



# wwPDB NMR Structure Validation Summary Report ⓘ

Nov 30, 2021 – 02:09 pm GMT

PDB ID : 7B3K  
Title : Dynamic complex between all-D-enantiomeric peptide D3 with L723P mutant of amyloid precursor protein (APP) 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 wwPDB NMR Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto: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

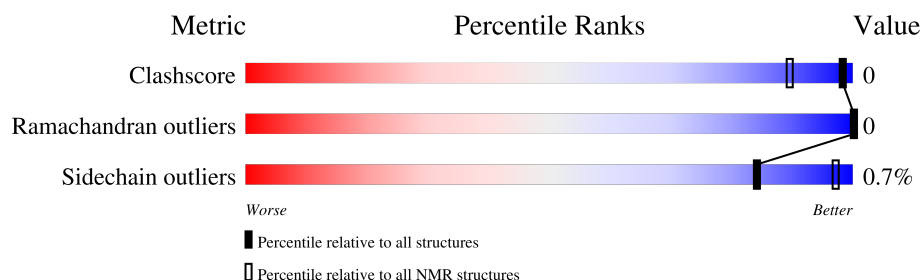
# 1 Overall quality at a glance

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  $\geq 3$ , 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	<div> <div>40%</div> <div>60%</div> </div>

## 2 Ensemble composition and analysis

This entry contains 8 models. Model 8 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:51 (22)	0.49	8

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

### 3 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 1089 atoms, of which 556 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
1	A	55	Total	C	H	N	O	S	0
			857	272	437	71	75	2	

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	52	PRO	LEU	engineered mutation	UNP P05067

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

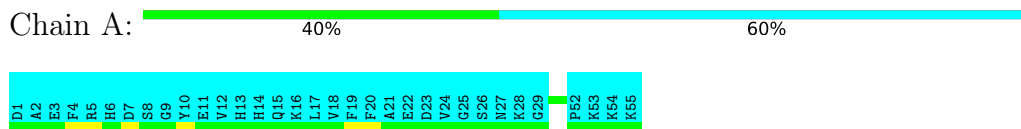
Mol	Chain	Residues	Atoms					Trace
2	B	12	Total	C	H	N	O	0
			232	65	119	32	16	

## 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.

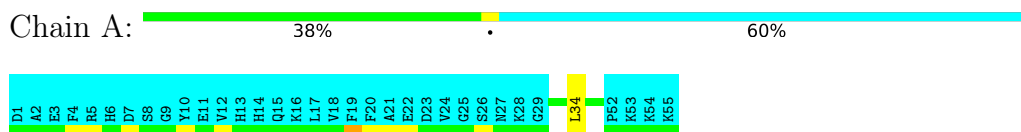
- Molecule 1: Isoform L-APP677 of Amyloid-beta precursor protein



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

The representative model is number 8. Colouring as in section 4.1 above.

- 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 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
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	645
Number of shifts mapped to atoms	645
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: DAR, DLE, DSG, DPR, DHI, 2TL

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±0.10	0±0/150 ( 0.1± 0.2%)	1.78±0.10	2±2/207 ( 1.2± 0.9%)
2	B	0.00±0.00	-	0.00±0.00	-
All	All	1.41	1/1200 ( 0.1%)	1.78	20/1656 ( 1.2%)

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
2	B	0.0±0.0	0.1±0.3
All	All	0	3

All unique bond outliers are listed below.

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)	Models	
								Worst	Total
1	A	39	VAL	CA-CB	5.06	1.65	1.54	6	1

5 of 17 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	50	VAL	CG1-CB-CG2	-9.15	96.26	110.90	2	1
1	A	36	VAL	CG1-CB-CG2	-7.96	98.16	110.90	6	2
1	A	42	ALA	N-CA-CB	-7.59	99.48	110.10	6	1
1	A	40	VAL	CG1-CB-CG2	-6.94	99.79	110.90	2	2
1	A	43	THR	OG1-CB-CG2	-6.45	95.17	110.00	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	31	ILE	Mainchain	2
2	B	101	DAR	Peptide	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	150	182	182	0±0
2	B	113	119	114	0±0
All	All	2104	2408	2360	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
2:B:106:DLE:HD21	2:B:110:DAR:NH1	0.41	2.30	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	22/55 (40%)	20±1 (93±3%)	2±1 (7±3%)	0±0 (0±0%)	100	100
2	B	0	-	-	-	-	-
All	All	176/536 (33%)	164 (93%)	12 (7%)	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	A	17/45 (38%)	17±0 (99±2%)	0±0 (1±2%)	84 97
2	B	0	-	-	-
All	All	136/360 (38%)	135 (99%)	1 (1%)	84 97

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

Mol	Chain	Res	Type	Models (Total)
1	A	35	MET	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	645
Number of shifts mapped to atoms	645
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	12

#### 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	$1.69 \pm 0.22$	Should be applied
$^{13}\text{C}_\beta$	48	$2.86 \pm 0.05$	Should be applied
$^{13}\text{C}'$	52	$2.37 \pm 0.16$	Should be applied
$^{15}\text{N}$	55	$1.11 \pm 0.29$	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 96%, i.e. 228 atoms were assigned a chemical shift out of a possible 237. 8 out of 8 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	$^1\text{H}$	$^{13}\text{C}$	$^{15}\text{N}$
Backbone	110/110 (100%)	44/44 (100%)	44/44 (100%)	22/22 (100%)
Sidechain	118/127 (93%)	65/69 (94%)	53/58 (91%)	0/0 (—%)

*Continued on next page...*

Continued from previous page...

	Total	<sup>1</sup> H	<sup>13</sup> C	<sup>15</sup> N
Aromatic	0/0 (—%)	0/0 (—%)	0/0 (—%)	0/0 (—%)
Overall	228/237 (96%)	109/113 (96%)	97/102 (95%)	22/22 (100%)

#### 7.1.4 Statistically unusual chemical shifts [i](#)

The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

Mol	Chain	Res	Type	Atom	Shift, ppm	Expected range, ppm	Z-score
1	A	52	PRO	C	0.00	184.42 – 169.02	-114.8
1	A	52	PRO	CD	0.00	55.31 – 45.41	-50.9
1	A	52	PRO	CA	0.00	71.13 – 55.53	-40.6
1	A	52	PRO	CB	0.00	37.79 – 25.89	-26.8
1	A	52	PRO	CG	0.00	32.66 – 21.76	-25.0
1	A	52	PRO	N	0.00	167.35 – 100.35	-20.0
1	A	52	PRO	HA	0.00	6.05 – 2.75	-13.3
1	A	52	PRO	HD2	0.00	5.45 – 1.85	-10.1
1	A	52	PRO	HD3	0.00	5.52 – 1.72	-9.5
1	A	52	PRO	HG2	0.00	3.48 – 0.38	-6.2
1	A	52	PRO	HB2	0.00	3.82 – 0.32	-5.9
1	A	52	PRO	HG3	0.00	3.56 – 0.26	-5.8

#### 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:

