



wwPDB NMR Structure Validation Summary Report ⓘ

Mar 19, 2026 – 10:31 PM UTC

PDB ID : 8BOO / pdb_00008boo
BMRB ID : 51673
Title : Spatial structure of amyloidogenic SEM1(45-67) peptide
Authors : Blokhin, D.S.; Osetrina, D.A.
Deposited on : 2022-11-15

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 ⓘ 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 ⓘ](#)) were used in the production of this report:

MolProbity : 4-5-2 with Phenix2.0
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
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) : 2.49

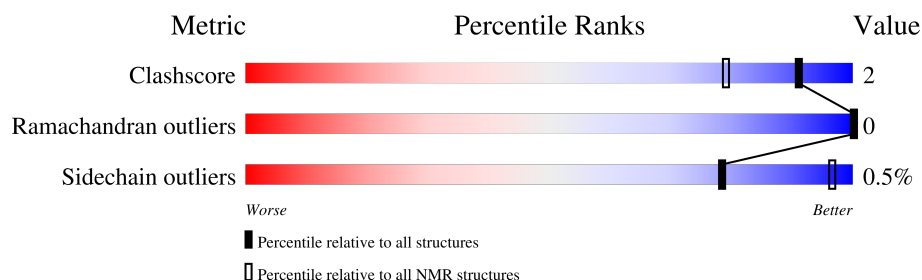
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

SOLUTION NMR

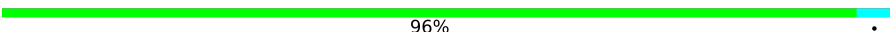
The overall completeness of chemical shifts assignment was not calculated.

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 (#Entries)	NMR archive (#Entries)
Clashscore	229148	14424
Ramachandran outliers	224038	12848
Sidechain outliers	223484	12823

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	23	 96%

2 Ensemble composition and analysis

This entry contains 11 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:46-A:67 (22)	0.37	6

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	1, 2, 4, 5, 6, 7, 8, 10, 11
2	3, 9

3 Entry composition [i](#)

There is only 1 type of molecule in this entry. The entry contains 354 atoms, of which 172 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Semenogelin-1.

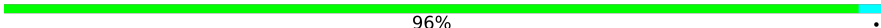
Mol	Chain	Residues	Atoms					Trace
1	A	23	Total	C	H	N	O	0
			354	111	172	33	38	

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: Semenogelin-1

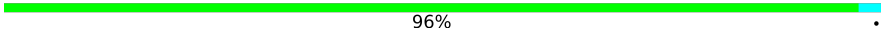
Chain A:  96% .



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: Semenogelin-1

Chain A:  96% .



5 Refinement protocol and experimental data overview

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

Of the 1000 calculated structures, 11 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	
X-PLOR NIH	structure calculation	

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	125
Number of shifts mapped to atoms	0
Number of unparsed shifts	0
Number of shifts with mapping errors	125
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	0%

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
1	A	178	167	166	1±1
All	All	1958	1837	1826	6

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 2.

All unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:58:GLU:O	1:A:64:SER:HB2	0.48	2.09	3	1
1:A:61:GLY:HA3	1:A:64:SER:OG	0.43	2.13	1	1
1:A:50:GLY:H	1:A:64:SER:HB3	0.41	1.75	5	4

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	21/23 (91%)	17±1 (80±3%)	4±1 (20±3%)	0±0 (0±0%)	100	100
All	All	231/253 (91%)	185 (80%)	46 (20%)	0 (0%)	100	100

There are no Ramachandran outliers.

6.3.2 Protein sidechains ⓘ

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	19/19 (100%)	19±0 (100±2%)	0±0 (0±2%)	78	96
All	All	209/209 (100%)	208 (100%)	1 (0%)	78	96

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

Mol	Chain	Res	Type	Models (Total)
1	A	57	THR	1

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 oligosaccharides in this entry.

6.6 Ligand geometry ⓘ

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

The completeness of assignment taking into account all chemical shift lists is 0% for the well-defined parts and 0% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *assigned_chemical_shifts_1*

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	125
Number of shifts mapped to atoms	0
Number of unparsed shifts	0
Number of shifts with mapping errors	125
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	0

The following assigned chemical shifts were not mapped to the molecules present in the coordinate file.

- No matching atom found in the structure. First 5 (of 125) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	1	GLY	HA2	3.763	0.004	1
1	A	1	GLY	C	169.876	0.011	1
1	A	1	GLY	CA	43.23	?	1
1	A	2	GLN	H	8.538	0.011	1
1	A	2	GLN	HA	4.213	0.004	1
1	A	2	GLN	HB2	1.825	0.016	2
1	A	2	GLN	HG2	2.18	0.005	1
1	A	2	GLN	C	175.533	0.005	1
1	A	2	GLN	CA	55.86	0.035	1
1	A	2	GLN	CB	29.798	0.005	1
1	A	2	GLN	CG	33.635	0.022	1
1	A	2	GLN	CD	180.211	?	1
1	A	2	GLN	N	119.551	?	1
1	A	3	HIS	H	8.583	0.014	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	3	HIS	HA	4.6	0.004	1
1	A	3	HIS	HB2	3.104	0.014	2
1	A	3	HIS	HB3	3.021	0.007	2
1	A	3	HIS	HD2	7.145	0.005	1
1	A	3	HIS	HE1	8.497	0.005	1
1	A	3	HIS	C	173.812	0.122	1
1	A	3	HIS	CA	55.092	0.102	1
1	A	3	HIS	CB	29.204	0.012	1
1	A	3	HIS	CG	130.843	0.083	1
1	A	3	HIS	CD2	120.062	0.034	1
1	A	3	HIS	CE1	136.417	0.041	1
1	A	3	HIS	N	120.361	?	1
1	A	4	TYR	H	8.343	0.014	1
1	A	4	TYR	HA	4.55	0.002	1
1	A	4	TYR	HB2	2.912	0.006	2
1	A	4	TYR	HB3	2.861	0.007	2
1	A	4	TYR	HD1	7.036	0.005	1
1	A	4	TYR	HD2	7.036	0.005	1
1	A	4	TYR	C	175.603	0.011	1
1	A	4	TYR	CA	57.819	0.006	1
1	A	4	TYR	CB	39.124	0.006	1
1	A	4	TYR	CG	130.284	0.003	1
1	A	4	TYR	CD1	133.255	0.003	1
1	A	4	TYR	CD2	133.255	0.003	1
1	A	4	TYR	N	123.045	?	1
1	A	5	SER	H	8.354	0.013	1
1	A	5	SER	HA	4.344	0.009	1
1	A	5	SER	HB2	3.736	0.014	1
1	A	5	SER	C	174.635	0.101	1
1	A	5	SER	CA	58.187	0.062	1
1	A	5	SER	CB	63.962	0.094	1
1	A	5	SER	N	119.126	?	1
1	A	6	GLY	H	7.718	0.007	1
1	A	6	GLY	N	110.237	?	1
1	A	7	GLN	HB3	1.865	0.008	2
1	A	8	LYS	H	8.396	0.003	1
1	A	8	LYS	HA	4.214	0.003	1
1	A	8	LYS	HB2	1.672	0.005	2
1	A	8	LYS	HB3	1.746	0.006	2
1	A	8	LYS	HG2	1.324	0.009	2
1	A	8	LYS	HG3	1.366	0.002	2

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	8	LYS	HD2	1.567	0.007	1
1	A	8	LYS	HE2	2.894	0.009	1
1	A	8	LYS	C	176.809	?	1
1	A	8	LYS	CA	56.069	?	1
1	A	8	LYS	CB	33.064	?	1
1	A	8	LYS	CG	24.833	?	1
1	A	8	LYS	CE	42.208	?	1
1	A	8	LYS	N	123.076	?	1
1	A	10	LYS	CD	29.065	0.01	1
1	A	12	GLN	HE21	7.432	0.001	1
1	A	12	GLN	HE22	6.775	0.004	1
1	A	12	GLN	NE2	112.286	0.011	1
1	A	13	THR	H	8.159	0.012	1
1	A	13	THR	HA	4.251	0.009	1
1	A	13	THR	HB	4.113	0.005	1
1	A	13	THR	HG21	1.11	0.005	1
1	A	13	THR	HG22	1.11	0.005	1
1	A	13	THR	HG23	1.11	0.005	1
1	A	13	THR	C	174.439	?	1
1	A	13	THR	CA	61.987	0.017	1
1	A	13	THR	CB	69.824	0.036	1
1	A	13	THR	CG2	21.627	?	1
1	A	13	THR	N	115.807	?	1
1	A	14	GLU	H	8.351	0.004	1
1	A	14	GLU	HB2	2.029	0.012	2
1	A	14	GLU	HB3	1.884	0.005	2
1	A	14	GLU	HG2	2.34	0.015	1
1	A	14	GLU	C	176.0	0.083	1
1	A	14	GLU	CA	55.809	?	1
1	A	14	GLU	CB	29.276	0.133	1
1	A	14	GLU	CG	33.385	0.075	1
1	A	14	GLU	CD	180.437	0.053	1
1	A	14	GLU	N	122.987	?	1
1	A	16	LYS	HZ1	7.429	0.002	1
1	A	16	LYS	HZ2	7.429	0.002	1
1	A	16	LYS	HZ3	7.429	0.002	1
1	A	19	PHE	H	8.141	0.003	1
1	A	19	PHE	HA	4.604	0.008	1
1	A	19	PHE	HB2	3.083	0.007	2
1	A	19	PHE	HB3	2.945	0.005	2
1	A	19	PHE	HD1	7.164	0.004	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	19	PHE	HD2	7.164	0.004	1
1	A	19	PHE	C	175.444	0.036	1
1	A	19	PHE	CA	57.649	0.042	1
1	A	19	PHE	CB	39.77	?	1
1	A	19	PHE	CG	138.854	0.016	1
1	A	19	PHE	CD1	131.907	0.017	1
1	A	19	PHE	CD2	131.907	0.017	1
1	A	19	PHE	N	121.696	?	1
1	A	21	ILE	H	7.988	0.01	1
1	A	21	ILE	HA	4.015	0.005	1
1	A	21	ILE	HB	1.668	0.004	1
1	A	21	ILE	HG12	1.322	0.004	2
1	A	21	ILE	HG13	1.02	0.004	2
1	A	21	ILE	HG21	0.642	0.004	1
1	A	21	ILE	HG22	0.642	0.004	1
1	A	21	ILE	HG23	0.642	0.004	1
1	A	21	ILE	HD11	0.759	0.003	1
1	A	21	ILE	HD12	0.759	0.003	1
1	A	21	ILE	HD13	0.759	0.003	1
1	A	21	ILE	C	175.43	0.002	1
1	A	21	ILE	CA	61.149	0.047	1
1	A	21	ILE	CB	38.884	0.005	1
1	A	21	ILE	CG1	27.36	0.024	1
1	A	21	ILE	CG2	17.331	0.038	1
1	A	21	ILE	CD1	13.012	?	1
1	A	21	ILE	N	122.606	?	1
1	A	23	TYR	CE1	118.144	?	1
1	A	23	TYR	CE2	118.144	?	1
1	A	23	TYR	CZ	156.998	?	1

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 0%, i.e. 0 atoms were assigned a chemical shift out of a possible 288. 0 out of 0 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	¹ H	¹³ C	¹⁵ N
Backbone	0/113 (0%)	0/47 (0%)	0/44 (0%)	0/22 (0%)
Sidechain	0/139 (0%)	0/87 (0%)	0/44 (0%)	0/8 (0%)
Aromatic	0/36 (0%)	0/17 (0%)	0/17 (0%)	0/2 (0%)
Overall	0/288 (0%)	0/151 (0%)	0/105 (0%)	0/32 (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. If well-defined core and ill-defined regions are not identified then it is shown as gray bars.

Random coil index (RCI) for chain A:

