



Full wwPDB NMR Structure Validation Report ⓘ

Nov 30, 2021 – 02:10 pm GMT

PDB ID : 7B3J
Title : Dynamic complex between all-D-enantiomeric peptide D3 with wild-type amyloid precursor protein 672-726 fragment (amyloid beta 1-55)
Authors : Bocharov, E.V.; Volynsky, P.E.; Okhrimenko, I.S.; Urban, A.S.
Deposited on : 2020-12-01

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 ⓘ symbol.

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

MolProbity : 4.02b-467
Mogul : 1.8.4 (270009), CSD as541be (2020)
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 : 2.23.2
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.23.2

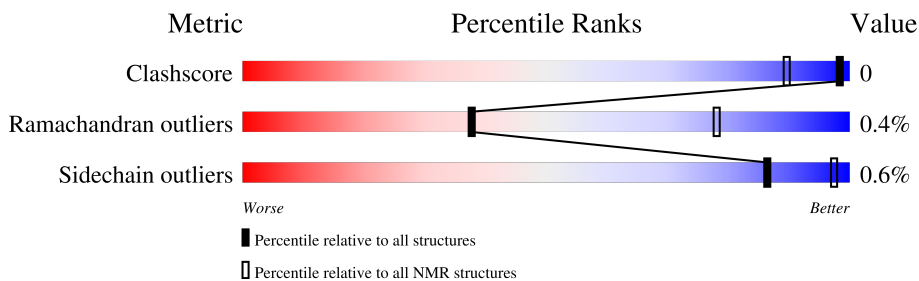
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

SOLUTION NMR

The overall completeness of chemical shifts assignment is 96%.

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 $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	A	55	

2 Ensemble composition and analysis

This entry contains 10 models. Model 10 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *target function*.

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:30-A:52 (23)	0.37	10

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

3 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 1094 atoms, of which 560 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Isoform L-APP677 of Amyloid-beta precursor protein.

Mol	Chain	Residues	Atoms					Trace	
			Total	C	H	N	O		S
1	A	55	862	273	441	71	75	2	0

- Molecule 2 is a protein (with D amino acids) called D3 all D-enantiomeric peptide.

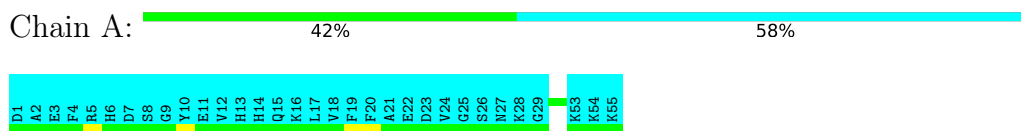
Mol	Chain	Residues	Atoms					Trace
			Total	C	H	N	O	
2	B	12	232	65	119	32	16	0

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: Isoform L-APP677 of Amyloid-beta precursor protein

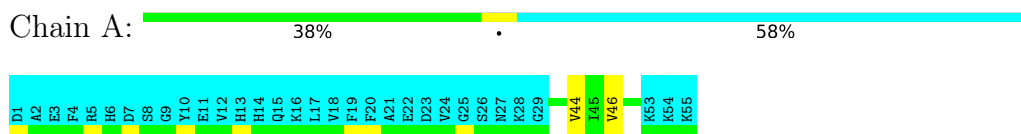


4.2 Scores per residue for each member of the ensemble

Colouring as in section [4.1](#) above.

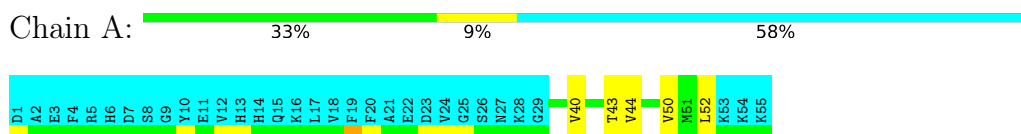
4.2.1 Score per residue for model 1

- Molecule 1: Isoform L-APP677 of Amyloid-beta precursor protein



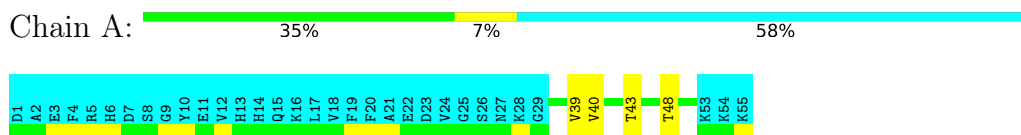
4.2.2 Score per residue for model 2

- Molecule 1: Isoform L-APP677 of Amyloid-beta precursor protein



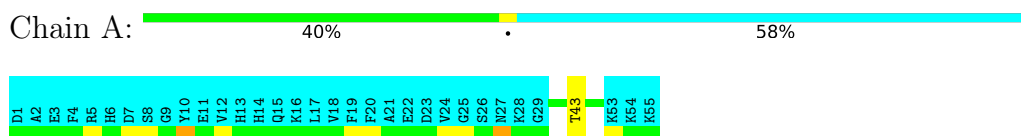
4.2.3 Score per residue for model 3

- Molecule 1: Isoform L-APP677 of Amyloid-beta precursor protein



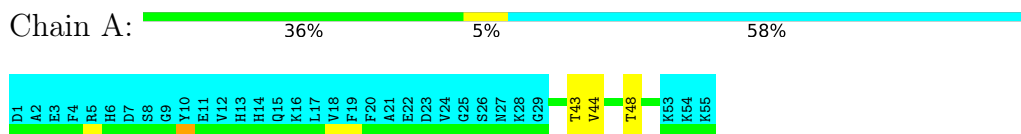
4.2.4 Score per residue for model 4

- Molecule 1: Isoform L-APP677 of Amyloid-beta precursor protein



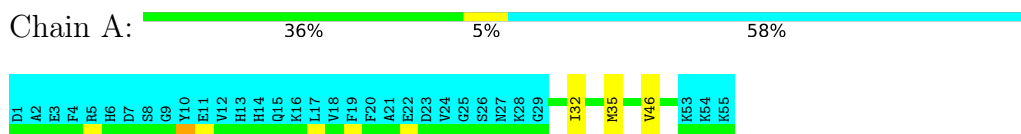
4.2.5 Score per residue for model 5

- Molecule 1: Isoform L-APP677 of Amyloid-beta precursor protein



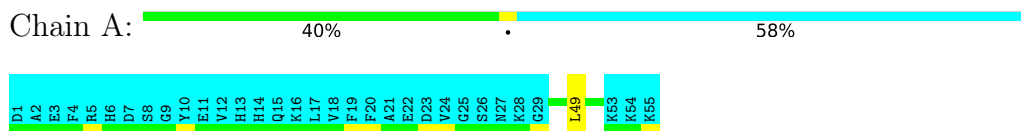
4.2.6 Score per residue for model 6

- Molecule 1: Isoform L-APP677 of Amyloid-beta precursor protein



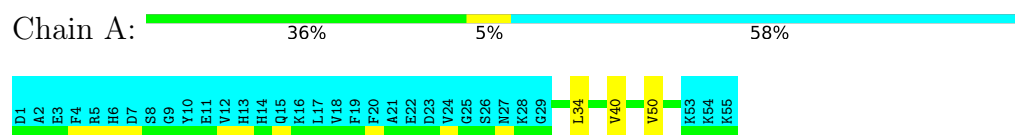
4.2.7 Score per residue for model 7

- Molecule 1: Isoform L-APP677 of Amyloid-beta precursor protein



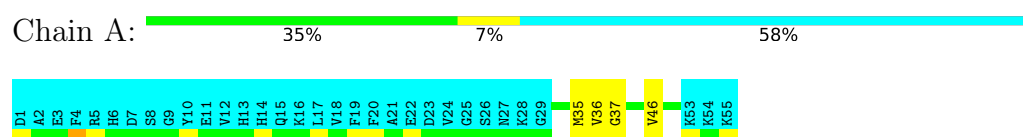
4.2.8 Score per residue for model 8

- Molecule 1: Isoform L-APP677 of Amyloid-beta precursor protein



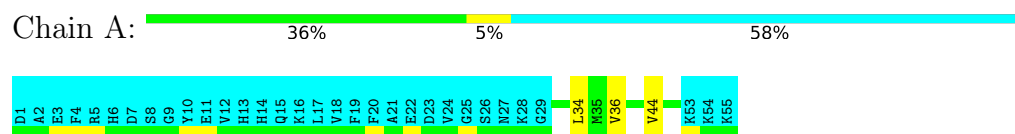
4.2.9 Score per residue for model 9

- Molecule 1: Isoform L-APP677 of Amyloid-beta precursor protein



4.2.10 Score per residue for model 10 (medoid)

- Molecule 1: Isoform L-APP677 of Amyloid-beta precursor protein



5 Refinement protocol and experimental data overview

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

Of the 10 calculated structures, 10 were deposited, based on the following criterion: *structures with acceptable covalent geometry*.

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

Software name	Classification	Version
CYANA	structure calculation	
GROMACS	structure calculation	
GROMACS	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)	working_cs.cif
Number of chemical shift lists	1
Total number of shifts	651
Number of shifts mapped to atoms	651
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	96%

6 Model quality i

6.1 Standard geometry i

Bond lengths and bond angles in the following residue types are not validated in this section: DPR, DAR, 2TL, DHI, DSG, DLE

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	1.41±0.06	0±0/158 (0.1± 0.2%)	1.75±0.09	2±1/217 (1.1± 0.6%)
2	B	0.00±0.00	-	0.00±0.00	-
All	All	1.41	1/1580 (0.1%)	1.75	24/2170 (1.1%)

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	0.2±0.4
All	All	0	2

All unique bond outliers are listed below.

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)	Models	
								Worst	Total
1	A	37	GLY	CA-C	5.53	1.60	1.51	9	1

All 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(°)	Ideal(°)	Models	
								Worst	Total
1	A	44	VAL	CA-CB-CG2	7.00	121.41	110.90	5	1
1	A	36	VAL	CG1-CB-CG2	-6.99	99.72	110.90	9	1
1	A	36	VAL	CA-CB-CG2	6.90	121.25	110.90	9	2
1	A	39	VAL	CG1-CB-CG2	-6.66	100.25	110.90	3	1
1	A	46	VAL	CA-CB-CG1	6.57	120.75	110.90	1	1
1	A	44	VAL	CG1-CB-CG2	-6.12	101.10	110.90	1	3
1	A	35	MET	CG-SD-CE	5.70	109.31	100.20	6	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)	Models	
								Worst	Total
1	A	50	VAL	CG1-CB-CG2	-5.59	101.96	110.90	8	1
1	A	52	LEU	CB-CG-CD2	-5.50	101.64	111.00	2	1
1	A	44	VAL	CA-CB-CG1	5.47	119.11	110.90	10	1
1	A	34	LEU	CB-CG-CD1	5.29	119.99	111.00	10	1
1	A	50	VAL	CA-CB-CG2	5.14	118.61	110.90	2	1
1	A	48	THR	OG1-CB-CG2	-5.13	98.20	110.00	3	2
1	A	49	LEU	CB-CG-CD1	-5.08	102.36	111.00	7	1
1	A	43	THR	OG1-CB-CG2	-5.06	98.36	110.00	5	1
1	A	40	VAL	CG1-CB-CG2	-5.05	102.81	110.90	2	1
1	A	40	VAL	CA-CB-CG2	-5.05	103.32	110.90	8	1
1	A	46	VAL	CA-CB-CG2	-5.03	103.35	110.90	9	1
1	A	43	THR	CA-CB-CG2	5.01	119.41	112.40	2	1
1	A	46	VAL	CG1-CB-CG2	-5.01	102.88	110.90	6	1

There are no chirality outliers.

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	43	THR	Mainchain	1
1	A	35	MET	Mainchain	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	158	193	193	0±0
2	B	113	119	116	0±0
All	All	2710	3120	3057	1

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:VAL:O	1:A:43:THR:HB	0.42	2.14	3	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	23/55 (42%)	23±0 (98±2%)	0±0 (1±2%)	0±0 (0±1%)	38 78
2	B	0	-	-	-	-
All	All	230/670 (34%)	226 (98%)	3 (1%)	1 (0%)	38 78

All 1 unique Ramachandran outliers are listed below.

Mol	Chain	Res	Type	Models (Total)
1	A	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	Percentiles
1	A	18/45 (40%)	18±0 (99±2%)	0±0 (1±2%)	86 97
2	B	0	-	-	-
All	All	180/450 (40%)	179 (99%)	1 (1%)	86 97

All 1 unique residues with a non-rotameric sidechain are listed below.

Mol	Chain	Res	Type	Models (Total)
1	A	34	LEU	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](#)

12 non-standard protein/DNA/RNA residues are modelled in this entry.

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

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](#)

The completeness of assignment taking into account all chemical shift lists is 96% for the well-defined parts and 84% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: starch_output

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	651
Number of shifts mapped to atoms	651
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 \pm precision, ppm	Suggested action
$^{13}\text{C}_\alpha$	54	1.59 ± 0.13	Should be applied
$^{13}\text{C}_\beta$	48	2.79 ± 0.05	Should be applied
$^{13}\text{C}'$	52	2.46 ± 0.14	Should be applied
^{15}N	55	1.28 ± 0.16	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 96%, i.e. 242 atoms were assigned a chemical shift out of a possible 251. 9 out of 9 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^1H	^{13}C	^{15}N
Backbone	115/115 (100%)	46/46 (100%)	46/46 (100%)	23/23 (100%)
Sidechain	127/136 (93%)	70/74 (95%)	57/62 (92%)	0/0 (—%)

Continued on next page...

Continued from previous page...

	Total	¹ H	¹³ C	¹⁵ N
Aromatic	0/0 (—%)	0/0 (—%)	0/0 (—%)	0/0 (—%)
Overall	242/251 (96%)	116/120 (97%)	103/108 (95%)	23/23 (100%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 84%, i.e. 556 atoms were assigned a chemical shift out of a possible 658. 13 out of 13 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	¹ H	¹³ C	¹⁵ N
Backbone	269/275 (98%)	108/110 (98%)	106/110 (96%)	55/55 (100%)
Sidechain	287/327 (88%)	172/187 (92%)	115/130 (88%)	0/10 (0%)
Aromatic	0/56 (0%)	0/31 (0%)	0/22 (0%)	0/3 (0%)
Overall	556/658 (84%)	280/328 (85%)	221/262 (84%)	55/68 (81%)

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:

