



# Full wwPDB NMR Structure Validation Report ⓘ

Aug 9, 2020 – 11:52 PM BST

PDB ID : 6QTC  
Title : tSH2 domain of transcription elongation factor Spt6 complexed with tyrosine phosphorylated CTD  
Authors : Brazda, P.; Krejcikova, M.; Smirakova, E.; Kubicek, K.; Stefl, R.  
Deposited on : 2019-02-24

This is a Full wwPDB NMR Structure Validation 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 : 1.8.5 (274361), 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 : 2.13.1  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.13.1

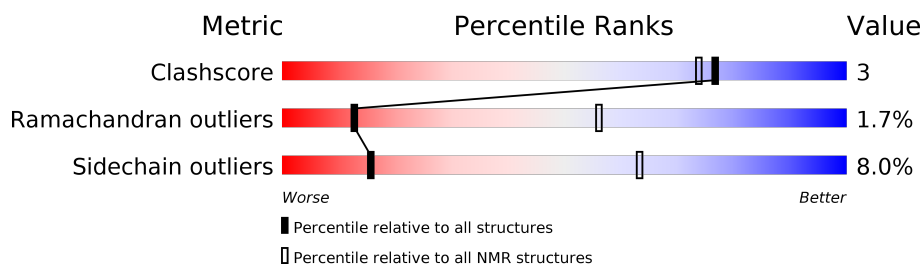
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*SOLUTION NMR*

The overall completeness of chemical shifts assignment is 59%.

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	196	 86% 9% . .
2	B	16	 100%

## 2 Ensemble composition and analysis

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:4-A:190 (187)	0.56	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 1 single-model cluster was found.

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

### 3 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 3504 atoms, of which 1737 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called tSH2 domain of transcription elongation factor Spt6.

Mol	Chain	Residues	Atoms						Trace
1	A	195	Total	C	H	N	O	S	0
			3276	1067	1633	278	297	1	

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	SER	THR	conflict	UNP Q6FLB1

- Molecule 2 is a protein called Tyrosine phosphorylated CTD.

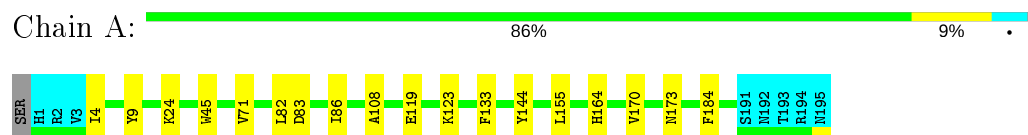
Mol	Chain	Residues	Atoms						Trace
2	B	16	Total	C	H	N	O	P	0
			228	72	104	16	34	2	

## 4 Residue-property plots [i](#)

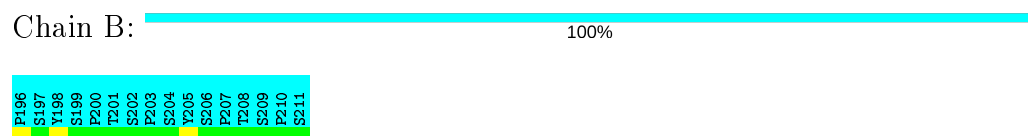
### 4.1 Average score per residue in the NMR ensemble

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: tSH2 domain of transcription elongation factor Spt6



- Molecule 2: Tyrosine phosphorylated CTD

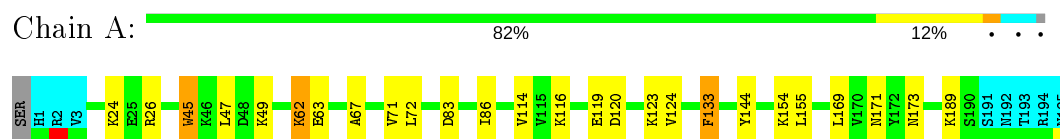


### 4.2 Scores per residue for each member of the ensemble

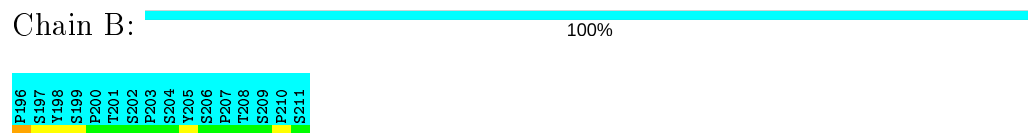
Colouring as in section 4.1 above.

#### 4.2.1 Score per residue for model 1

- Molecule 1: tSH2 domain of transcription elongation factor Spt6

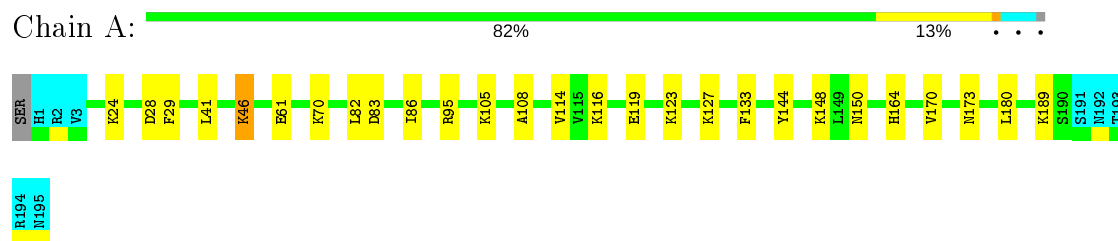


- Molecule 2: Tyrosine phosphorylated CTD

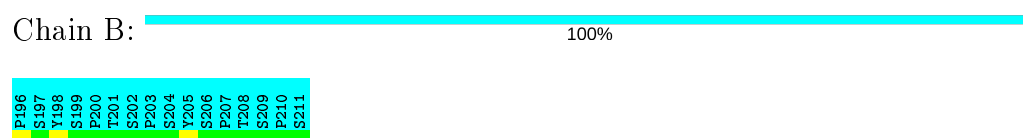


### 4.2.2 Score per residue for model 2

- Molecule 1: tSH2 domain of transcription elongation factor Spt6

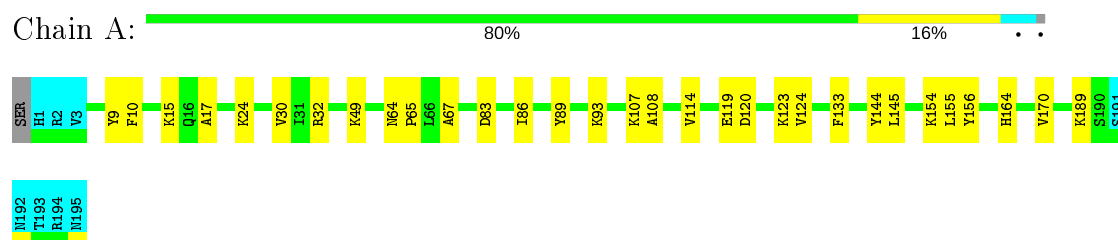


- Molecule 2: Tyrosine phosphorylated CTD

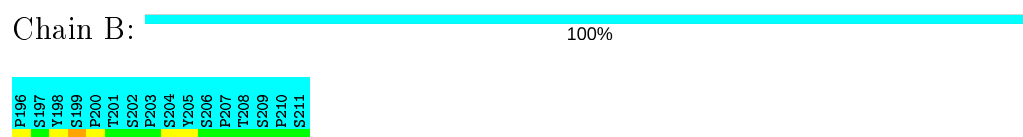


### 4.2.3 Score per residue for model 3

- Molecule 1: tSH2 domain of transcription elongation factor Spt6

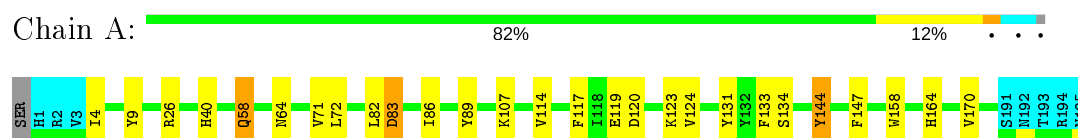


- Molecule 2: Tyrosine phosphorylated CTD



### 4.2.4 Score per residue for model 4

- Molecule 1: tSH2 domain of transcription elongation factor Spt6





- Molecule 2: Tyrosine phosphorylated CTD

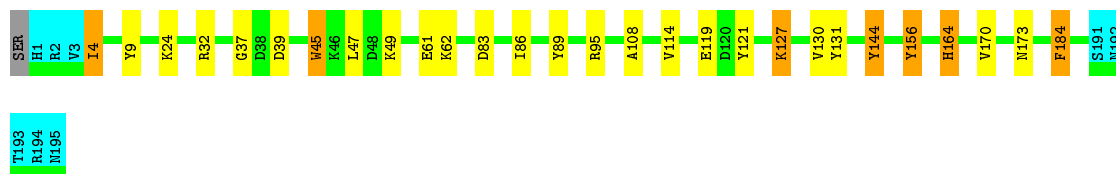
Chain B:  100%



#### 4.2.5 Score per residue for model 5

- Molecule 1: tSH2 domain of transcription elongation factor Spt6

Chain A:  81%  11% . . .





- Molecule 2: Tyrosine phosphorylated CTD

Chain B:  100%



#### 4.2.6 Score per residue for model 6

- Molecule 1: tSH2 domain of transcription elongation factor Spt6

Chain A:  83%  12% . . .





- Molecule 2: Tyrosine phosphorylated CTD

Chain B:  100%



#### 4.2.7 Score per residue for model 7

- Molecule 1: tSH2 domain of transcription elongation factor Spt6

Chain A:  81%  13% . . .



- Molecule 2: Tyrosine phosphorylated CTD

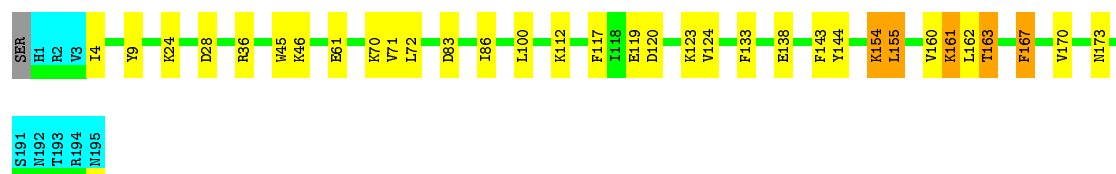
Chain B: 100%



#### 4.2.8 Score per residue for model 8

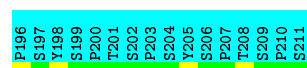
- Molecule 1: tSH2 domain of transcription elongation factor Spt6

Chain A: 79% 14% . . .



- Molecule 2: Tyrosine phosphorylated CTD

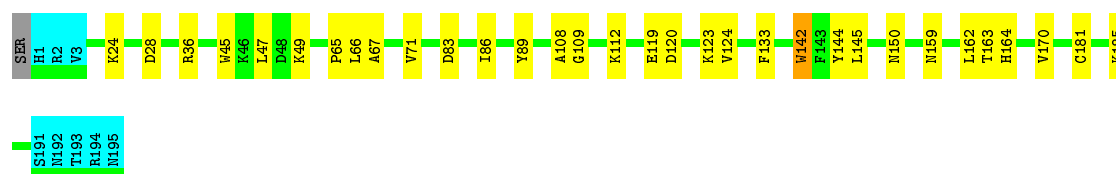
Chain B: 100%



#### 4.2.9 Score per residue for model 9 (medoid)

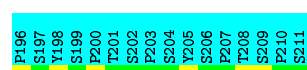
- Molecule 1: tSH2 domain of transcription elongation factor Spt6

Chain A: 79% 16% . . .



- Molecule 2: Tyrosine phosphorylated CTD

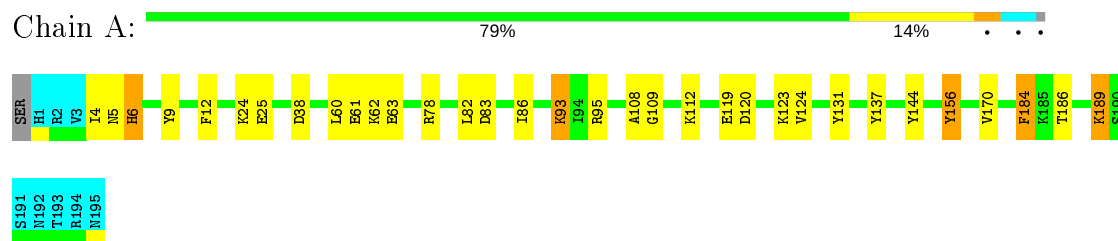
Chain B: 100%



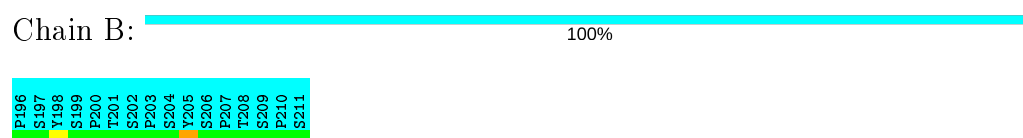


### 4.2.10 Score per residue for model 10

- Molecule 1: tSH2 domain of transcription elongation factor Spt6

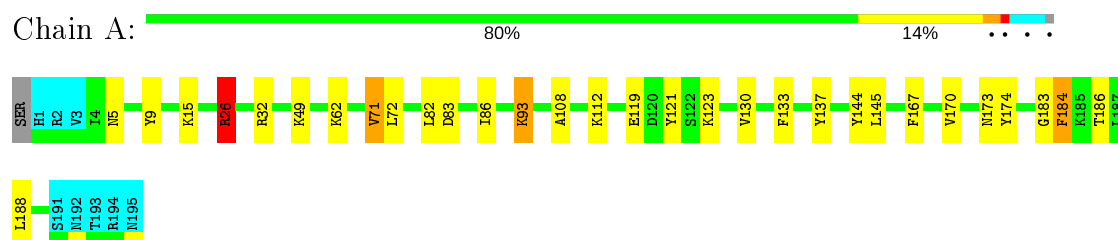


- Molecule 2: Tyrosine phosphorylated CTD

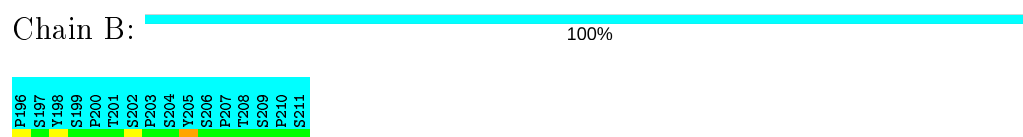


### 4.2.11 Score per residue for model 11

- Molecule 1: tSH2 domain of transcription elongation factor Spt6

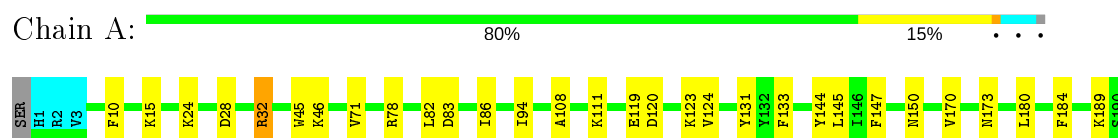


- Molecule 2: Tyrosine phosphorylated CTD



### 4.2.12 Score per residue for model 12

- Molecule 1: tSH2 domain of transcription elongation factor Spt6





- Molecule 2: Tyrosine phosphorylated CTD

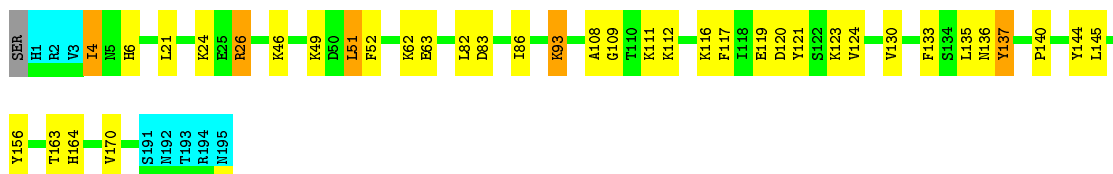
Chain B: 100%



#### 4.2.13 Score per residue for model 13

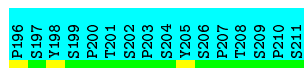
- Molecule 1: tSH2 domain of transcription elongation factor Spt6

Chain A: 76% 17% . . .



- Molecule 2: Tyrosine phosphorylated CTD

Chain B: 100%



#### 4.2.14 Score per residue for model 14

- Molecule 1: tSH2 domain of transcription elongation factor Spt6

Chain A: 81% 13% . . .



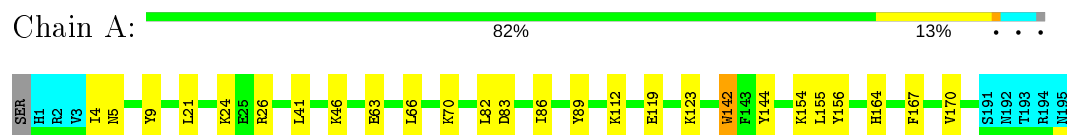
- Molecule 2: Tyrosine phosphorylated CTD

Chain B: 100%

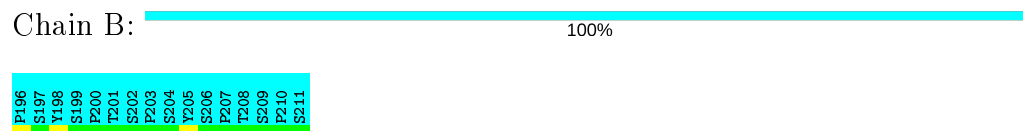


### 4.2.15 Score per residue for model 15

- Molecule 1: tSH2 domain of transcription elongation factor Spt6

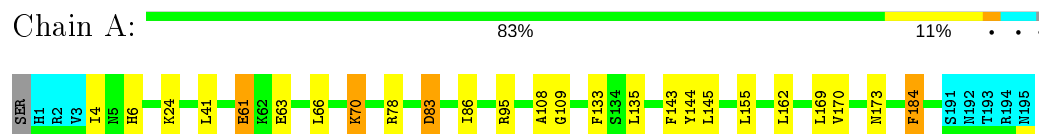


- Molecule 2: Tyrosine phosphorylated CTD

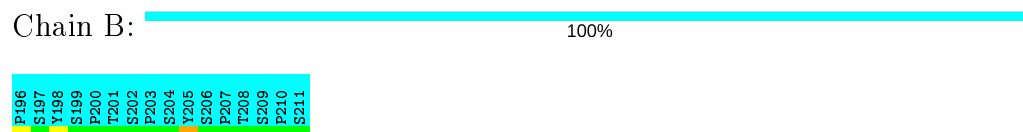


### 4.2.16 Score per residue for model 16

- Molecule 1: tSH2 domain of transcription elongation factor Spt6

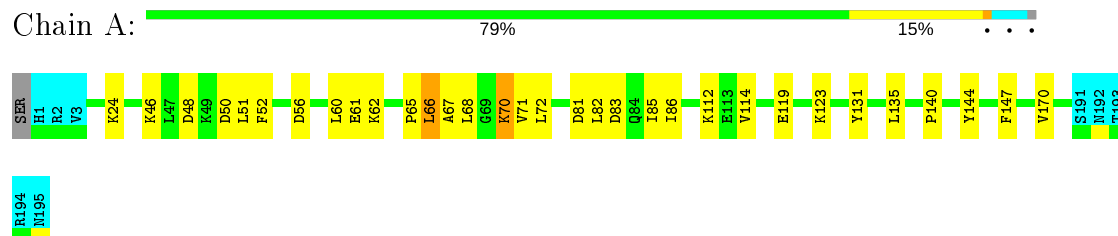


- Molecule 2: Tyrosine phosphorylated CTD



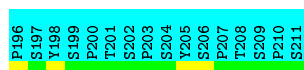
### 4.2.17 Score per residue for model 17

- Molecule 1: tSH2 domain of transcription elongation factor Spt6



- Molecule 2: Tyrosine phosphorylated CTD





#### 4.2.18 Score per residue for model 18

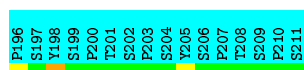
- Molecule 1: tSH2 domain of transcription elongation factor Spt6

Chain A: 81% 14% . . .



- Molecule 2: Tyrosine phosphorylated CTD

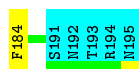
Chain B: 100%



#### 4.2.19 Score per residue for model 19

- Molecule 1: tSH2 domain of transcription elongation factor Spt6

Chain A: 78% 17% . . .



- Molecule 2: Tyrosine phosphorylated CTD

Chain B: 100%



#### 4.2.20 Score per residue for model 20

- Molecule 1: tSH2 domain of transcription elongation factor Spt6

Chain A: 82% 13% . . .



- Molecule 2: Tyrosine phosphorylated CTD

Chain B:

100%

P196	P197	P198	P199	P200	T201	S202	P203	S204	Y205	S206	P207	T208	S209	P210	S211

## 5 Refinement protocol and experimental data overview

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

Of the 500 calculated structures, 20 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
CYANA	structure calculation	3.97-3.98.5
Amber	refinement	16

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	1706
Number of shifts mapped to atoms	1706
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	59%

No validations of the models with respect to experimental NMR restraints is performed at this time.

## 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: PTR

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	0.57±0.00	0±0/1616 ( 0.0± 0.0%)	0.81±0.01	0±0/2185 ( 0.0± 0.0%)
All	All	0.57	0/32320 ( 0.0%)	0.81	3/43700 ( 0.0%)

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	0.1±0.2
All	All	0	1

There are no bond-length outliers.

All 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	9	TYR	CB-CG-CD2	-6.26	117.25	121.00	7	1
1	A	9	TYR	CB-CG-CD1	6.18	124.71	121.00	7	1
1	A	131	TYR	CB-CG-CD2	-5.49	117.71	121.00	20	1

There are no chirality outliers.

All unique planar outliers are listed below.

Mol	Chain	Res	Type	Group	Models (Total)
1	A	9	TYR	Sidechain	1

## 6.2 Too-close contacts ⓘ

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	1574	1567	1567	9±3
2	B	0	0	0	0±0
All	All	31480	31340	31340	189

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 3.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:10:PHE:CE1	1:A:17:ALA:HB2	0.63	2.28	3	1
1:A:10:PHE:CZ	1:A:17:ALA:HB2	0.56	2.35	3	1
1:A:160:VAL:HG22	1:A:161:LYS:H	0.55	1.61	8	1
1:A:160:VAL:HG22	1:A:161:LYS:N	0.55	2.16	8	1
1:A:189:LYS:HA	1:A:189:LYS:HE3	0.55	1.78	1	1
1:A:89:TYR:CD1	1:A:164:HIS:CD2	0.53	2.96	4	2
1:A:10:PHE:CD2	1:A:30:VAL:HB	0.52	2.38	3	1
1:A:6:HIS:HB2	1:A:9:TYR:CG	0.52	2.39	7	1
1:A:70:LYS:HA	1:A:70:LYS:HE2	0.52	1.82	16	2
1:A:50:ASP:C	1:A:51:LEU:HD22	0.52	2.25	17	1
1:A:131:TYR:CE1	1:A:147:PHE:CD2	0.52	2.98	12	2
1:A:83:ASP:HA	1:A:86:ILE:HG22	0.51	1.82	11	20
1:A:4:ILE:CG2	1:A:9:TYR:CG	0.51	2.94	5	2
1:A:133:PHE:CD2	1:A:145:LEU:HD22	0.51	2.41	11	4
1:A:93:LYS:HA	1:A:93:LYS:HE3	0.51	1.82	11	1
1:A:189:LYS:HE2	1:A:189:LYS:HA	0.50	1.84	10	2
1:A:120:ASP:O	1:A:124:VAL:HG22	0.50	2.07	12	9
1:A:21:LEU:HD13	1:A:46:LYS:HB2	0.49	1.84	13	4
1:A:6:HIS:CB	1:A:9:TYR:CZ	0.49	2.95	7	1
1:A:82:LEU:HD13	1:A:82:LEU:C	0.49	2.27	10	1
1:A:9:TYR:N	1:A:9:TYR:CD1	0.49	2.81	7	1
1:A:51:LEU:HD13	1:A:52:PHE:N	0.49	2.23	13	1
1:A:189:LYS:HA	1:A:189:LYS:CE	0.49	2.38	12	1
1:A:89:TYR:CE1	1:A:93:LYS:CE	0.49	2.96	3	1
1:A:6:HIS:HB2	1:A:9:TYR:CD2	0.48	2.43	7	1
1:A:93:LYS:HE3	1:A:93:LYS:HA	0.48	1.84	10	2

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:133:PHE:CD1	1:A:145:LEU:HD22	0.48	2.44	3	3
1:A:27:GLY:HA2	1:A:45:TRP:CZ2	0.48	2.44	7	1
1:A:181:CYS:SG	1:A:185:LYS:HE2	0.48	2.49	9	1
1:A:71:VAL:C	1:A:72:LEU:HD22	0.47	2.29	7	6
1:A:10:PHE:CZ	1:A:17:ALA:CB	0.47	2.97	3	1
1:A:60:LEU:HB3	1:A:71:VAL:HG23	0.47	1.85	17	1
1:A:189:LYS:HA	1:A:189:LYS:HE2	0.47	1.84	12	1
1:A:154:LYS:CD	1:A:155:LEU:H	0.47	2.22	3	4
1:A:131:TYR:CE1	1:A:147:PHE:CD1	0.47	3.03	17	1
1:A:4:ILE:CG2	1:A:6:HIS:H	0.46	2.23	16	3
1:A:189:LYS:CE	1:A:189:LYS:HA	0.46	2.40	3	1
1:A:89:TYR:CD1	1:A:164:HIS:CG	0.46	3.03	5	1
1:A:9:TYR:CD1	1:A:31:ILE:HD12	0.46	2.45	7	1
1:A:46:LYS:HE3	1:A:48:ASP:O	0.46	2.10	18	1
1:A:10:PHE:HB3	1:A:32:ARG:CB	0.46	2.40	12	1
1:A:10:PHE:CZ	1:A:32:ARG:HB3	0.46	2.45	3	1
1:A:4:ILE:HG22	1:A:9:TYR:CD2	0.46	2.46	4	3
1:A:82:LEU:C	1:A:82:LEU:HD13	0.46	2.30	13	2
1:A:42:ALA:CB	1:A:54:HIS:CE1	0.46	2.99	20	1
1:A:180:LEU:HD11	1:A:184:PHE:CZ	0.45	2.46	12	1
1:A:156:TYR:CD1	1:A:156:TYR:N	0.45	2.84	10	3
1:A:60:LEU:HB2	1:A:71:VAL:CG2	0.45	2.40	18	1
1:A:63:GLU:H	1:A:63:GLU:CD	0.45	2.15	1	1
1:A:45:TRP:CE3	1:A:93:LYS:HE3	0.45	2.46	6	1
1:A:156:TYR:N	1:A:156:TYR:CD1	0.45	2.85	13	2
1:A:65:PRO:C	1:A:67:ALA:H	0.45	2.14	9	3
1:A:70:LYS:HA	1:A:70:LYS:CE	0.45	2.41	15	1
1:A:100:LEU:HD11	1:A:143:PHE:CE1	0.44	2.47	8	1
1:A:60:LEU:HG	1:A:61:GLU:H	0.44	1.72	10	1
1:A:160:VAL:CG2	1:A:161:LYS:H	0.44	2.26	8	1
1:A:174:TYR:CE1	1:A:183:GLY:HA3	0.44	2.47	11	1
1:A:4:ILE:HG22	1:A:9:TYR:CG	0.44	2.47	5	1
1:A:160:VAL:CG2	1:A:161:LYS:N	0.44	2.81	8	1
1:A:46:LYS:HE3	1:A:48:ASP:C	0.44	2.32	17	1
1:A:167:PHE:CD1	1:A:167:PHE:C	0.44	2.91	14	1
1:A:127:LYS:HE3	1:A:127:LYS:HA	0.44	1.88	5	1
1:A:162:LEU:O	1:A:163:THR:HG22	0.44	2.13	8	1
1:A:142:TRP:CD1	1:A:142:TRP:C	0.44	2.91	9	1
1:A:121:TYR:CD1	1:A:130:VAL:HG13	0.43	2.47	13	2
1:A:70:LYS:HE2	1:A:70:LYS:HA	0.43	1.89	17	1
1:A:180:LEU:HD12	1:A:180:LEU:C	0.43	2.33	14	1

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:144:TYR:C	1:A:144:TYR:CD1	0.43	2.92	5	1
1:A:9:TYR:CE1	1:A:31:ILE:HD12	0.43	2.48	7	1
1:A:62:LYS:HA	1:A:62:LYS:HE2	0.43	1.89	19	1
1:A:4:ILE:HG23	1:A:6:HIS:H	0.43	1.73	13	1
1:A:34:SER:C	1:A:36:ARG:H	0.43	2.17	19	1
1:A:62:LYS:CG	1:A:67:ALA:HB3	0.43	2.44	1	1
1:A:28:ASP:O	1:A:45:TRP:HA	0.43	2.13	12	6
1:A:6:HIS:CD2	1:A:82:LEU:HD12	0.43	2.49	10	1
1:A:135:LEU:HG	1:A:143:PHE:CE1	0.43	2.49	16	1
1:A:6:HIS:CB	1:A:9:TYR:CE1	0.43	3.02	7	1
1:A:28:ASP:C	1:A:46:LYS:HE3	0.42	2.34	2	1
1:A:135:LEU:H	1:A:135:LEU:HD23	0.42	1.74	17	1
1:A:52:PHE:H	1:A:140:PRO:HG2	0.42	1.73	17	1
1:A:4:ILE:HG22	1:A:9:TYR:CD1	0.42	2.49	19	1
1:A:149:LEU:HD22	1:A:149:LEU:N	0.42	2.29	7	1
1:A:44:THR:HG22	1:A:54:HIS:ND1	0.42	2.29	14	1
1:A:79:TYR:CD2	1:A:85:ILE:HG22	0.42	2.48	20	1
1:A:9:TYR:CE2	1:A:82:LEU:CB	0.42	3.02	7	1
1:A:154:LYS:HG3	1:A:155:LEU:H	0.42	1.74	8	1
1:A:29:PHE:N	1:A:46:LYS:HE3	0.42	2.29	2	1
1:A:26:ARG:CD	1:A:26:ARG:H	0.42	2.27	11	1
1:A:142:TRP:C	1:A:142:TRP:CD1	0.42	2.93	15	1
1:A:121:TYR:CD2	1:A:130:VAL:HG13	0.42	2.50	14	4
1:A:66:LEU:HD12	1:A:66:LEU:H	0.42	1.74	16	1
1:A:156:TYR:N	1:A:156:TYR:CD2	0.42	2.87	15	1
1:A:45:TRP:CZ2	1:A:47:LEU:CD2	0.41	3.03	1	3
1:A:107:LYS:CB	1:A:114:VAL:HG23	0.41	2.45	3	1
1:A:111:LYS:HE2	1:A:137:TYR:CD1	0.41	2.50	13	1
1:A:89:TYR:HA	1:A:164:HIS:CB	0.41	2.45	14	1
1:A:133:PHE:CD2	1:A:145:LEU:HD12	0.41	2.49	12	1
1:A:135:LEU:H	1:A:135:LEU:CD2	0.41	2.27	17	1
1:A:135:LEU:N	1:A:135:LEU:HD23	0.41	2.31	17	1
1:A:66:LEU:N	1:A:66:LEU:HD12	0.41	2.30	15	2
1:A:6:HIS:HB2	1:A:9:TYR:CE2	0.41	2.49	7	1
1:A:10:PHE:CB	1:A:32:ARG:CB	0.41	2.99	12	1
1:A:161:LYS:HG2	1:A:167:PHE:CD2	0.41	2.50	8	1
1:A:120:ASP:O	1:A:124:VAL:HG23	0.41	2.16	13	1
1:A:22:ARG:CD	1:A:50:ASP:HA	0.41	2.46	20	1
1:A:135:LEU:HD22	1:A:135:LEU:N	0.41	2.30	13	1
1:A:62:LYS:HE2	1:A:62:LYS:HA	0.41	1.93	1	1
1:A:154:LYS:CG	1:A:155:LEU:H	0.41	2.29	8	1

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:89:TYR:CE2	1:A:164:HIS:ND1	0.41	2.89	9	1
1:A:45:TRP:CZ2	1:A:94:ILE:HG12	0.41	2.51	12	1
1:A:39:ASP:CB	1:A:62:LYS:HE3	0.41	2.46	5	1
1:A:40:HIS:CE1	1:A:58:GLN:HE21	0.40	2.35	4	1
1:A:4:ILE:CG2	1:A:9:TYR:HB2	0.40	2.47	10	1
1:A:61:GLU:CD	1:A:61:GLU:C	0.40	2.80	16	1
1:A:144:TYR:CD1	1:A:144:TYR:O	0.40	2.74	18	1
1:A:133:PHE:CD1	1:A:133:PHE:N	0.40	2.89	1	1
1:A:107:LYS:HE3	1:A:117:PHE:C	0.40	2.37	4	1
1:A:134:SER:HB3	1:A:144:TYR:CE2	0.40	2.51	4	1

## 6.3 Torsion angles

### 6.3.1 Protein backbone

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	187/196 (95%)	172±2 (92±1%)	11±2 (6±1%)	3±1 (2±1%)	13	56
2	B	0	-	-	-	-	-
All	All	3740/4240 (88%)	3448 (92%)	227 (6%)	65 (2%)	13	56

All 21 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	170	VAL	19
1	A	108	ALA	11
1	A	109	GLY	5
1	A	26	ARG	4
1	A	154	LYS	3
1	A	4	ILE	2
1	A	5	ASN	2
1	A	61	GLU	2
1	A	62	LYS	2
1	A	24	LYS	2
1	A	66	LEU	2

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type	Models (Total)
1	A	63	GLU	2
1	A	68	LEU	1
1	A	161	LYS	1
1	A	37	GLY	1
1	A	163	THR	1
1	A	140	PRO	1
1	A	25	GLU	1
1	A	171	ASN	1
1	A	137	TYR	1
1	A	35	SER	1

### 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	173/182 (95%)	159±3 (92±2%)	14±3 (8±2%)	16 63
2	B	0	-	-	-
All	All	3460/3920 (88%)	3182 (92%)	278 (8%)	16 63

All 69 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	144	TYR	20
1	A	119	GLU	19
1	A	123	LYS	18
1	A	24	LYS	16
1	A	173	ASN	12
1	A	82	LEU	10
1	A	112	LYS	10
1	A	114	VAL	7
1	A	184	PHE	7
1	A	61	GLU	6
1	A	164	HIS	6
1	A	155	LEU	6
1	A	49	LYS	6
1	A	71	VAL	5

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Models (Total)
1	A	95	ARG	5
1	A	62	LYS	5
1	A	167	PHE	5
1	A	78	ARG	4
1	A	25	GLU	4
1	A	133	PHE	4
1	A	41	LEU	4
1	A	32	ARG	4
1	A	70	LYS	4
1	A	9	TYR	4
1	A	83	ASP	3
1	A	26	ARG	3
1	A	93	LYS	3
1	A	15	LYS	3
1	A	189	LYS	3
1	A	186	THR	3
1	A	131	TYR	3
1	A	45	TRP	3
1	A	150	ASN	3
1	A	116	LYS	3
1	A	162	LEU	3
1	A	46	LYS	3
1	A	148	LYS	2
1	A	163	THR	2
1	A	105	LYS	2
1	A	89	TYR	2
1	A	136	ASN	2
1	A	142	TRP	2
1	A	36	ARG	2
1	A	117	PHE	2
1	A	64	ASN	2
1	A	169	LEU	2
1	A	180	LEU	2
1	A	137	TYR	2
1	A	63	GLU	2
1	A	111	LYS	2
1	A	18	GLU	2
1	A	53	GLN	2
1	A	156	TYR	2
1	A	127	LYS	2
1	A	96	LEU	1
1	A	154	LYS	1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Models (Total)
1	A	138	GLU	1
1	A	4	ILE	1
1	A	58	GLN	1
1	A	12	PHE	1
1	A	6	HIS	1
1	A	159	ASN	1
1	A	56	ASP	1
1	A	171	ASN	1
1	A	161	LYS	1
1	A	81	ASP	1
1	A	5	ASN	1
1	A	51	LEU	1
1	A	85	ILE	1

### 6.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 6.4 Non-standard residues in protein, DNA, RNA chains ⓘ

2 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds for which Mogul statistics could be retrieved, the number of bonds that are observed in the model and the number of bonds that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length is the number of standard deviations the observed value is removed from the expected value. A bond length with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond lengths.

Mol	Type	Chain	Res	Link	Bond lengths		
					Counts	RMSZ	#Z>2
2	PTR	B	205	2	15,16,17	1.40±0.00	0±0 (0±0%)
2	PTR	B	198	2	15,16,17	1.34±0.00	0±0 (0±0%)

In the following table, the Counts columns list the number of angles for which Mogul statistics could be retrieved, the number of angles that are observed in the model and the number of angles that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond angle is the number of standard deviations the observed value is removed from the expected value. A bond angle with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond angles.

Mol	Type	Chain	Res	Link	Bond angles		
					Counts	RMSZ	#Z>2
2	PTR	B	205	2	19,22,24	0.88±0.00	0±0 (0±0%)
2	PTR	B	198	2	19,22,24	0.92±0.00	0±0 (0±0%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	PTR	B	205	2	-	0±0,10,11,13	0±0,1,1,1
2	PTR	B	198	2	-	0±0,10,11,13	0±0,1,1,1

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

The completeness of assignment taking into account all chemical shift lists is 59% for the well-defined parts and 55% for the entire structure.

### 7.1 Chemical shift list 1

File name: working\_cs.cif

Chemical shift list name: *spt6.STAR*

#### 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	1706
Number of shifts mapped to atoms	1706
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	2

#### 7.1.2 Chemical shift referencing

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction $\pm$ precision, ppm	Suggested action
$^{13}\text{C}_\alpha$	152	$-0.24 \pm 0.14$	None needed ( $< 0.5$ ppm)
$^{13}\text{C}_\beta$	145	$-0.34 \pm 0.06$	None needed ( $< 0.5$ ppm)
$^{13}\text{C}'$	0	—	None (insufficient data)
$^{15}\text{N}$	156	$-0.36 \pm 0.29$	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 59%, i.e. 1464 atoms were assigned a chemical shift out of a possible 2479. 30 out of 35 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	$^1\text{H}$	$^{13}\text{C}$	$^{15}\text{N}$
Backbone	600/923 (65%)	300/368 (82%)	148/374 (40%)	152/181 (84%)
Sidechain	803/1273 (63%)	487/747 (65%)	316/466 (68%)	0/60 (0%)

*Continued on next page...*



Continued from previous page...

	<b>Total</b>	<b><sup>1</sup>H</b>	<b><sup>13</sup>C</b>	<b><sup>15</sup>N</b>
Aromatic	61/283 (22%)	39/150 (26%)	22/125 (18%)	0/8 (0%)
Overall	1464/2479 (59%)	826/1265 (65%)	486/965 (50%)	152/249 (61%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 55%, i.e. 1497 atoms were assigned a chemical shift out of a possible 2726. 30 out of 36 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	<b>Total</b>	<b><sup>1</sup>H</b>	<b><sup>13</sup>C</b>	<b><sup>15</sup>N</b>
Backbone	616/1023 (60%)	308/407 (76%)	152/418 (36%)	156/198 (79%)
Sidechain	820/1413 (58%)	498/834 (60%)	322/511 (63%)	0/68 (0%)
Aromatic	61/290 (21%)	39/154 (25%)	22/127 (17%)	0/9 (0%)
Overall	1497/2726 (55%)	845/1395 (61%)	496/1056 (47%)	156/275 (57%)

#### 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 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	107	LYS	HE3	1.48	3.86 – 1.96	-7.5
1	A	107	LYS	HE2	1.59	3.87 – 1.97	-7.0

#### 7.1.5 Random Coil Index (RCI) plots ⓘ

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:

