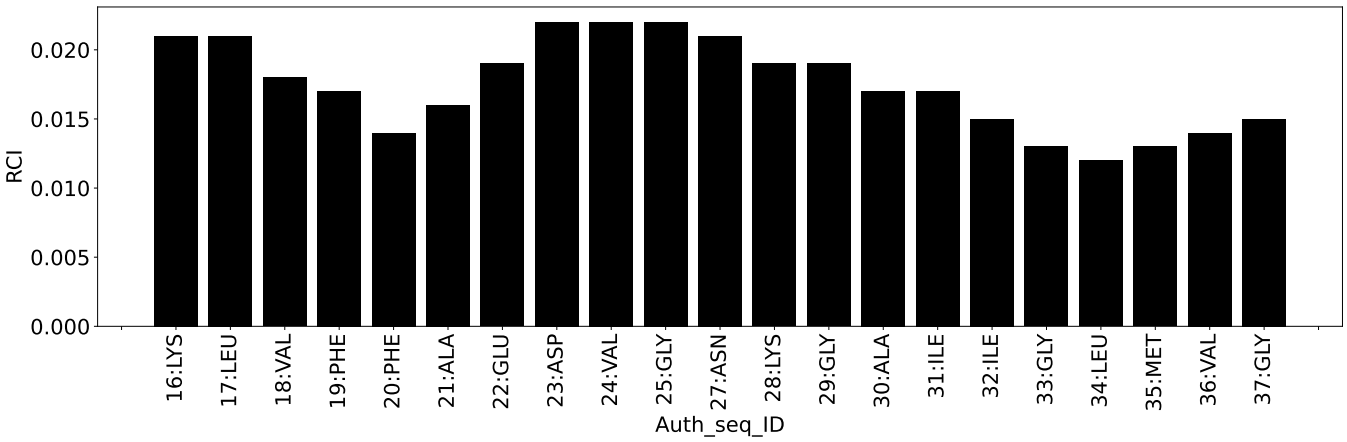
7.1.4 Statistically unusual chemical shifts (i)

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 1:



8 NMR restraints analysis

8.1 Conformationally restricting restraints

The following table provides the summary of experimentally observed NMR restraints in different categories. Restraints are classified into different categories based on the sequence separation of the atoms involved.

Description	Value
Total distance restraints	0
Intra-residue (|i-j|=0)	0
Sequential (|i-j|=1)	0
Medium range (|i-j|>1 and |i-j|<5)	0
Long range (|i-j|≥5)	0
Inter-chain	0
Hydrogen bond restraints	0
Disulfide bond restraints	0
Total dihedral-angle restraints	26
Number of unmapped restraints	0
Number of restraints per residue	0
Number of long range restraints per residue[1]	0

[1]Long range hydrogen bonds and disulfide bonds are counted as long range restraints while calculating the number of long range restraints per residue

8.2 Residual restraint violations

This section provides the overview of the restraint violations analysis. The violations are binned as small, medium and large violations based on its absolute value. Average number of violations per model is calculated by dividing the total number of violations in each bin by the size of the ensemble.

8.2.1 Average number of distance violations per model

Distance violations less than 0.1 Å
are not included in the calculation.There are no distance restraints

8.2.2 Average number of dihedral-angle violations per model

Dihedral-angle violations less than 1° are not included in the calculation.There are no dihedral-angle violations

9 Distance violation analysis (i)

No distance restraints data found

10 Dihedral-angle violation analysis

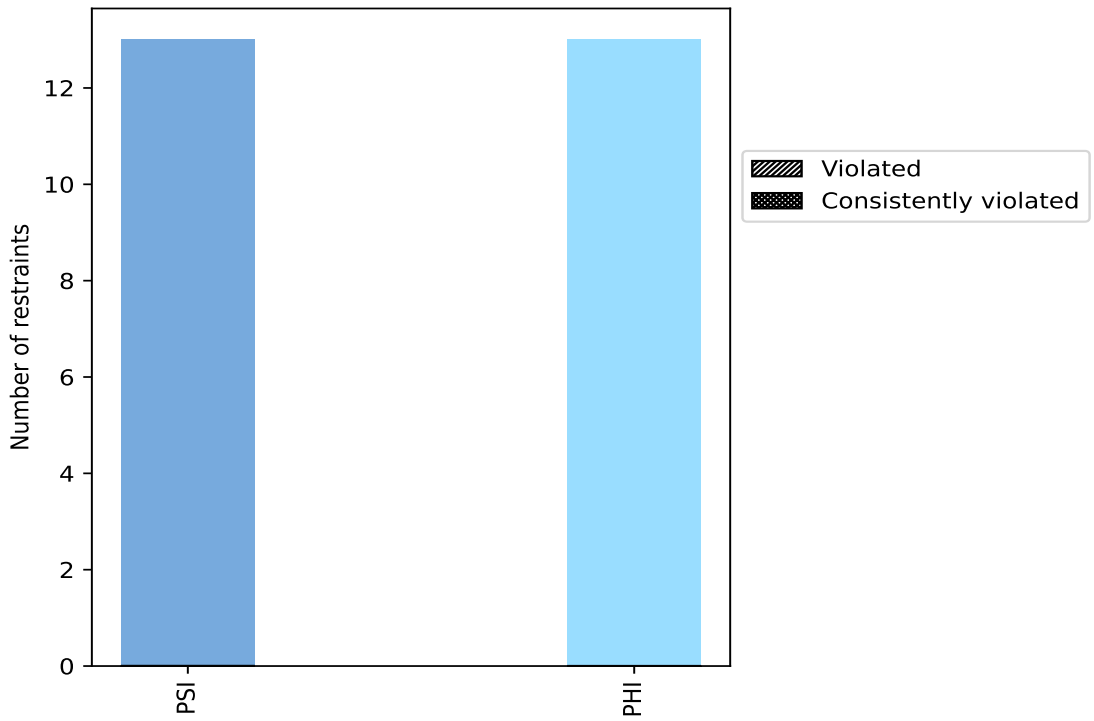
10.1 Summary of dihedral-angle violations

The following table provides the summary of dihedral-angle violations in different dihedral-angle types. Violations less than 1° are not included in the calculation.

Angle type	Count	%[1]	Violated[3]			Consistently Violated[4]		
			Count	%[2]	%[1]	Count	%[2]	%[1]
PSI	13	50.0	0	0.0	0.0	0	0.0	0.0
PHI	13	50.0	0	0.0	0.0	0	0.0	0.0
Total	26	100.0	0	0.0	0.0	0	0.0	0.0

[1] percentage calculated with respect to total number of dihedral-angle restraints, [2] percentage calculated with respect to number of restraints in a particular dihedral-angle type, [3] violated in at least one model, [4] violated in all the models

10.1.1 Bar chart : Distribution of dihedral-angles and violations



Violated and consistently violated restraints are shown using different hatch patterns in their respective categories

10.2 Dihedral-angle violation statistics for each model (i)

No violations found

10.3 Dihedral-angle violation statistics for the ensemble (i)

No violations found

10.4 Most violated dihedral-angle restraints in the ensemble (i)

No violations found

10.5 All violated dihedral-angle restraints (i)

No violations found