



# wwPDB NMR Structure Validation Summary Report ⓘ

Dec 17, 2022 – 04:25 pm GMT

PDB ID : 6YQ5  
EMDB ID : EMD-10792  
BMRB ID : 27468  
Title : Hybrid structure of the SPP1 tail tube by solid-state NMR and cryo EM - NMR Ensemble  
Authors : Zinke, M.; Sachowsky, K.A.A.; Zinn-Justin, S.; Ravelli, R.; Schroder, G.F.; Habeck, M.; Lange, A.  
Deposited on : 2020-04-16

This is a wwPDB NMR Structure Validation Summary Report for a publicly released PDB entry.

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

EMDB validation analysis : **NOT EXECUTED**  
MolProbity : 4.02b-467  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
MapQ : **NOT EXECUTED**  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.31.3

# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

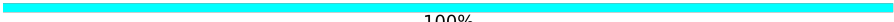
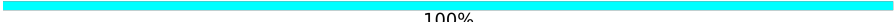
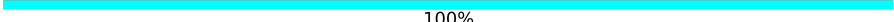
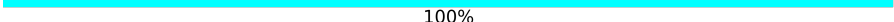
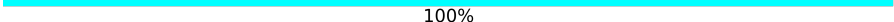
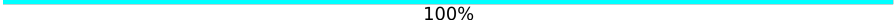
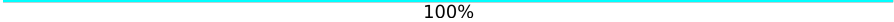
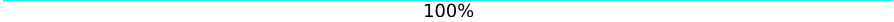
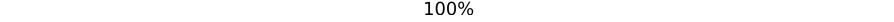
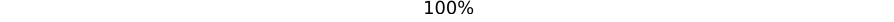
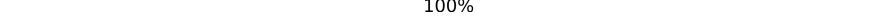
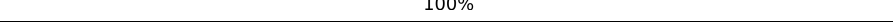
*ELECTRON MICROSCOPY, SOLID-STATE NMR*

The reported resolution of this entry is 4.00 Å.

The overall completeness of chemical shifts assignment is 4%.

There are no overall percentile quality scores available for this entry.

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	172	 100%
1	B	172	 100%
1	C	172	 100%
1	D	172	 100%
1	E	172	 100%
1	F	172	 100%
1	G	172	 100%
1	H	172	 100%
1	I	172	 100%
1	J	172	 100%
1	K	172	 100%
1	L	172	 100%

## 2 Ensemble composition and analysis ⓘ

This entry contains 10 models. The atoms present in the NMR models are not consistent. Some calculations may have failed as a result. All residues are included in the validation scores.

Cyrange was unable to find well-defined residues.

Error message: Cyrange did not run

NmrClust was unable to cluster the ensemble.

Error message: NmrClust did not run

### 3 Entry composition

There is only 1 type of molecule in this entry. The entry contains 30672 atoms, of which 14916 are hydrogens and 0 are deuteriums.

In the tables below, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Tail tube protein gp17.1\*.

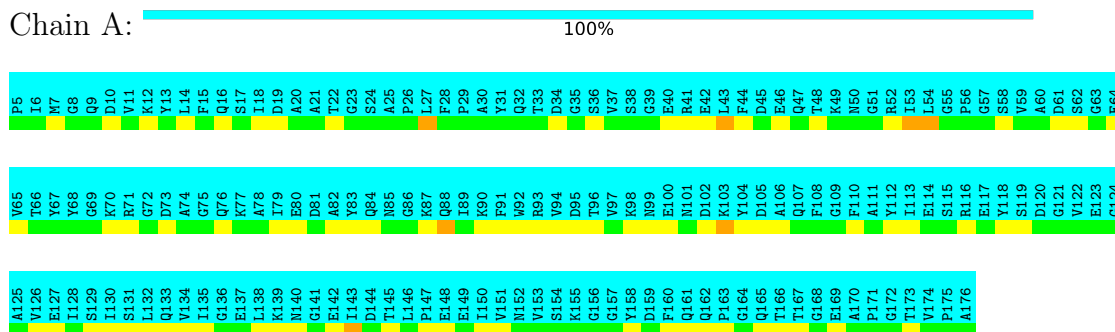
Mol	Chain	Residues	Atoms						AltConf	Trace
1	A	172	Total	C	H	N	O	S	0	
			2556	821	1243	216	275	1		
1	B	172	Total	C	H	N	O	S	0	
			2556	821	1243	216	275	1		
1	C	172	Total	C	H	N	O	S	0	
			2556	821	1243	216	275	1		
1	D	172	Total	C	H	N	O	S	0	
			2556	821	1243	216	275	1		
1	E	172	Total	C	H	N	O	S	0	
			2556	821	1243	216	275	1		
1	F	172	Total	C	H	N	O	S	0	
			2556	821	1243	216	275	1		
1	G	172	Total	C	H	N	O	S	0	
			2556	821	1243	216	275	1		
1	H	172	Total	C	H	N	O	S	0	
			2556	821	1243	216	275	1		
1	I	172	Total	C	H	N	O	S	0	
			2556	821	1243	216	275	1		
1	J	172	Total	C	H	N	O	S	0	
			2556	821	1243	216	275	1		
1	K	172	Total	C	H	N	O	S	0	
			2556	821	1243	216	275	1		
1	L	172	Total	C	H	N	O	S	0	
			2556	821	1243	216	275	1		

## 4 Residue-property plots

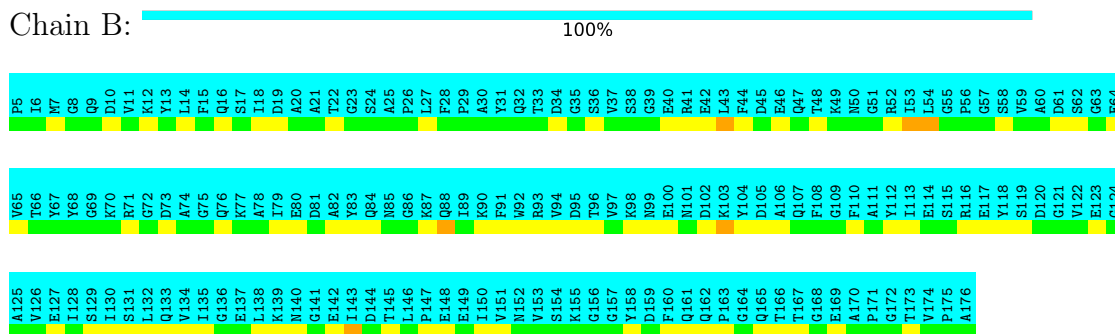
### 4.1 Average score per residue in the NMR ensemble

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: Tail tube protein gp17.1\*



- Molecule 1: Tail tube protein gp17.1\*



- Molecule 1: Tail tube protein gp17.1\*

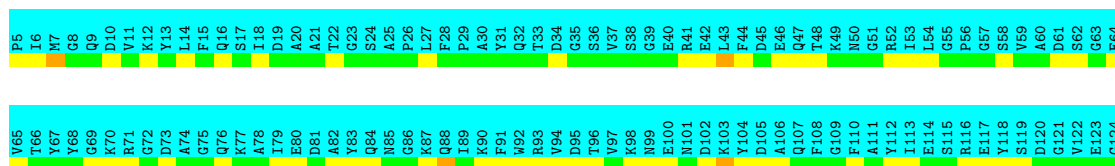






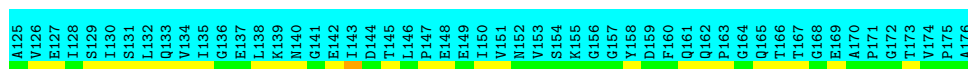
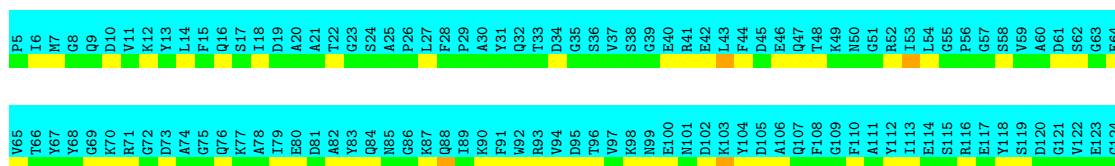
- Molecule 1: Tail tube protein gp17.1\*

Chain H: 100%



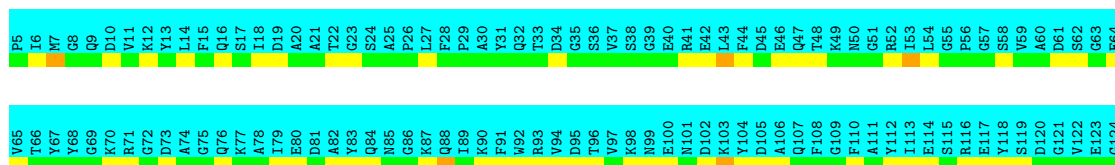
- Molecule 1: Tail tube protein gp17.1\*

Chain I: 100%



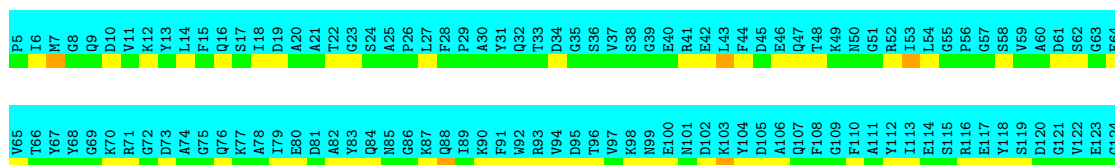
- Molecule 1: Tail tube protein gp17.1\*

Chain J: 100%



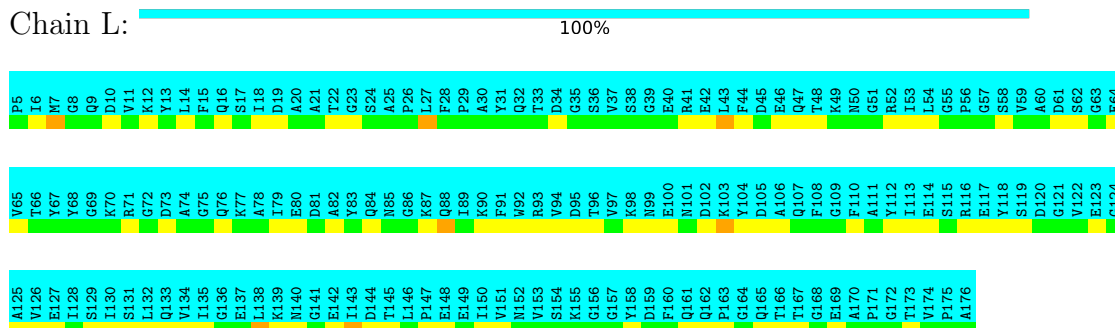
- Molecule 1: Tail tube protein gp17.1\*

Chain K: 100%



- Molecule 1: Tail tube protein gp17.1\*

Chain L:

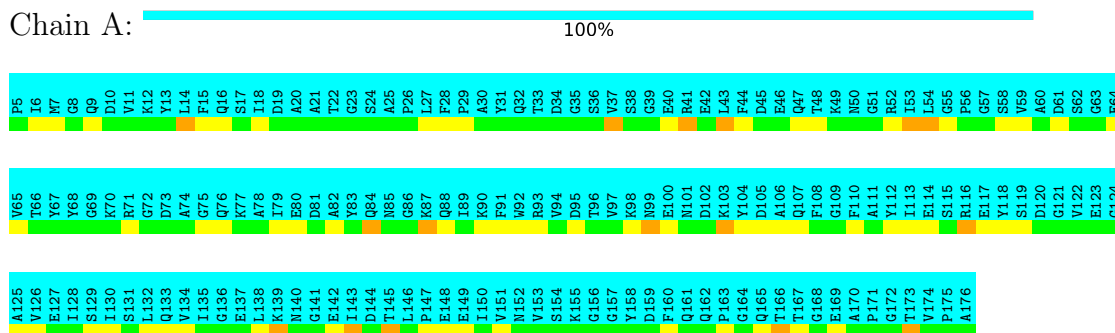


#### 4.2 Residue scores for the representative (author defined) model from the NMR ensemble

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

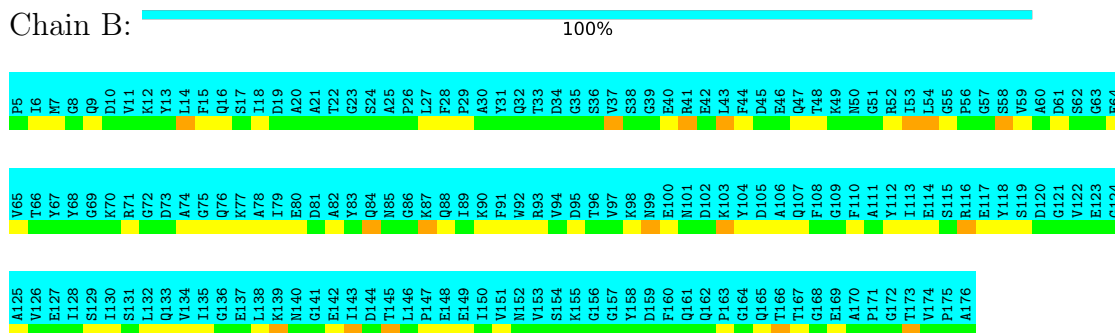
- Molecule 1: Tail tube protein gp17.1\*

Chain A:



- Molecule 1: Tail tube protein gp17.1\*

Chain B:

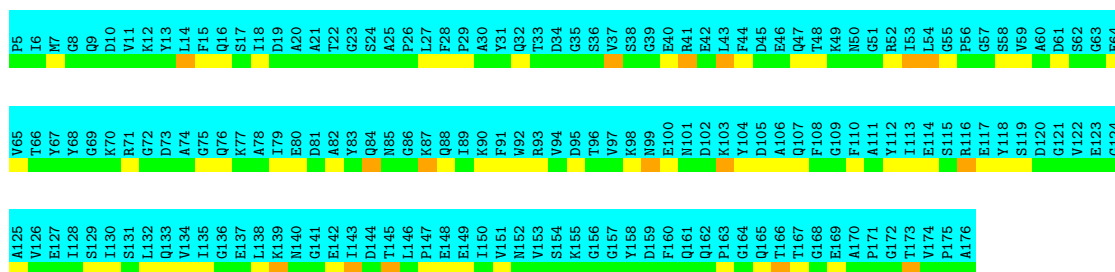


- Molecule 1: Tail tube protein gp17.1\*

Chain C:



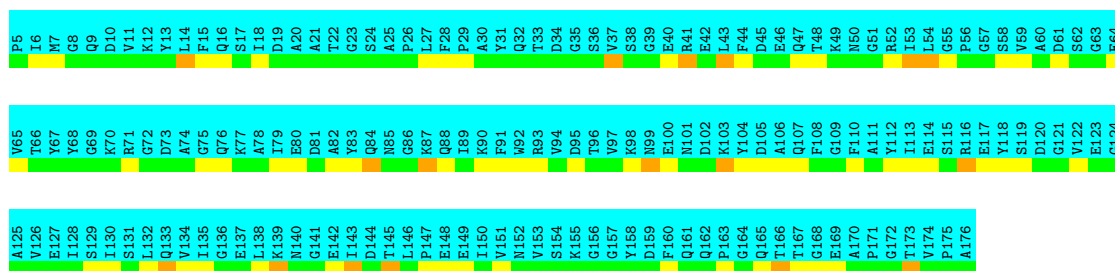




- Molecule 1: Tail tube protein gp17.1\*

Chain D:

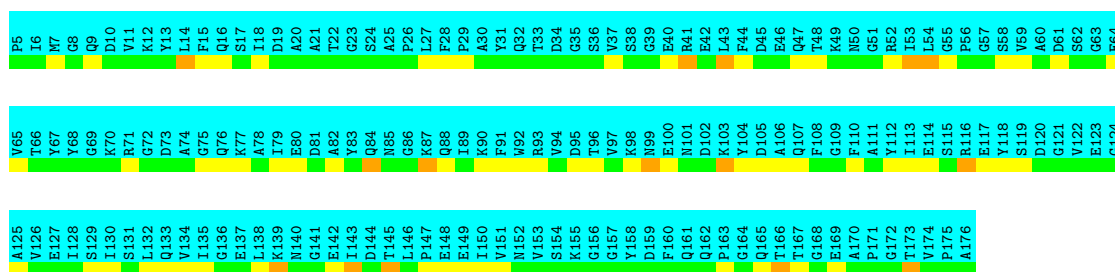
100%



- Molecule 1: Tail tube protein gp17.1\*

Chain E:

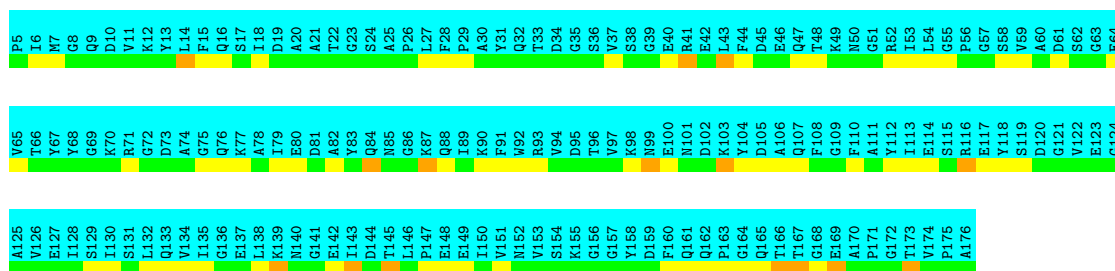
100%



- Molecule 1: Tail tube protein gp17.1\*

Chain F:

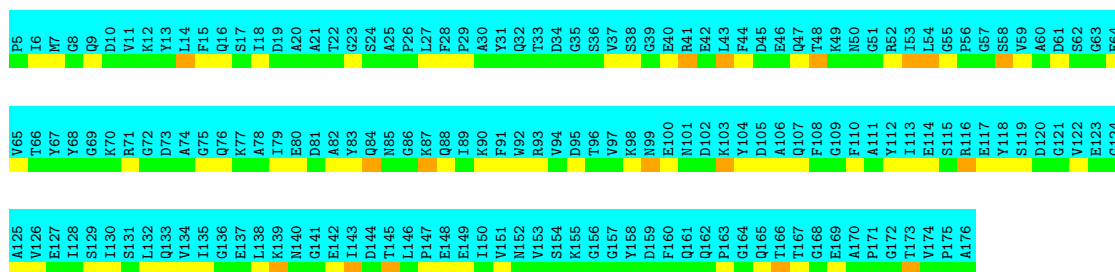
100%



- Molecule 1: Tail tube protein gp17.1\*

Chain G:

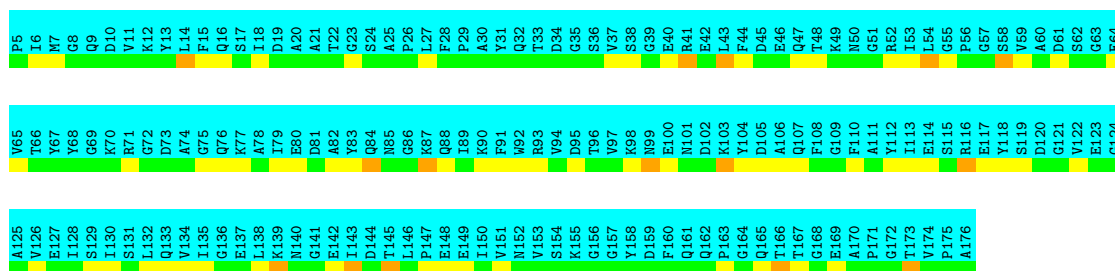
100%



- Molecule 1: Tail tube protein gp17.1\*

Chain H:

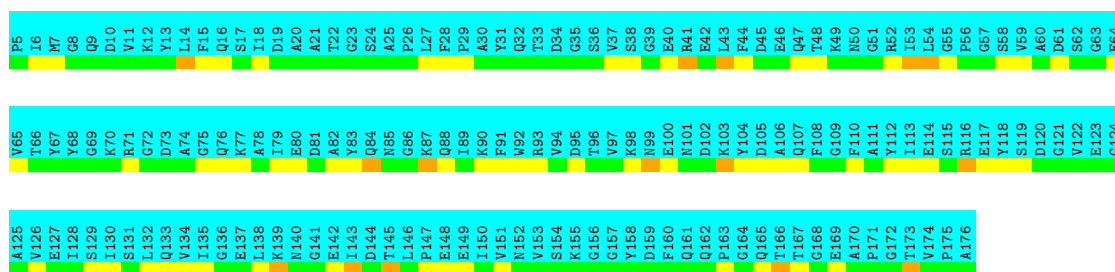
100%



- Molecule 1: Tail tube protein gp17.1\*

Chain I:

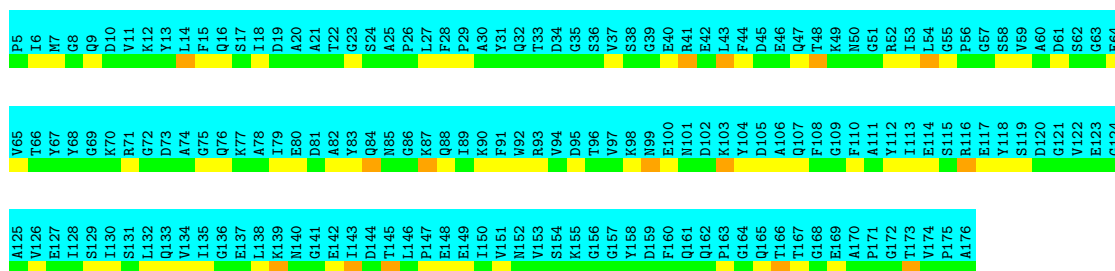
100%



- Molecule 1: Tail tube protein gp17.1\*

Chain J:

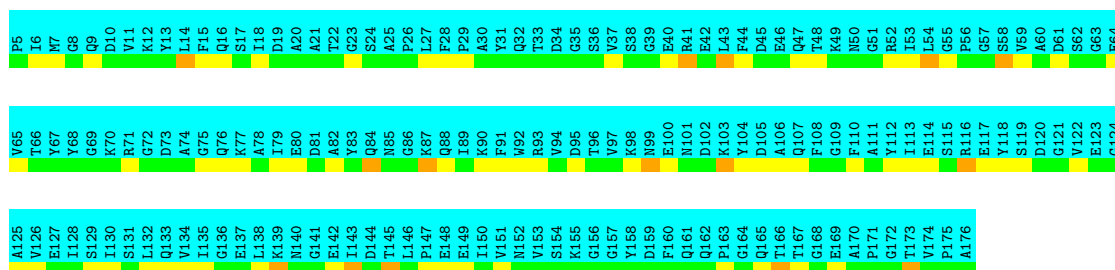
100%



- Molecule 1: Tail tube protein gp17.1\*

Chain K:

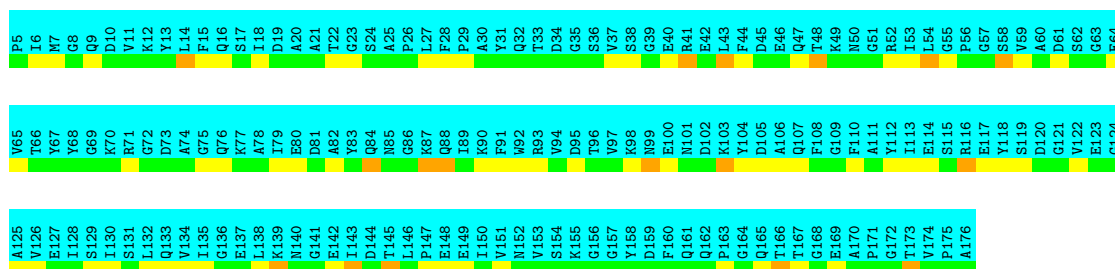
100%



- Molecule 1: Tail tube protein gp17.1\*

Chain L:

100%



## 5 Refinement protocol and experimental data overview

The models were refined using the following method: *na*.

Of the 500 calculated structures, 10 were deposited, based on the following criterion: *target function*.

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

Software name	Classification	Version
Inferential Structure Determination (ISD)	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	963
Number of shifts mapped to atoms	963
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	4%

Note: This is a solid-state NMR structure, where hydrogen atoms are typically not assigned a chemical shift value, which may lead to lower completeness of assignment measure.

## 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	0	0	0	0±0
1	B	0	0	0	0±0
1	C	0	0	0	0±0
1	D	0	0	0	0±0
1	E	0	0	0	0±0
1	F	0	0	0	0±0
1	G	0	0	0	0±0
1	H	0	0	0	0±0
1	I	0	0	0	0±0
1	J	0	0	0	0±0
1	K	0	0	0	0±0
1	L	0	0	0	0±0
All	All	0	0	0	-

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	0	-	-	-	-
1	B	0	-	-	-	-
1	C	0	-	-	-	-
1	D	0	-	-	-	-
1	E	0	-	-	-	-
1	F	0	-	-	-	-
1	G	0	-	-	-	-
1	H	0	-	-	-	-
1	I	0	-	-	-	-
1	J	0	-	-	-	-
1	K	0	-	-	-	-
1	L	0	-	-	-	-
All	All	0	-	-	-	-

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	0	-	-	-
1	B	0	-	-	-
1	C	0	-	-	-
1	D	0	-	-	-
1	E	0	-	-	-
1	F	0	-	-	-
1	G	0	-	-	-
1	H	0	-	-	-
1	I	0	-	-	-
1	J	0	-	-	-

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	K	0	-	-	-
1	L	0	-	-	-
All	All	0	-	-	-

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

The completeness of assignment taking into account all chemical shift lists is 4% for the well-defined parts and 4% for the entire structure.

### 7.1 Chemical shift list 1

File name: working\_cs.cif

Chemical shift list name: *assigned\_chem\_shift\_list\_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	963
Number of shifts mapped to atoms	963
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

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction $\pm$ precision, ppm	Suggested action
$^{13}\text{C}_\alpha$	161	$-0.10 \pm 0.14$	None needed ( $< 0.5$ ppm)
$^{13}\text{C}_\beta$	126	$0.58 \pm 0.12$	Should be applied
$^{13}\text{C}'$	161	$0.01 \pm 0.13$	None needed ( $< 0.5$ ppm)
$^{15}\text{N}$	151	$0.02 \pm 0.50$	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 4%, i.e. 861 atoms were assigned a chemical shift out of a possible 23880. 0 out of 228 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	$^1\text{H}$	$^{13}\text{C}$	$^{15}\text{N}$
Backbone	624/10128 (6%)	151/4032 (4%)	322/4128 (8%)	151/1968 (8%)
Sidechain	235/11988 (2%)	57/6996 (1%)	178/4476 (4%)	0/516 (0%)

*Continued on next page...*



Continued from previous page...

	Total	<sup>1</sup> H	<sup>13</sup> C	<sup>15</sup> N
Aromatic	2/1764 (0%)	1/924 (0%)	0/828 (0%)	1/12 (8%)
Overall	861/23880 (4%)	209/11952 (2%)	500/9432 (5%)	152/2496 (6%)

Note: This is a solid-state NMR structure, where hydrogen atoms are typically not assigned a chemical shift value, which may lead to lower completeness of assignment measure.

#### 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:

