



# Full wwPDB NMR Structure Validation Report ⓘ

Feb 17, 2018 – 11:21 pm GMT

PDB ID : 2KBZ  
Title : NMR structure of protein gp15 of bacteriophage SPP1  
Authors : Gallopin, M.; Gilquin, B.; Zinn-Justin, S.  
Deposited on : 2008-12-12

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:

Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Percentile statistics	:	20171227.v01 (using entries in the PDB archive December 27th 2017)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	trunk30686
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk30686

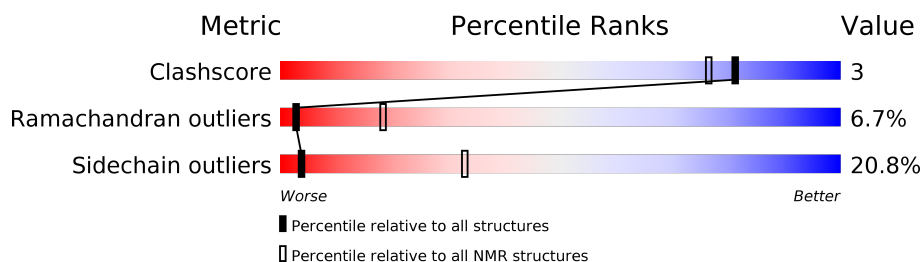
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*SOLUTION NMR*

The overall completeness of chemical shifts assignment is 84%.

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	136279	12091
Ramachandran outliers	132675	10835
Sidechain outliers	132484	10811

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	99	

## 2 Ensemble composition and analysis

This entry contains 20 models. Model 2 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:42-A:102, A:116-A:133 (79)	0.75	2

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

### 3 Entry composition [i](#)

There is only 1 type of molecule in this entry. The entry contains 1586 atoms, of which 797 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called 15 protein (Bacteriophage SPP1 complete nucleotide sequence).

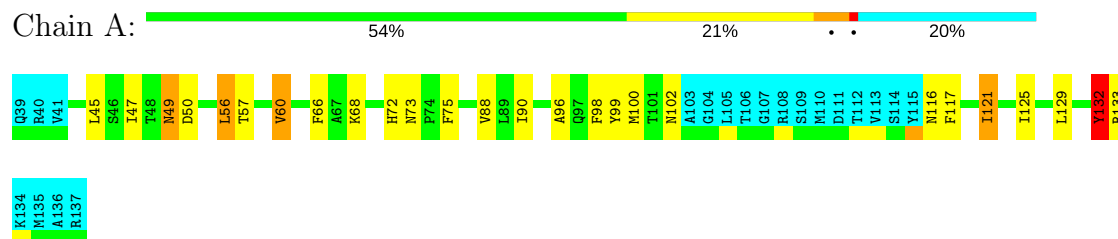
Mol	Chain	Residues	Atoms						Trace
1	A	99	Total	C	H	N	O	S	0
			1586	502	797	134	148	5	

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

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

These plots are provided for all protein, RNA and DNA 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: 15 protein (Bacteriophage SPP1 complete nucleotide sequence)

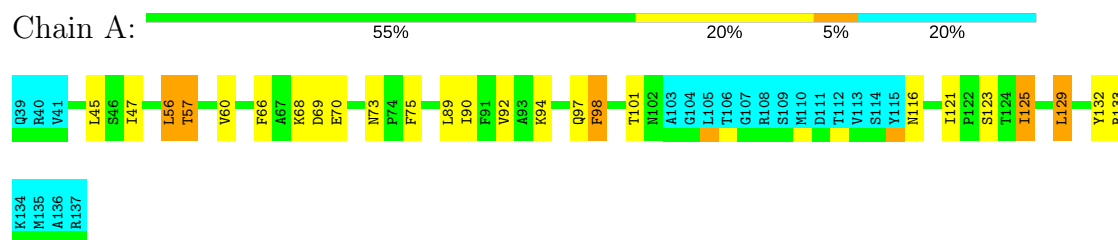


### 4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

#### 4.2.1 Score per residue for model 1

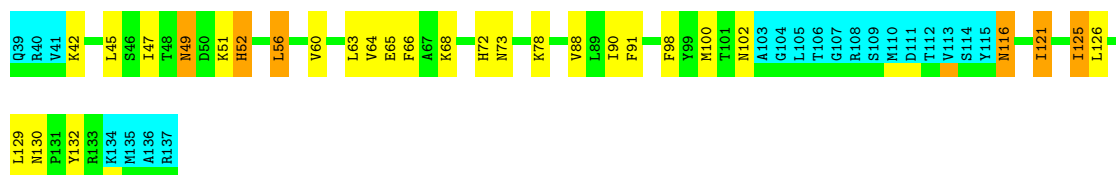
- Molecule 1: 15 protein (Bacteriophage SPP1 complete nucleotide sequence)



#### 4.2.2 Score per residue for model 2 (medoid)

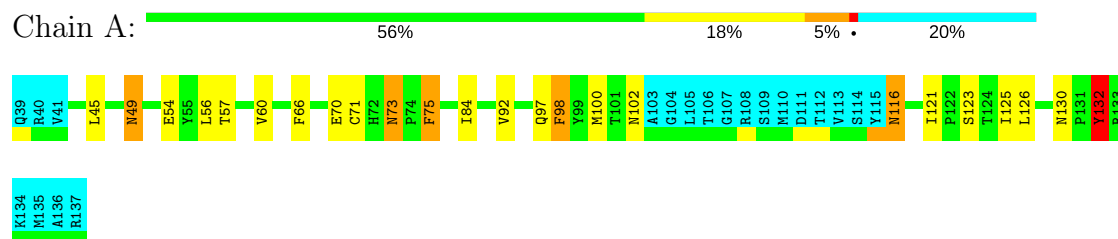
- Molecule 1: 15 protein (Bacteriophage SPP1 complete nucleotide sequence)





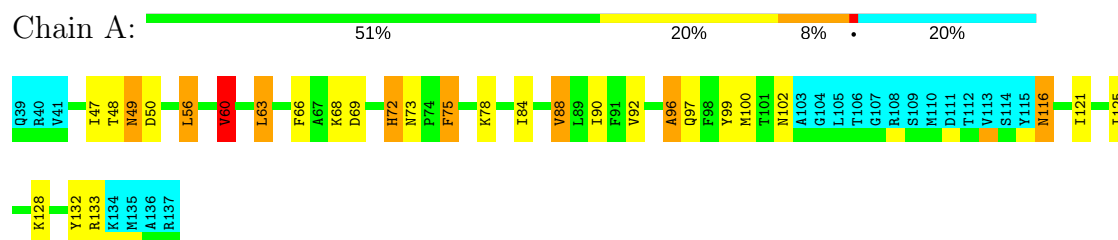
#### 4.2.3 Score per residue for model 3

- Molecule 1: 15 protein (Bacteriophage SPP1 complete nucleotide sequence)



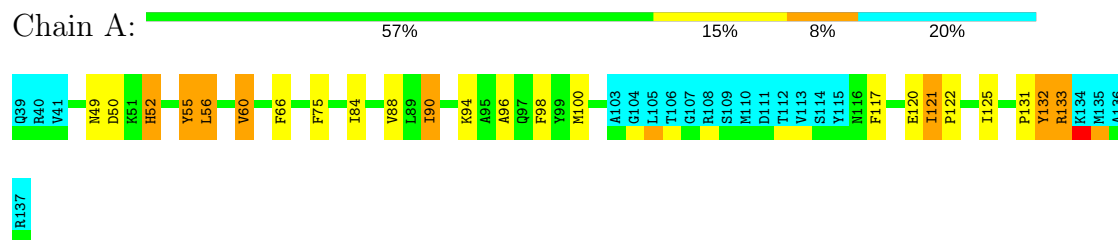
#### 4.2.4 Score per residue for model 4

- Molecule 1: 15 protein (Bacteriophage SPP1 complete nucleotide sequence)



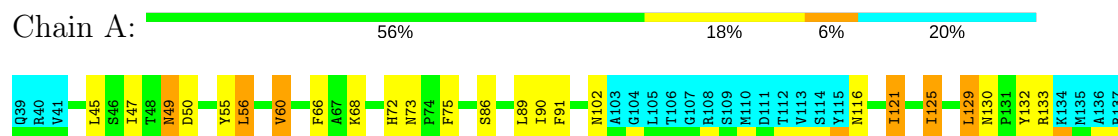
#### 4.2.5 Score per residue for model 5

- Molecule 1: 15 protein (Bacteriophage SPP1 complete nucleotide sequence)



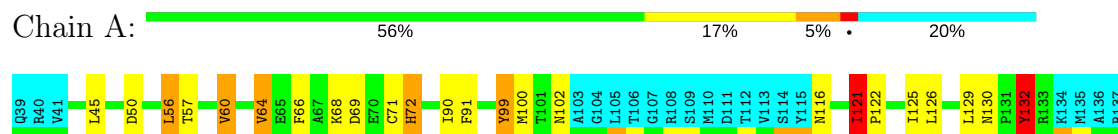
#### 4.2.6 Score per residue for model 6

- Molecule 1: 15 protein (Bacteriophage SPP1 complete nucleotide sequence)



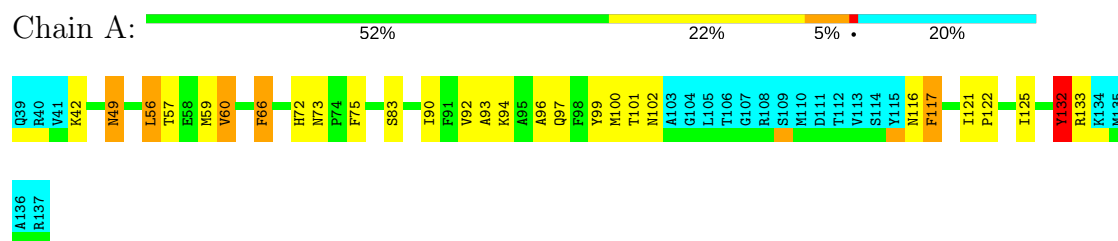
#### 4.2.7 Score per residue for model 7

- Molecule 1: 15 protein (Bacteriophage SPP1 complete nucleotide sequence)



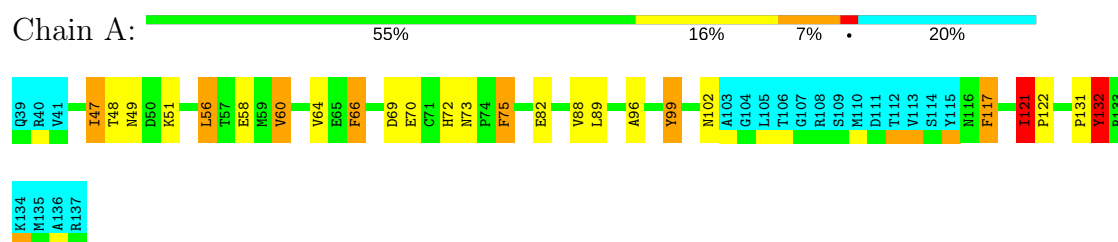
#### 4.2.8 Score per residue for model 8

- Molecule 1: 15 protein (Bacteriophage SPP1 complete nucleotide sequence)



#### 4.2.9 Score per residue for model 9

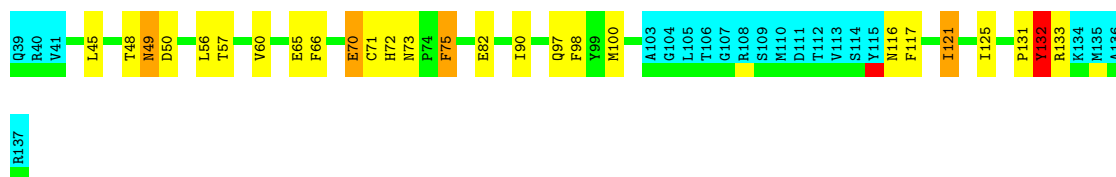
- Molecule 1: 15 protein (Bacteriophage SPP1 complete nucleotide sequence)



#### 4.2.10 Score per residue for model 10

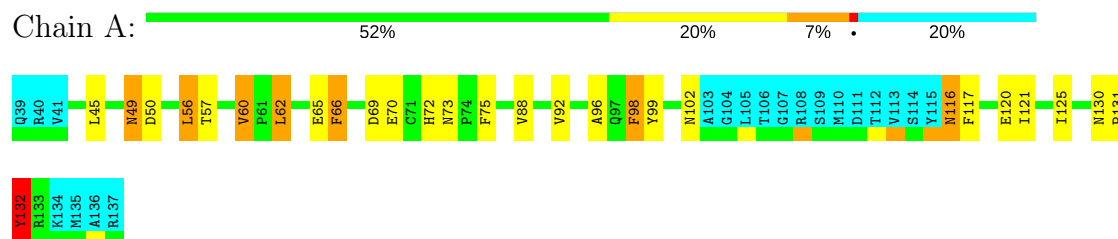
- Molecule 1: 15 protein (Bacteriophage SPP1 complete nucleotide sequence)





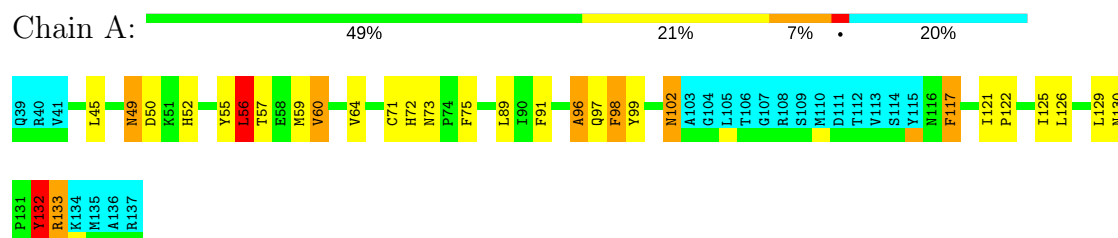
#### 4.2.11 Score per residue for model 11

- Molecule 1: 15 protein (Bacteriophage SPP1 complete nucleotide sequence)



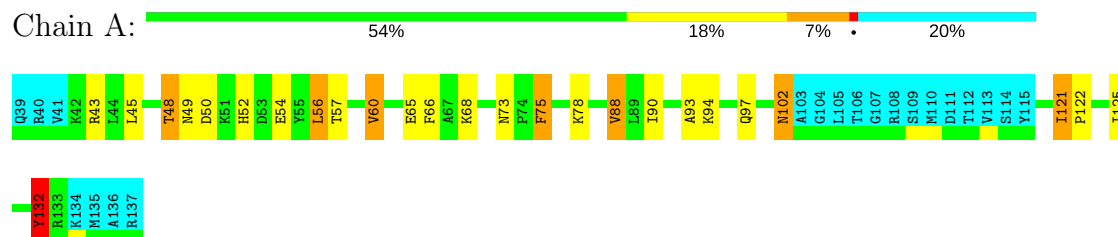
#### 4.2.12 Score per residue for model 12

- Molecule 1: 15 protein (Bacteriophage SPP1 complete nucleotide sequence)



#### 4.2.13 Score per residue for model 13

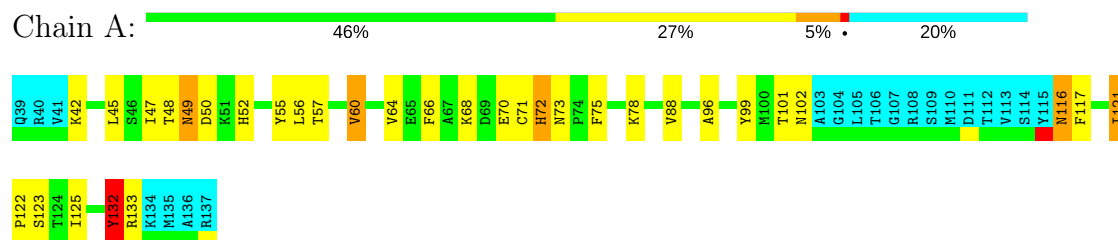
- Molecule 1: 15 protein (Bacteriophage SPP1 complete nucleotide sequence)



#### 4.2.14 Score per residue for model 14

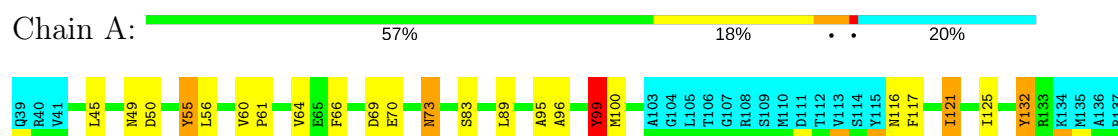
- Molecule 1: 15 protein (Bacteriophage SPP1 complete nucleotide sequence)





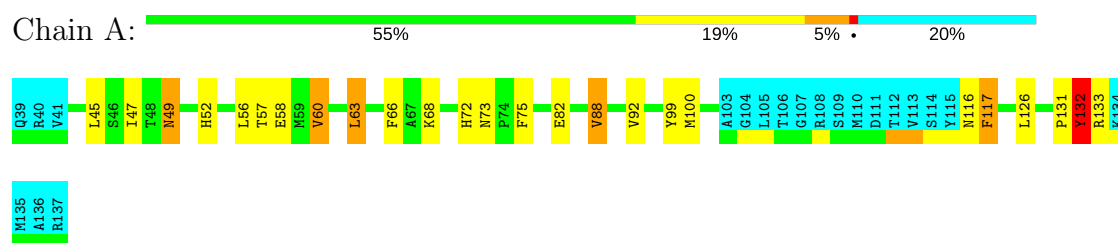
#### 4.2.15 Score per residue for model 15

- Molecule 1: 15 protein (Bacteriophage SPP1 complete nucleotide sequence)



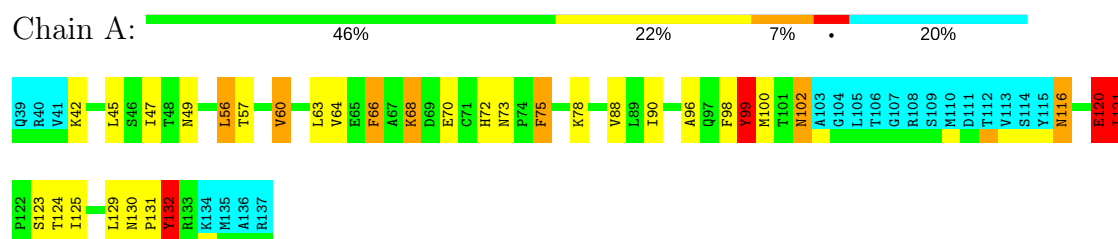
#### 4.2.16 Score per residue for model 16

- Molecule 1: 15 protein (Bacteriophage SPP1 complete nucleotide sequence)



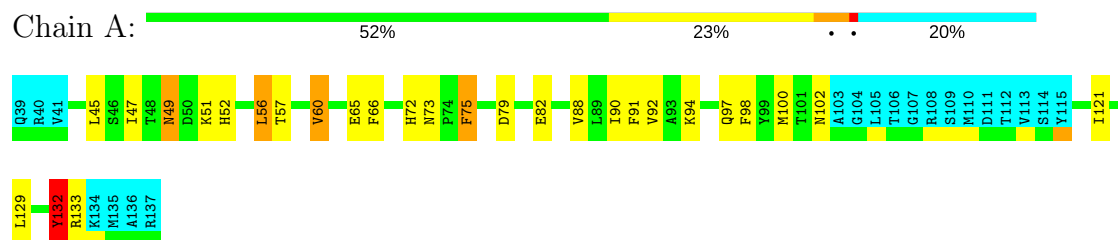
#### 4.2.17 Score per residue for model 17

- Molecule 1: 15 protein (Bacteriophage SPP1 complete nucleotide sequence)



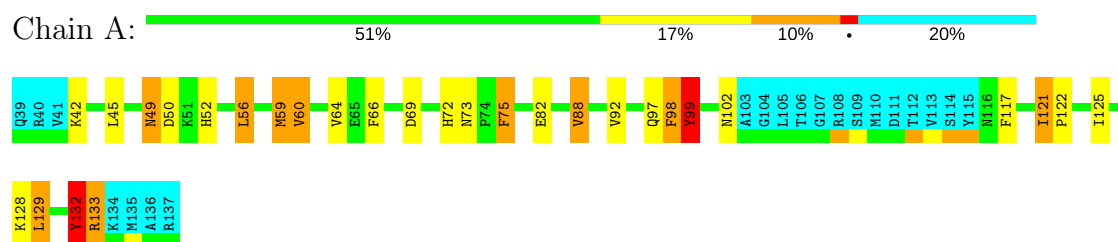
#### 4.2.18 Score per residue for model 18

- Molecule 1: 15 protein (Bacteriophage SPP1 complete nucleotide sequence)



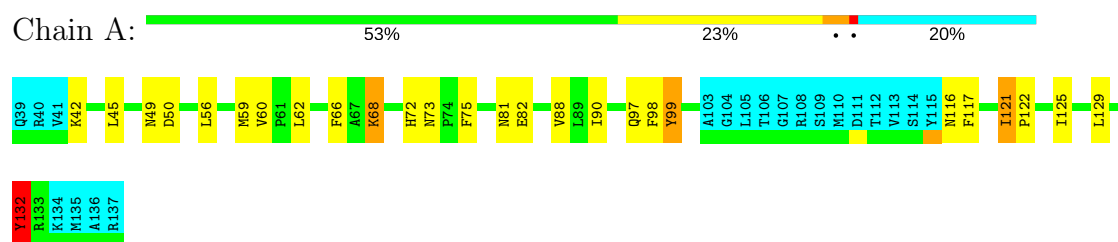
#### 4.2.19 Score per residue for model 19

- Molecule 1: 15 protein (Bacteriophage SPP1 complete nucleotide sequence)



#### 4.2.20 Score per residue for model 20

- Molecule 1: 15 protein (Bacteriophage SPP1 complete nucleotide sequence)



## 5 Refinement protocol and experimental data overview

The models were refined using the following method: *molecular dynamics*.

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
INCA	refinement	1.0

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)	BMRB entry 16057
Number of chemical shift lists	1
Total number of shifts	1141
Number of shifts mapped to atoms	1141
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	84%

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

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	0.98±0.02	0±0/651 (0.0±0.0%)	1.64±0.04	10±3/883 (1.1±0.3%)
All	All	0.98	0/13020 (0.0%)	1.64	192/17660 (1.1%)

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	2.2±0.9
All	All	0	44

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	98	PHE	CB-CG-CD1	9.90	127.73	120.80	3	7
1	A	132	TYR	CB-CG-CD1	9.46	126.67	121.00	20	9
1	A	99	TYR	CB-CG-CD1	-8.78	115.73	121.00	14	5
1	A	60	VAL	CA-C-N	8.53	140.97	117.10	18	18
1	A	98	PHE	CB-CG-CD2	-8.17	115.08	120.80	11	8
1	A	59	MET	CG-SD-CE	-7.84	87.66	100.20	19	1
1	A	91	PHE	CB-CG-CD1	7.78	126.24	120.80	2	2
1	A	129	LEU	CB-CG-CD2	7.75	124.18	111.00	1	2
1	A	132	TYR	CB-CG-CD2	-7.48	116.51	121.00	10	8
1	A	99	TYR	CB-CG-CD2	-7.29	116.63	121.00	17	3
1	A	133	ARG	C-N-CA	7.29	139.92	121.70	5	1
1	A	100	MET	CG-SD-CE	-7.02	88.97	100.20	15	2
1	A	73	ASN	N-CA-CB	-6.95	98.10	110.60	8	2
1	A	75	PHE	CB-CG-CD1	6.90	125.63	120.80	3	1
1	A	117	PHE	CB-CG-CD2	-6.86	116.00	120.80	8	2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)	Models	
								Worst	Total
1	A	116	ASN	C-N-CA	6.84	138.81	121.70	20	1
1	A	55	TYR	CB-CG-CD1	6.81	125.08	121.00	5	4
1	A	125	ILE	CA-CB-CG1	6.75	123.83	111.00	6	2
1	A	96	ALA	CB-CA-C	6.74	120.22	110.10	15	9
1	A	129	LEU	CA-CB-CG	6.59	130.47	115.30	1	1
1	A	132	TYR	CA-CB-CG	6.53	125.80	113.40	19	1
1	A	65	GLU	CB-CA-C	6.47	123.33	110.40	2	4
1	A	91	PHE	CB-CG-CD2	-6.46	116.28	120.80	2	2
1	A	121	ILE	CA-C-N	6.40	135.02	117.10	6	1
1	A	95	ALA	CB-CA-C	-6.39	100.51	110.10	15	1
1	A	56	LEU	CB-CG-CD2	6.26	121.64	111.00	5	7
1	A	121	ILE	CB-CA-C	6.25	124.11	111.60	6	1
1	A	60	VAL	CA-C-O	-6.18	107.13	120.10	18	11
1	A	66	PHE	CB-CG-CD1	6.15	125.11	120.80	9	4
1	A	69	ASP	C-N-CA	6.15	137.07	121.70	1	7
1	A	72	HIS	N-CA-CB	5.99	121.39	110.60	14	2
1	A	100	MET	C-N-CA	5.95	136.57	121.70	16	2
1	A	102	ASN	N-CA-CB	5.88	121.19	110.60	12	2
1	A	130	ASN	CA-C-N	5.88	133.56	117.10	2	6
1	A	49	ASN	N-CA-CB	5.84	121.11	110.60	4	11
1	A	75	PHE	CB-CG-CD2	-5.80	116.74	120.80	3	2
1	A	99	TYR	N-CA-CB	5.73	120.92	110.60	15	2
1	A	120	GLU	C-N-CA	5.71	135.97	121.70	17	3
1	A	56	LEU	CB-CA-C	5.64	120.93	110.20	12	2
1	A	70	GLU	N-CA-CB	-5.59	100.54	110.60	9	2
1	A	60	VAL	O-C-N	-5.52	110.61	121.10	12	1
1	A	70	GLU	CA-CB-CG	5.50	125.51	113.40	10	2
1	A	79	ASP	N-CA-CB	5.47	120.46	110.60	18	1
1	A	55	TYR	CB-CG-CD2	-5.45	117.73	121.00	5	2
1	A	48	THR	C-N-CA	5.39	135.18	121.70	13	1
1	A	96	ALA	N-CA-CB	-5.38	102.56	110.10	12	1
1	A	57	THR	CA-CB-CG2	-5.33	104.94	112.40	10	1
1	A	60	VAL	CG1-CB-CG2	5.30	119.38	110.90	12	1
1	A	90	ILE	N-CA-CB	5.29	122.96	110.80	5	1
1	A	56	LEU	CA-CB-CG	5.24	127.36	115.30	6	4
1	A	117	PHE	CB-CG-CD1	5.23	124.46	120.80	8	1
1	A	131	PRO	C-N-CA	5.23	134.77	121.70	17	4
1	A	88	VAL	CB-CA-C	5.21	121.31	111.40	13	1
1	A	132	TYR	C-N-CA	5.21	134.72	121.70	19	1
1	A	68	LYS	CB-CA-C	5.21	120.81	110.40	1	1
1	A	73	ASN	CB-CA-C	5.20	120.80	110.40	3	1

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)	Models	
								Worst	Total
1	A	101	THR	C-N-CA	5.19	134.68	121.70	14	1
1	A	62	LEU	N-CA-CB	5.14	120.67	110.40	11	1
1	A	90	ILE	CA-CB-CG1	5.12	120.73	111.00	5	1
1	A	121	ILE	CA-CB-CG1	5.08	120.66	111.00	20	1
1	A	72	HIS	CA-CB-CG	5.06	122.21	113.60	7	1
1	A	99	TYR	N-CA-C	5.02	124.55	111.00	9	1
1	A	57	THR	CA-C-N	5.02	128.24	117.20	16	1

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	75	PHE	Sidechain	17
1	A	132	TYR	Sidechain	15
1	A	99	TYR	Sidechain	8
1	A	117	PHE	Sidechain	2
1	A	55	TYR	Sidechain	2

## 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	636	637	637	4±2
All	All	12720	12740	12740	72

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:121:ILE:HG22	1:A:122:PRO:HD2	0.67	1.65	19	1
1:A:121:ILE:HG22	1:A:122:PRO:CD	0.63	2.22	19	1
1:A:88:VAL:O	1:A:92:VAL:HG13	0.59	1.97	11	2
1:A:88:VAL:O	1:A:92:VAL:HG23	0.53	2.02	4	1

Continued on next page...

*Continued from previous page...*

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:57:THR:HG22	1:A:92:VAL:HG12	0.53	1.81	18	1
1:A:56:LEU:HD13	1:A:57:THR:H	0.52	1.63	1	1
1:A:63:LEU:HD22	1:A:121:ILE:HD12	0.51	1.82	17	1
1:A:56:LEU:HD21	1:A:93:ALA:HA	0.51	1.82	13	2
1:A:47:ILE:HG23	1:A:48:THR:H	0.51	1.65	9	1
1:A:132:TYR:CG	1:A:133:ARG:N	0.51	2.78	14	2
1:A:60:VAL:CG1	1:A:92:VAL:HG12	0.50	2.36	11	2
1:A:60:VAL:HG13	1:A:63:LEU:HD23	0.50	1.81	4	1
1:A:88:VAL:HG11	1:A:132:TYR:CG	0.50	2.42	19	1
1:A:59:MET:HE3	1:A:99:TYR:CE2	0.50	2.42	20	1
1:A:121:ILE:H	1:A:121:ILE:HD13	0.49	1.67	9	1
1:A:121:ILE:HB	1:A:122:PRO:HD2	0.49	1.83	20	6
1:A:56:LEU:HB2	1:A:96:ALA:HB1	0.49	1.83	12	2
1:A:89:LEU:HA	1:A:92:VAL:HG22	0.49	1.83	1	1
1:A:125:ILE:HD12	1:A:125:ILE:H	0.48	1.68	2	1
1:A:60:VAL:HG11	1:A:92:VAL:HA	0.48	1.84	16	1
1:A:56:LEU:HD23	1:A:57:THR:HG23	0.48	1.86	17	2
1:A:60:VAL:HG13	1:A:61:PRO:HD3	0.48	1.85	15	1
1:A:64:VAL:HG12	1:A:91:PHE:CE2	0.48	2.43	7	1
1:A:60:VAL:HG11	1:A:92:VAL:HG12	0.48	1.86	11	1
1:A:68:LYS:HG2	1:A:132:TYR:CE2	0.47	2.45	20	1
1:A:56:LEU:O	1:A:60:VAL:HG12	0.47	2.10	15	1
1:A:60:VAL:HB	1:A:92:VAL:HG12	0.46	1.86	8	1
1:A:63:LEU:HD22	1:A:117:PHE:CB	0.46	2.41	16	1
1:A:68:LYS:HD3	1:A:132:TYR:CE2	0.46	2.46	7	1
1:A:57:THR:HG22	1:A:92:VAL:HG23	0.45	1.88	3	2
1:A:62:LEU:HD12	1:A:117:PHE:CE2	0.45	2.46	11	1
1:A:60:VAL:HG21	1:A:92:VAL:HA	0.45	1.89	11	1
1:A:99:TYR:CE1	1:A:121:ILE:HG23	0.44	2.47	19	1
1:A:52:HIS:CE1	1:A:55:TYR:CD2	0.44	3.05	5	1
1:A:122:PRO:HG2	1:A:125:ILE:HD11	0.44	1.89	8	2
1:A:60:VAL:HG21	1:A:92:VAL:O	0.44	2.12	8	1
1:A:99:TYR:CD1	1:A:121:ILE:HG23	0.44	2.47	19	1
1:A:99:TYR:CE1	1:A:120:GLU:HB2	0.44	2.47	17	1
1:A:117:PHE:N	1:A:117:PHE:CD1	0.43	2.85	9	1
1:A:63:LEU:HD22	1:A:121:ILE:HB	0.43	1.89	2	1
1:A:59:MET:HB3	1:A:117:PHE:CZ	0.43	2.49	8	1
1:A:64:VAL:HG21	1:A:88:VAL:HG13	0.43	1.91	2	1
1:A:121:ILE:CD1	1:A:121:ILE:H	0.42	2.26	9	1
1:A:64:VAL:O	1:A:68:LYS:HB2	0.42	2.15	17	1
1:A:88:VAL:HG21	1:A:132:TYR:HA	0.42	1.91	19	1

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:68:LYS:HG2	1:A:132:TYR:CD2	0.42	2.50	13	1
1:A:132:TYR:CD2	1:A:133:ARG:N	0.42	2.88	12	1
1:A:99:TYR:CE1	1:A:121:ILE:CG2	0.41	3.03	15	1
1:A:91:PHE:CD2	1:A:129:LEU:HB2	0.41	2.50	6	1
1:A:60:VAL:CG1	1:A:92:VAL:HG22	0.41	2.45	18	1
1:A:63:LEU:HD22	1:A:117:PHE:HB3	0.41	1.92	16	1
1:A:94:LYS:HA	1:A:97:GLN:HG2	0.41	1.91	13	3
1:A:129:LEU:O	1:A:132:TYR:CD2	0.41	2.73	17	1
1:A:47:ILE:HG22	1:A:52:HIS:HB3	0.41	1.93	14	1
1:A:52:HIS:CE1	1:A:100:MET:SD	0.41	3.13	2	1
1:A:91:PHE:CE2	1:A:129:LEU:HD23	0.40	2.51	12	1
1:A:60:VAL:HG12	1:A:64:VAL:HG23	0.40	1.93	14	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	79/99 (80%)	67±2 (85±2%)	7±2 (8±2%)	5±2 (7±2%)	3	18
All	All	1580/1980 (80%)	1342 (85%)	132 (8%)	106 (7%)	3	18

All 20 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	49	ASN	18
1	A	72	HIS	14
1	A	132	TYR	13
1	A	50	ASP	12
1	A	121	ILE	11
1	A	102	ASN	6
1	A	116	ASN	6
1	A	78	LYS	4
1	A	133	ARG	4
1	A	75	PHE	3

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Models (Total)
1	A	123	SER	3
1	A	131	PRO	2
1	A	51	LYS	2
1	A	47	ILE	2
1	A	52	HIS	1
1	A	82	GLU	1
1	A	48	THR	1
1	A	120	GLU	1
1	A	117	PHE	1
1	A	71	CYS	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	72/88 (82%)	57±3 (79±4%)	15±3 (21±4%)	4	33
All	All	1440/1760 (82%)	1140 (79%)	300 (21%)	4	33

All 53 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	66	PHE	19
1	A	56	LEU	19
1	A	73	ASN	16
1	A	45	LEU	16
1	A	132	TYR	15
1	A	125	ILE	15
1	A	116	ASN	13
1	A	90	ILE	12
1	A	121	ILE	11
1	A	88	VAL	10
1	A	102	ASN	9
1	A	60	VAL	9
1	A	98	PHE	7
1	A	97	GLN	7
1	A	129	LEU	7

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Models (Total)
1	A	68	LYS	7
1	A	52	HIS	6
1	A	42	LYS	6
1	A	133	ARG	6
1	A	47	ILE	6
1	A	100	MET	6
1	A	82	GLU	5
1	A	70	GLU	5
1	A	57	THR	5
1	A	126	LEU	5
1	A	64	VAL	5
1	A	71	CYS	4
1	A	89	LEU	4
1	A	99	TYR	4
1	A	117	PHE	4
1	A	75	PHE	3
1	A	48	THR	3
1	A	84	ILE	3
1	A	63	LEU	2
1	A	128	LYS	2
1	A	54	GLU	2
1	A	94	LYS	2
1	A	101	THR	2
1	A	59	MET	2
1	A	58	GLU	2
1	A	83	SER	2
1	A	65	GLU	1
1	A	43	ARG	1
1	A	49	ASN	1
1	A	51	LYS	1
1	A	72	HIS	1
1	A	123	SER	1
1	A	86	SER	1
1	A	81	ASN	1
1	A	78	LYS	1
1	A	124	THR	1
1	A	62	LEU	1
1	A	130	ASN	1

### 6.3.3 RNA ⓘ

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 carbohydrates 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 84% for the well-defined parts and 81% for the entire structure.

### 7.1 Chemical shift list 1

File name: BMRB entry 16057

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	1141
Number of shifts mapped to atoms	1141
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	3

#### 7.1.2 Chemical shift referencing

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction $\pm$ precision, ppm	Suggested action
$^{13}\text{C}_\alpha$	95	$2.33 \pm 0.13$	Should be applied
$^{13}\text{C}_\beta$	90	$3.04 \pm 0.18$	Should be applied
$^{13}\text{C}'$	90	$2.43 \pm 0.06$	Should be applied
$^{15}\text{N}$	86	$1.26 \pm 0.31$	Should be applied

#### 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 84%, i.e. 833 atoms were assigned a chemical shift out of a possible 995. 12 out of 12 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	$^1\text{H}$	$^{13}\text{C}$	$^{15}\text{N}$
Backbone	366/385 (95%)	147/153 (96%)	150/158 (95%)	69/74 (93%)
Sidechain	439/527 (83%)	276/309 (89%)	156/198 (79%)	7/20 (35%)

*Continued on next page...*

Continued from previous page...

	Total	<sup>1</sup> H	<sup>13</sup> C	<sup>15</sup> N
Aromatic	28/83 (34%)	26/45 (58%)	2/36 (6%)	0/2 (0%)
Overall	833/995 (84%)	449/507 (89%)	308/392 (79%)	76/96 (79%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 81%, i.e. 1002 atoms were assigned a chemical shift out of a possible 1242. 15 out of 15 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	<sup>1</sup> H	<sup>13</sup> C	<sup>15</sup> N
Backbone	451/485 (93%)	180/193 (93%)	185/198 (93%)	86/94 (91%)
Sidechain	521/666 (78%)	327/391 (84%)	187/244 (77%)	7/31 (23%)
Aromatic	30/91 (33%)	28/49 (57%)	2/40 (5%)	0/2 (0%)
Overall	1002/1242 (81%)	535/633 (85%)	374/482 (78%)	93/127 (73%)

#### 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	133	ARG	NE	117.30	92.63 – 76.73	20.5
1	A	75	PHE	CZ	113.60	137.04 – 121.44	-10.0
1	A	121	ILE	CG1	18.50	36.54 – 18.94	-5.3

#### 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:

