



wwPDB NMR Structure Validation Summary Report ⓘ

Jun 23, 2024 – 12:12 AM EDT

PDB ID : 6ER0
BMRB ID : 34185
Title : 6th KOW domain of human hSpt5
Authors : Hahn, L.; Schweimer, K.; Roesch, P.; Woehrl, B.M.
Deposited on : 2017-10-16

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 : 4.02b-467
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
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) : 2.37.1

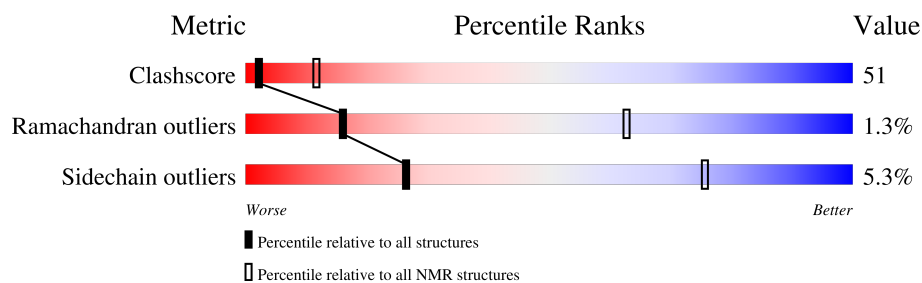
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

SOLUTION NMR

The overall completeness of chemical shifts assignment is 87%.

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	127	

2 Ensemble composition and analysis

This entry contains 15 models. Model 1 is the overall representative, medoid model (most similar to other models).

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:18-A:127 (110)	0.78	1

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, 3, 4, 5, 6, 7, 8, 10, 12, 13, 14, 15
2	9, 11

3 Entry composition [i](#)

There is only 1 type of molecule in this entry. The entry contains 1785 atoms, of which 904 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Transcription elongation factor SPT5.

Mol	Chain	Residues	Atoms						Trace
1	A	114	Total	C	H	N	O	S	0
			1785	549	904	149	180	3	

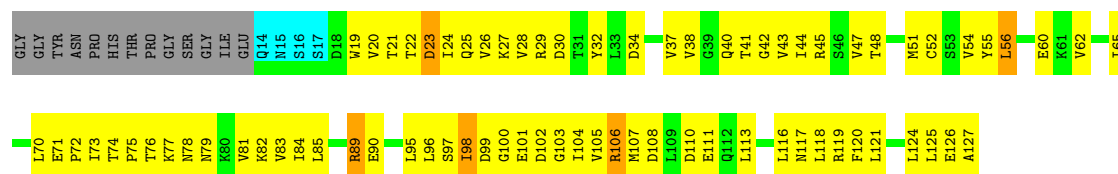
4 Residue-property plots [i](#)

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: Transcription elongation factor SPT5

Chain A: 



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

The representative model is number 1. Colouring as in section 4.1 above.

- Molecule 1: Transcription elongation factor SPT5

Chain A: 



5 Refinement protocol and experimental data overview

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

Of the 120 calculated structures, 15 were deposited, based on the following criterion: *structures with the lowest energy*.

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

Software name	Classification	Version
Xplor-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	1457
Number of shifts mapped to atoms	1349
Number of unparsed shifts	0
Number of shifts with mapping errors	108
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	87%

6 Model quality [i](#)

6.1 Standard geometry [i](#)

There are no covalent bond-length or bond-angle outliers.

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	4.9±0.2
All	All	0	74

There are no bond-length outliers.

There are no bond-angle outliers.

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	29	ARG	Sidechain	15
1	A	45	ARG	Sidechain	15
1	A	106	ARG	Sidechain	15
1	A	119	ARG	Sidechain	15
1	A	89	ARG	Sidechain	14

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	852	880	880	89±7
All	All	12780	13200	13200	1337

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 51.

5 of 600 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:95:LEU:CD2	1:A:105:VAL:HG12	0.99	1.86	15	1
1:A:105:VAL:HG11	1:A:121:LEU:CD2	0.97	1.89	7	4
1:A:94:VAL:HG23	1:A:108:ASP:OD2	0.94	1.60	14	2
1:A:95:LEU:HD21	1:A:105:VAL:HG12	0.94	1.37	15	1
1:A:26:VAL:CG2	1:A:54:VAL:HG11	0.93	1.94	1	3

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	109/127 (86%)	104±2 (95±2%)	4±2 (4±2%)	1±1 (1±1%)	16	63
All	All	1635/1905 (86%)	1554 (95%)	59 (4%)	22 (1%)	16	63

5 of 8 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	100	GLY	10
1	A	49	GLY	3
1	A	89	ARG	3
1	A	75	PRO	2
1	A	34	ASP	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	99/112 (88%)	94±1 (95±1%)	5±1 (5±1%)	26	75
All	All	1485/1680 (88%)	1407 (95%)	78 (5%)	26	75

5 of 20 unique residues with a non-rotameric sidechain are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	56	LEU	15
1	A	23	ASP	11
1	A	98	ILE	10
1	A	74	THR	7
1	A	85	LEU	6

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

The completeness of assignment taking into account all chemical shift lists is 87% for the well-defined parts and 87% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *kow6.bmrB*

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	1457
Number of shifts mapped to atoms	1349
Number of unparsed shifts	0
Number of shifts with mapping errors	108
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	1

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 108) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	1	GLY	C	174.74	0.20	1
1	A	2	GLY	H	8.26	0.03	1
1	A	2	GLY	HA2	3.92	0.03	2
1	A	2	GLY	HA3	3.92	0.03	2
1	A	2	GLY	C	173.63	0.20	1
1	A	2	GLY	CA	45.36	0.20	1
1	A	2	GLY	N	108.73	0.20	1
1	A	3	TYR	H	8.16	0.03	1
1	A	3	TYR	HA	4.52	0.03	1
1	A	3	TYR	HB2	2.94	0.03	2
1	A	3	TYR	HB3	2.94	0.03	2
1	A	3	TYR	HD1	7.07	0.03	1
1	A	3	TYR	HD2	7.07	0.03	1
1	A	3	TYR	HE1	6.78	0.03	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	3	TYR	HE2	6.78	0.03	1
1	A	3	TYR	C	174.9	0.20	1
1	A	3	TYR	CA	57.9	0.20	1
1	A	3	TYR	CB	38.97	0.20	1
1	A	3	TYR	CD1	133.27	0.20	1
1	A	3	TYR	CE1	118.24	0.20	1
1	A	3	TYR	N	120.7	0.20	1
1	A	4	ASN	H	8.33	0.03	1
1	A	4	ASN	CA	50.49	0.20	1
1	A	4	ASN	CB	39.35	0.20	1
1	A	4	ASN	N	123.9	0.20	1
1	A	5	PRO	HA	4.27	0.03	1
1	A	5	PRO	HB2	2.19	0.03	2
1	A	5	PRO	HB3	2.19	0.03	2
1	A	5	PRO	C	176.6	0.20	1
1	A	5	PRO	CA	63.27	0.20	1
1	A	5	PRO	CB	32.15	0.20	1
1	A	5	PRO	CG	26.52	0.20	1
1	A	5	PRO	CD	50.72	0.20	1
1	A	6	HIS	H	8.29	0.03	1
1	A	6	HIS	HA	4.64	0.03	1
1	A	6	HIS	HB2	3.16	0.03	2
1	A	6	HIS	HB3	3.16	0.03	2
1	A	6	HIS	HD2	7.12	0.03	1
1	A	6	HIS	HE1	8.16	0.03	1
1	A	6	HIS	C	174.96	0.20	1
1	A	6	HIS	CA	55.84	0.20	1
1	A	6	HIS	CB	29.75	0.20	1
1	A	6	HIS	N	118.38	0.20	1
1	A	7	THR	H	8.01	0.03	1
1	A	7	THR	CA	59.74	0.20	1
1	A	7	THR	CB	69.75	0.20	1
1	A	7	THR	N	117.67	0.20	1
1	A	8	PRO	HA	4.41	0.03	1
1	A	8	PRO	HB2	1.96	0.03	2
1	A	8	PRO	HB3	2.33	0.03	2
1	A	8	PRO	HG2	2.08	0.03	2
1	A	8	PRO	HG3	2.0	0.03	2
1	A	8	PRO	HD2	3.81	0.03	2
1	A	8	PRO	HD3	3.7	0.03	2
1	A	8	PRO	C	177.58	0.20	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	8	PRO	CA	63.87	0.20	1
1	A	8	PRO	CB	32.26	0.20	1
1	A	8	PRO	CG	27.43	0.20	1
1	A	8	PRO	CD	51.11	0.20	1
1	A	9	GLY	H	8.6	0.03	1
1	A	9	GLY	HA2	4.0	0.03	2
1	A	9	GLY	HA3	4.0	0.03	2
1	A	9	GLY	C	174.39	0.20	1
1	A	9	GLY	CA	45.36	0.20	1
1	A	9	GLY	N	110.07	0.20	1
1	A	10	SER	H	8.21	0.03	1
1	A	10	SER	HA	4.46	0.03	1
1	A	10	SER	HB2	3.91	0.03	2
1	A	10	SER	HB3	3.91	0.03	2
1	A	10	SER	C	175.13	0.20	1
1	A	10	SER	CA	58.73	0.20	1
1	A	10	SER	CB	64.12	0.20	1
1	A	10	SER	N	115.47	0.20	1
1	A	11	GLY	H	8.54	0.03	1
1	A	11	GLY	HA2	4.0	0.03	2
1	A	11	GLY	HA3	4.0	0.03	2
1	A	11	GLY	C	174.14	0.20	1
1	A	11	GLY	CA	45.38	0.20	1
1	A	11	GLY	N	111.04	0.20	1
1	A	12	ILE	H	7.97	0.03	1
1	A	12	ILE	HA	4.2	0.03	1
1	A	12	ILE	HB	1.84	0.03	1
1	A	12	ILE	HG12	1.16	0.03	2
1	A	12	ILE	HG13	1.42	0.03	2
1	A	12	ILE	HG21	0.88	0.03	1
1	A	12	ILE	HG22	0.88	0.03	1
1	A	12	ILE	HG23	0.88	0.03	1
1	A	12	ILE	HD11	0.82	0.03	1
1	A	12	ILE	HD12	0.82	0.03	1
1	A	12	ILE	HD13	0.82	0.03	1
1	A	12	ILE	C	176.39	0.20	1
1	A	12	ILE	CA	61.36	0.20	1
1	A	12	ILE	CB	38.94	0.20	1
1	A	12	ILE	CG1	27.43	0.20	1
1	A	12	ILE	CG2	17.93	0.20	1
1	A	12	ILE	CD1	13.12	0.20	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	12	ILE	N	119.6	0.20	1
1	A	13	GLU	H	8.58	0.03	1
1	A	13	GLU	HA	4.28	0.03	1
1	A	13	GLU	HB2	1.95	0.03	2
1	A	13	GLU	HB3	2.05	0.03	2
1	A	13	GLU	HG2	2.28	0.03	2
1	A	13	GLU	HG3	2.19	0.03	2
1	A	13	GLU	C	176.53	0.20	1
1	A	13	GLU	CA	56.78	0.20	1
1	A	13	GLU	CB	30.0	0.20	1
1	A	13	GLU	CG	36.29	0.20	1
1	A	13	GLU	N	124.76	0.20	1

7.1.2 Chemical shift referencing ⓘ

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction \pm precision, ppm	Suggested action
$^{13}\text{C}_\alpha$	126	-0.15 ± 0.24	None needed (< 0.5 ppm)
$^{13}\text{C}_\beta$	112	0.20 ± 0.23	None needed (< 0.5 ppm)
$^{13}\text{C}'$	112	0.17 ± 0.17	None needed (< 0.5 ppm)
^{15}N	121	0.21 ± 0.32	None needed (< 0.5 ppm)

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 87%, i.e. 1314 atoms were assigned a chemical shift out of a possible 1505. 0 out of 27 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^1H	^{13}C	^{15}N
Backbone	539/555 (97%)	225/227 (99%)	207/220 (94%)	107/108 (99%)
Sidechain	742/903 (82%)	497/588 (85%)	243/285 (85%)	2/30 (7%)
Aromatic	33/47 (70%)	21/23 (91%)	11/22 (50%)	1/2 (50%)
Overall	1314/1505 (87%)	743/838 (89%)	461/527 (87%)	110/140 (79%)

7.1.4 Statistically unusual chemical shifts ⓘ

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 con-

taining paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

List Id	Chain	Res	Type	Atom	Shift, <i>ppm</i>	Expected range, <i>ppm</i>	Z-score
1	A	72	PRO	HD2	1.62	1.93 – 5.38	-5.9

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:

