Full wwPDB NMR Structure Validation Report (i)

Dec 19, 2020 – 08:04 am GMT

PDB ID	:	6YHX
Title	:	Solution NMR Structure of APP I45T mutant TMD
Authors	:	Silber, M.; Muhle-Goll, C.
Deposited on	:	2020-03-31

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

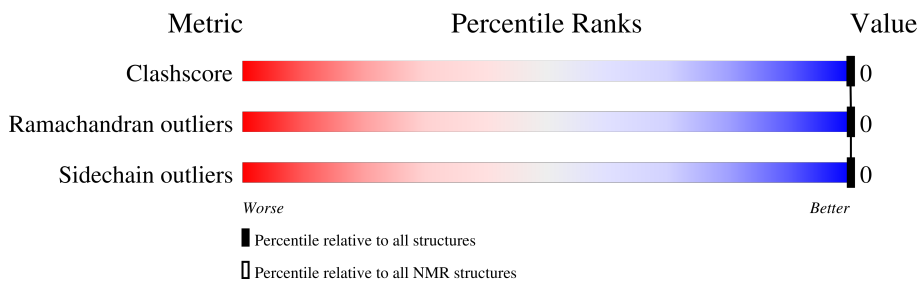|		
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.16
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.16

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 74%.

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	Percentile Ranks	Value
Clashscore		0
Ramachandran outliers		0
Sidechain outliers		0

Worse   *Better*

▋ Percentile relative to all structures

▯ Percentile relative to all NMR structures

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 <=5%

Mol	Chain	Length	Quality of chain
1	A	30	40% 60%

2 Ensemble composition and analysis (i)

This entry contains 20 models. Model 9 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	A:37-A:48 (12)	0.18	9

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 3 clusters and 4 single-model clusters were found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 469 atoms, of which 257 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Amyloid-beta precursor protein I45T mutant.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	30	469	139	257	35	36	2	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	45	THR	ILE	variant	UNP P05067

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 I45T mutant

4.2   Scores per residue for each member of the ensemble
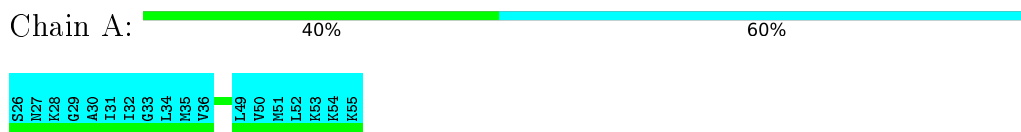
Colouring as in section 4.1 above.

4.2.1   Score per residue for model 1
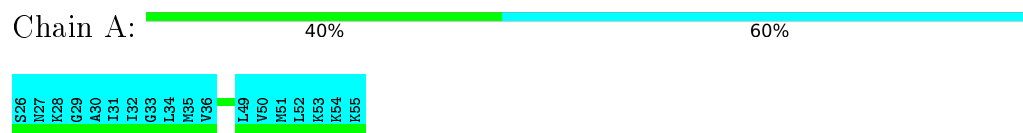
- Molecule 1: Amyloid-beta precursor protein I45T mutant
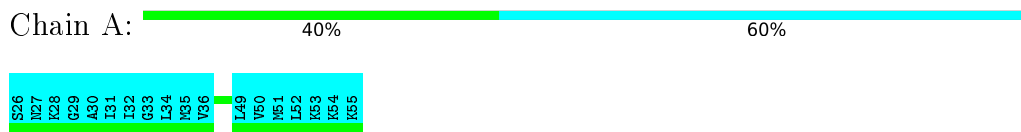
4.2.2   Score per residue for model 2

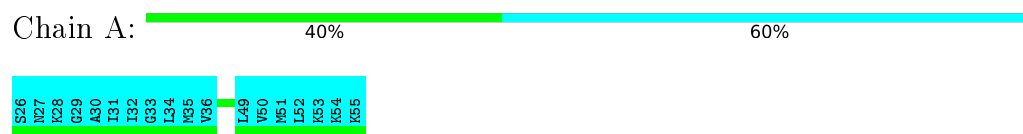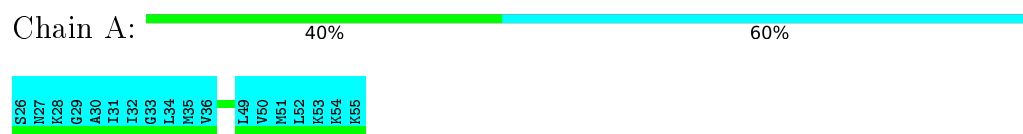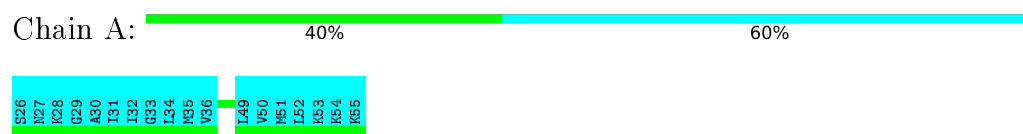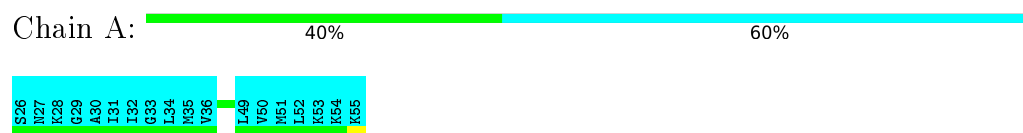- Molecule 1: Amyloid-beta precursor protein I45T mutant

4.2.3 Score per residue for model 3

• Molecule 1: Amyloid-beta precursor protein I45T mutant

Chain A:

4.2.4 Score per residue for model 4

• Molecule 1: Amyloid-beta precursor protein I45T mutant

Chain A:

4.2.5 Score per residue for model 5

• Molecule 1: Amyloid-beta precursor protein I45T mutant

Chain A:

4.2.6 Score per residue for model 6

• Molecule 1: Amyloid-beta precursor protein I45T mutant

Chain A:

4.2.7 Score per residue for model 7

• Molecule 1: Amyloid-beta precursor protein I45T mutant
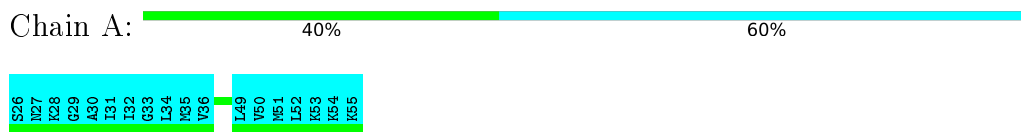
Chain A:

4.2.8 Score per residue for model 8

• Molecule 1: Amyloid-beta precursor protein I45T mutant

Chain A:

4.2.9 Score per residue for model 9 (medoid)

• Molecule 1: Amyloid-beta precursor protein I45T mutant

Chain A:

4.2.10 Score per residue for model 10

• Molecule 1: Amyloid-beta precursor protein I45T mutant

Chain A:

4.2.11 Score per residue for model 11

• Molecule 1: Amyloid-beta precursor protein I45T mutant

Chain A:

4.2.12 Score per residue for model 12

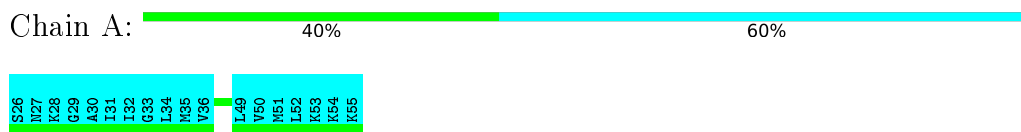• Molecule 1: Amyloid-beta precursor protein I45T mutant
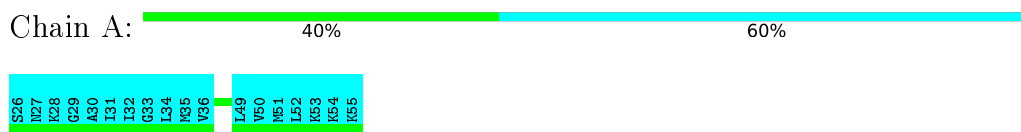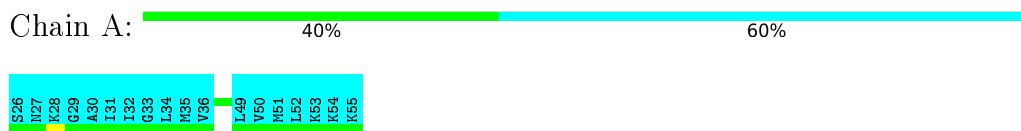
Chain A:

4.2.13 Score per residue for model 13

• Molecule 1: Amyloid-beta precursor protein I45T mutant

Chain A:

40%   60%

S26 N27 K28 G29 A30 I31 I32 G33 L34 M35 V36   L49 V50 M51 L52 K53 K54 K55

4.2.14 Score per residue for model 14

• Molecule 1: Amyloid-beta precursor protein I45T mutant

Chain A:

40%   60%

S26 N27 K28 G29 A30 I31 I32 G33 L34 M35 V36   L49 V50 M51 L52 K53 K54 K55

4.2.15 Score per residue for model 15

• Molecule 1: Amyloid-beta precursor protein I45T mutant

Chain A:

40%   60%

S26 N27 K28 G29 A30 I31 I32 G33 L34 M35 V36   L49 V50 M51 L52 K53 K54 K55

4.2.16 Score per residue for model 16

• Molecule 1: Amyloid-beta precursor protein I45T mutant

Chain A:

40%   60%

S26 N27 K28 G29 A30 I31 I32 G33 L34 M35 V36   L49 V50 M51 L52 K53 K54 K55

4.2.17 Score per residue for model 17

• Molecule 1: Amyloid-beta precursor protein I45T mutant

Chain A:

40%   60%

S26 N27 K28 G29 A30 I31 I32 G33 L34 M35 V36   L49 V50 M51 L52 K53 K54 K55

4.2.18 Score per residue for model 18

• Molecule 1: Amyloid-beta precursor protein I45T mutant

Chain A:

4.2.19 Score per residue for model 19
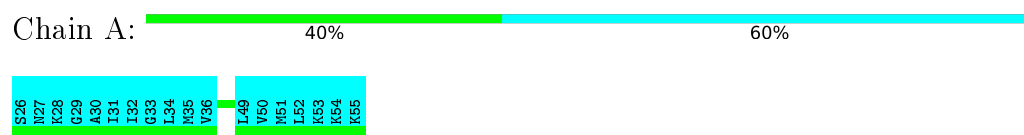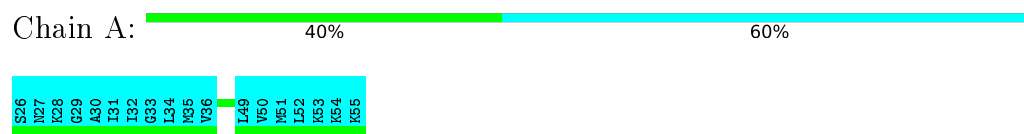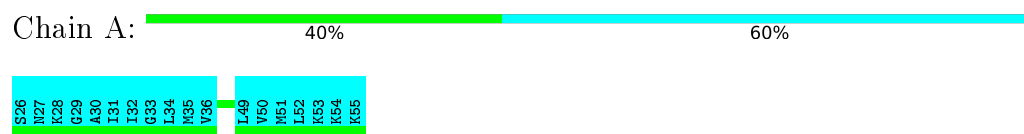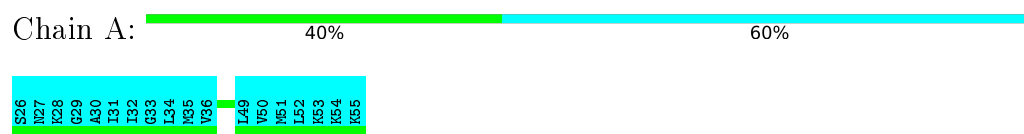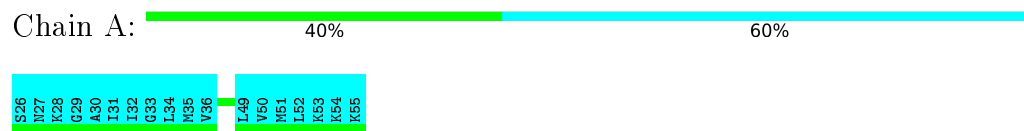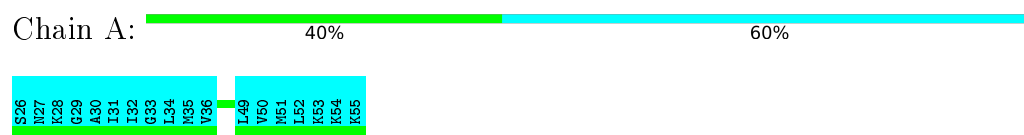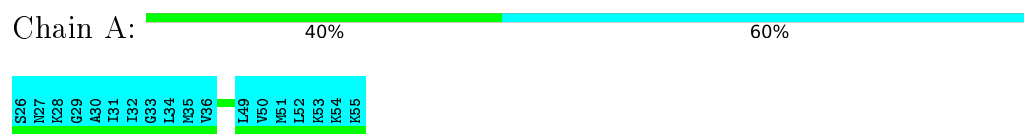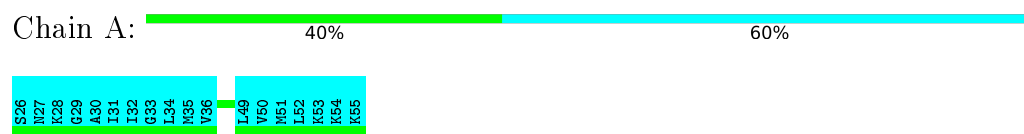
• Molecule 1: Amyloid-beta precursor protein I45T mutant

Chain A:

4.2.20 Score per residue for model 20

• Molecule 1: Amyloid-beta precursor protein I45T mutant

Chain A:

5 Refinement protocol and experimental data overview (i)

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

Of the 400 calculated structures, 20 were deposited, based on the following criterion: *20 structures for lowest energy.*

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

Software name	Classification	Version
CNS	refinement	1.2.1
ARIA	structure calculation	2.3.2

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	310
Number of shifts mapped to atoms	310
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	74%

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
All	All	1560	1800	1800	-

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 -.

There are no clashes.

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	12/30 (40%)	12±0 (100±0%)	0±0 (0±0%)	0±0 (0±0%)	100	100
All	All	240/600 (40%)	240 (100%)	0 (0%)	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	9/24 (38%)	9±0 (100±0%)	0±0 (0±0%)	100	100
All	All	180/480 (38%)	180 (100%)	0 (0%)	100	100

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

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)

The completeness of assignment taking into account all chemical shift lists is 74% for the well-defined parts and 69% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *i45t_shifts.str*

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	310
Number of shifts mapped to atoms	310
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
$^{13}C_\alpha$	30	-0.25 ± 0.22	None needed (< 0.5 ppm)
$^{13}C_\beta$	26	0.90 ± 0.10	Should be applied
$^{13}C'$	0	—	None (insufficient data)
^{15}N	0	—	None (insufficient data)

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 74%, i.e. 86 atoms were assigned a chemical shift out of a possible 116. 4 out of 4 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^1H	^{13}C	^{15}N
Backbone	36/60 (60%)	24/24 (100%)	12/24 (50%)	0/12 (0%)
Sidechain	50/56 (89%)	26/29 (90%)	24/27 (89%)	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	86/116 (74%)	50/53 (94%)	36/51 (71%)	0/12 (0%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 69%, i.e. 235 atoms were assigned a chemical shift out of a possible 343. 6 out of 9 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	89/150 (59%)	59/60 (98%)	30/60 (50%)	0/30 (0%)
Sidechain	146/193 (76%)	85/109 (78%)	61/79 (77%)	0/5 (0%)
Aromatic	0/0 (—%)	0/0 (—%)	0/0 (—%)	0/0 (—%)
Overall	235/343 (69%)	144/169 (85%)	91/139 (65%)	0/35 (0%)

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: