



Full wwPDB NMR Structure Validation Report ⓘ

May 29, 2020 – 07:12 am BST

PDB ID : 5IRT
Title : Dimerization interface of the noncrystalline HIV-1 capsid protein lattice from solid state NMR spectroscopy of tubular assemblies
Authors : Bayro, M.J.; Tycko, R.
Deposited on : 2016-03-14

This is a Full wwPDB NMR Structure Validation 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 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 : 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 : 2.11
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.11

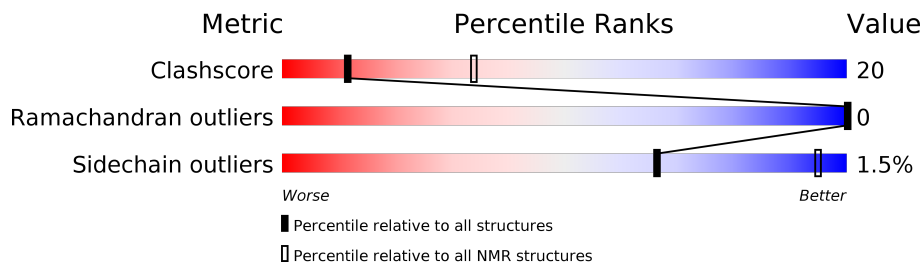
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

SOLID-STATE NMR

The overall completeness of chemical shifts assignment is 12%.

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	231	92%
1	B	231	92%

2 Ensemble composition and analysis i

This entry contains 20 models. Model 10 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *closest to the average*.

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:180-A:192, B:180-B:193 (27)	0.09	10

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 and 3 single-model clusters were found.

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

3 Entry composition

There is only 1 type of molecule in this entry. The entry contains 612 atoms, of which 300 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Capsid protein p24.

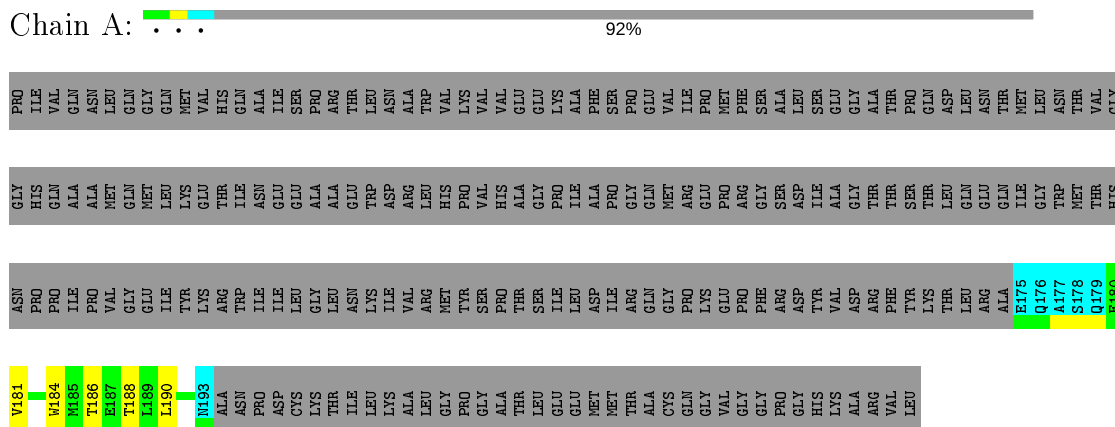
Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	19	306	96	150	26	33	1	0
1	B	19	306	96	150	26	33	1	0

4 Residue-property plots

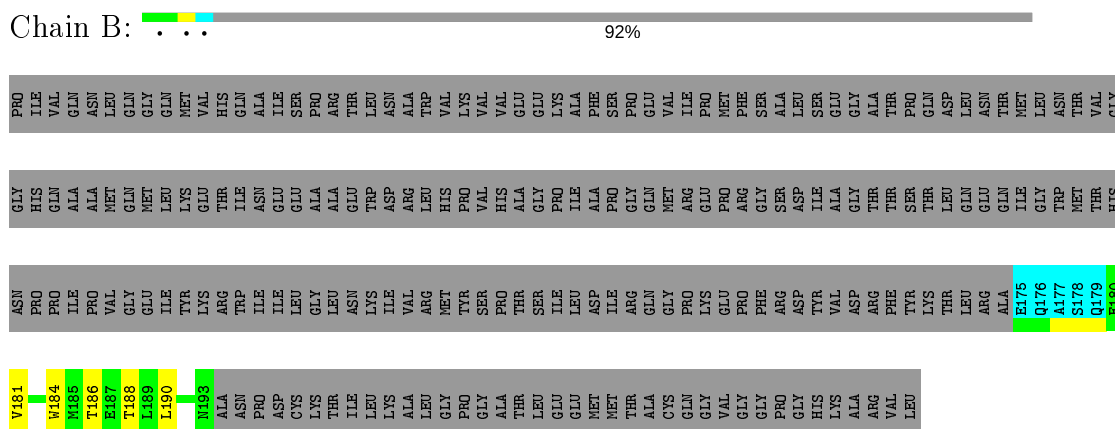
4.1 Average score per residue in the NMR ensemble

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: Capsid protein p24



- Molecule 1: Capsid protein p24



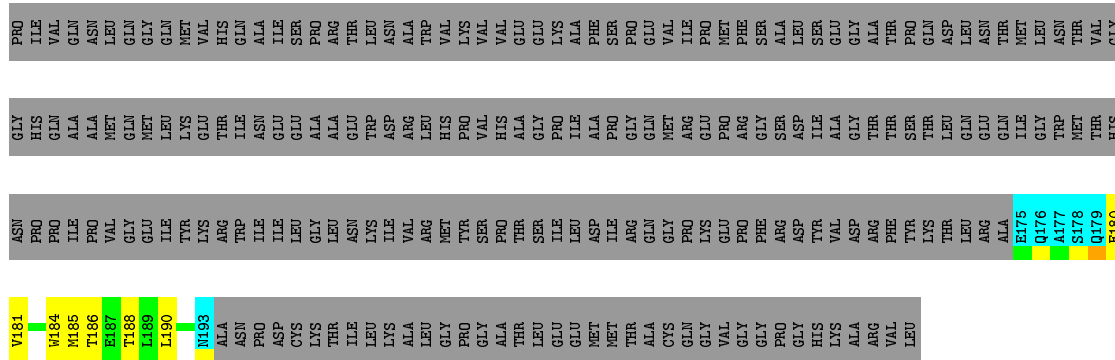
4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1

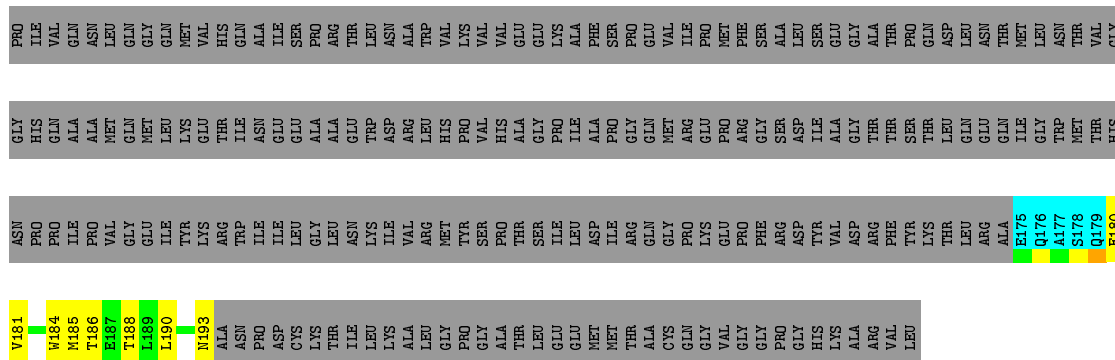
- Molecule 1: Capsid protein p24

Chain A:  92%



- Molecule 1: Capsid protein p24

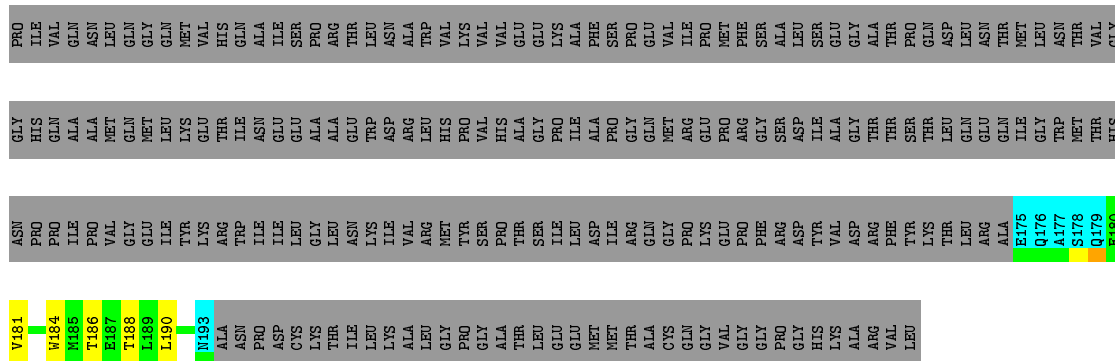
Chain B:  92%



4.2.2 Score per residue for model 2

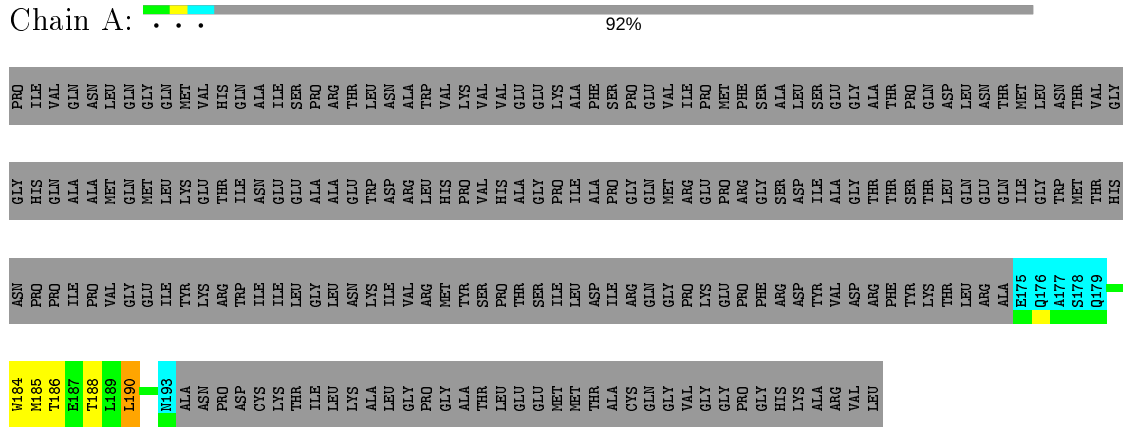
- Molecule 1: Capsid protein p24

Chain A:  92%

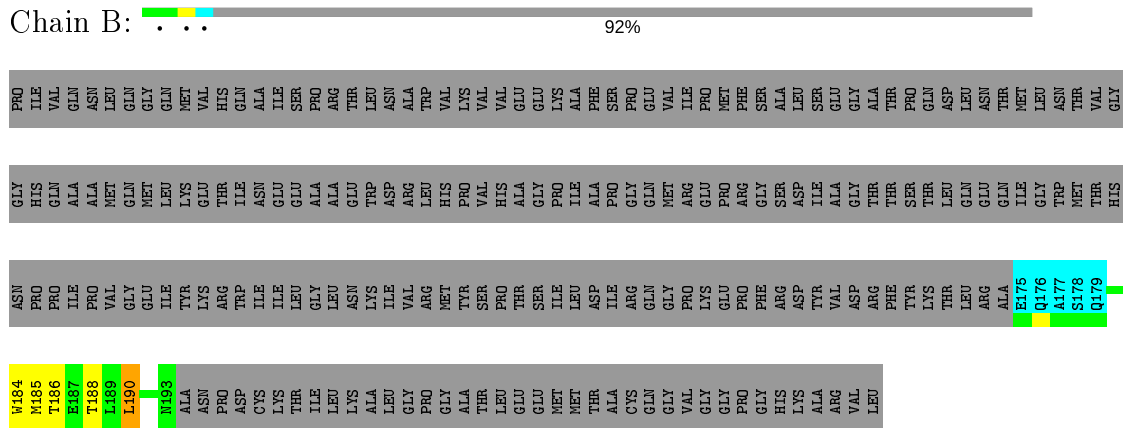


4.2.4 Score per residue for model 4

- Molecule 1: Capsid protein p24

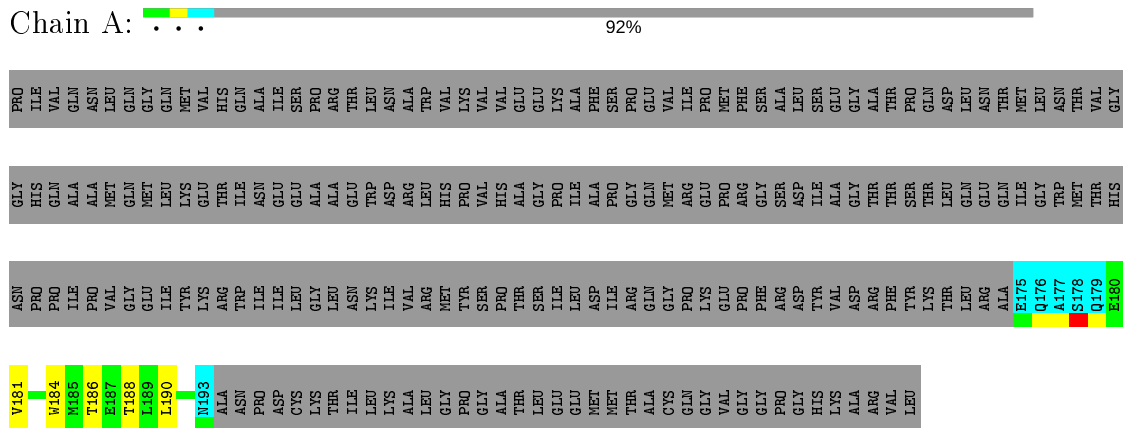


- Molecule 1: Capsid protein p24



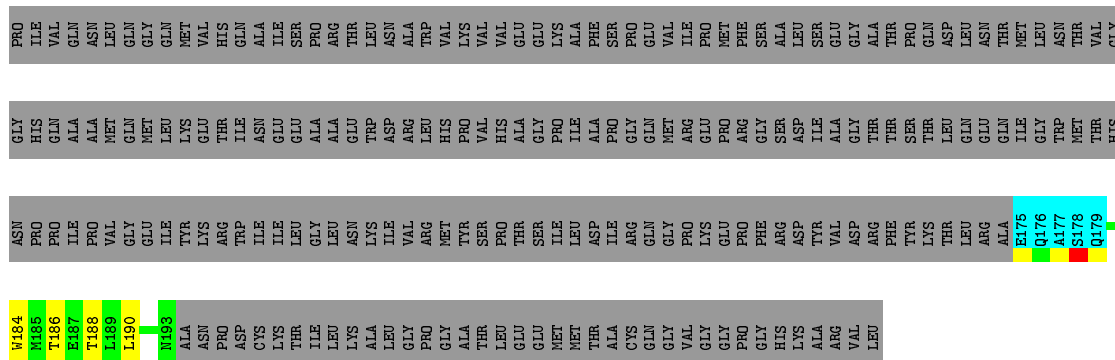
4.2.5 Score per residue for model 5

- Molecule 1: Capsid protein p24



- Molecule 1: Capsid protein p24

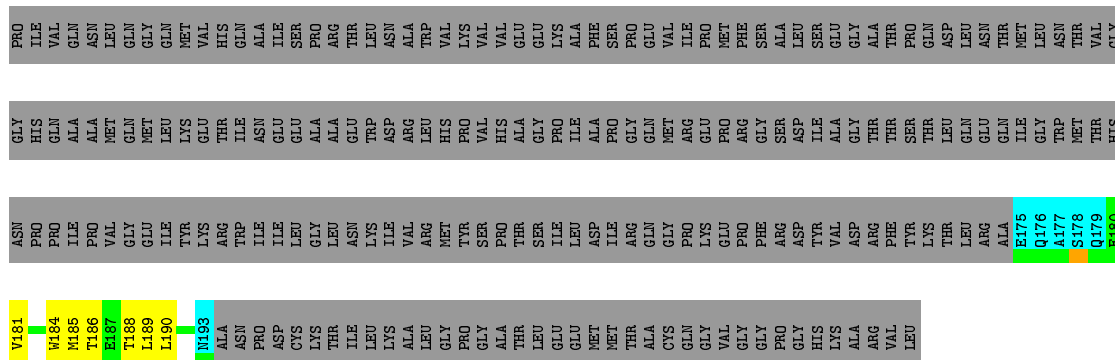
Chain B:  92%



4.2.6 Score per residue for model 6

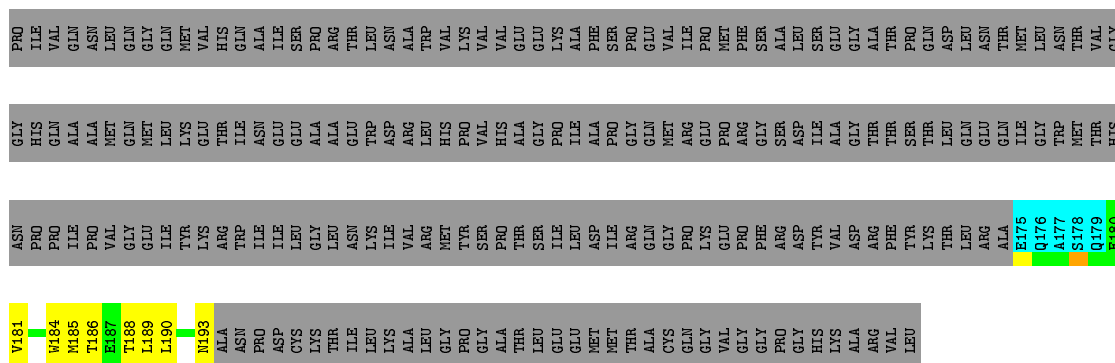
- Molecule 1: Capsid protein p24

Chain A:  92%



- Molecule 1: Capsid protein p24

Chain B:  92%



• Molecule 1: Capsid protein p24

Chain B:  92%

PRO	ILE	VAL	GLN	ASN	ALA	LEU	GLN	GLY	LEU	MET	LYS	VAL	HIS	GLN	ALA	ILE	GLU	SER	LEU	PRO	ALA	ARG	THR	LEU	ASN	ASP	LEU	TRP	GLY	ALA	ARG	LEU	VAL	PRO	LYS	ILE	ALA	PHE	PRO	SER	ILE	GLN	GLY	VAL	ILE	PRO	MET	PRO	ARG	GLY	SER	ALA	LEU	SER	ASP	ILE	GLU	GLY	THR	THR	ALA	ARG	TYR	SER	PRO	GLN	ASP	LEU	LEU	ASN	GLN	THR	ILE	GLY	LEU	ASN	THR	VAL	GLY								
GLY	HIS	GLN	ALA	ALA	MET	GLN	MET	GLY	LEU	LYS	THR	THR	ILE	ASN	GLN	ILE	GLU	LEU	ALA	ALA	ARG	TRP	GLY	GLU	THR	ASP	ILE	VAL	VAL	VAL	HIS	GLY	PRO	ALA	ILE	LEU	ASP	PRO	ILE	ALA	PHE	PRO	SER	GLY	GLN	VAL	ILE	PRO	MET	PRO	ARG	GLY	SER	ALA	LEU	SER	ASP	ILE	GLU	GLY	THR	THR	ALA	ARG	TYR	SER	PRO	GLN	ASP	LEU	LEU	ASN	GLN	THR	ILE	GLY	LEU	ASN	THR	VAL	GLY						
ASN	PRO	PRO	ILE	GLN	VAL	PRO	VAL	GLY	GLU	ILE	TYR	LYS	THR	ARG	TRP	ILE	LEU	LEU	GLY	ALA	ARG	LEU	ASN	GLY	TRP	LYS	ILE	ALA	PRO	PRO	THR	THR	GLY	VAL	ARG	GLY	LEU	ASP	ILE	ALA	PHE	PRO	ILE	ALA	PRO	SER	GLY	GLN	VAL	ILE	PRO	GLY	GLY	PRO	PHE	PRO	ARG	GLY	SER	ALA	LEU	SER	ASP	ILE	GLU	GLY	THR	THR	ALA	ARG	TYR	SER	PRO	GLN	ASP	LEU	LEU	ASN	GLN	THR	ALA	E175	Q176	A177	S178	Q179	
H184	R185	T186	E187	T188	L189	L190	N193	ALA	ASN	PRO	PRO	ASP	CYS	LYS	THR	ILE	GLU	LEU	ILE	GLY	ALA	ARG	LEU	ASN	GLY	TRP	LYS	ILE	ALA	PRO	PRO	THR	THR	GLY	VAL	ARG	GLY	LEU	ASP	ILE	ALA	PHE	PRO	ILE	ALA	PRO	SER	GLY	GLN	VAL	ILE	PRO	GLY	GLY	PRO	PHE	PRO	ARG	GLY	SER	ALA	LEU	SER	ASP	ILE	GLU	GLY	THR	THR	ALA	ARG	TYR	SER	PRO	GLN	ASP	LEU	LEU	ASN	GLN	THR	ALA	E175	Q176	A177	S178	Q179

4.2.12 Score per residue for model 12

• Molecule 1: Capsid protein p24

Chain A:  92%

PRO	ILE	VAL	GLN	ASN	ALA	LEU	GLN	GLY	LEU	MET	LYS	VAL	HIS	GLN	ALA	ILE	GLU	SER	LEU	PRO	ALA	ARG	THR	LEU	ASN	ASP	LEU	TRP	GLY	ALA	ARG	LEU	VAL	PRO	LYS	ILE	ALA	PHE	PRO	SER	ILE	GLN	GLY	VAL	ILE	PRO	MET	PRO	ARG	GLY	SER	ALA	LEU	SER	ASP	ILE	GLU	GLY	THR	THR	ALA	ARG	TYR	SER	PRO	GLN	ASP	LEU	LEU	ASN	GLN	THR	ILE	GLY	LEU	ASN	THR	VAL	GLY								
GLY	HIS	GLN	ALA	ALA	ALA	MET	GLN	MET	GLY	LEU	LYS	THR	THR	ILE	ASN	GLN	ILE	GLU	LEU	GLY	ALA	ARG	TRP	GLY	GLU	THR	ASP	ILE	VAL	VAL	HIS	GLY	PRO	ALA	ILE	LEU	ASP	PRO	ILE	ALA	PHE	PRO	SER	GLY	GLN	VAL	ILE	PRO	MET	PRO	ARG	GLY	SER	ALA	LEU	SER	ASP	ILE	GLU	GLY	THR	THR	ALA	ARG	TYR	SER	PRO	GLN	ASP	LEU	LEU	ASN	GLN	THR	ILE	GLY	LEU	ASN	THR	VAL	GLY						
ASN	PRO	PRO	ILE	GLN	VAL	PRO	VAL	GLY	GLU	ILE	TYR	LYS	THR	ARG	TRP	ILE	LEU	LEU	GLY	ALA	ARG	LEU	ASN	GLY	TRP	LYS	ILE	ALA	PRO	PRO	THR	THR	GLY	VAL	ARG	GLY	LEU	ASP	ILE	ALA	PHE	PRO	ILE	ALA	PRO	SER	GLY	GLN	VAL	ILE	PRO	GLY	GLY	PRO	PHE	PRO	ARG	GLY	SER	ALA	LEU	SER	ASP	ILE	GLU	GLY	THR	THR	ALA	ARG	TYR	SER	PRO	GLN	ASP	LEU	LEU	ASN	GLN	THR	ALA	E175	Q176	A177	S178	Q179	E180
V181	K182	M183	M184	M185	T186	E187	T188	L189	L190	N193	ALA	ASN	PRO	PRO	ASP	CYS	LYS	THR	ILE	GLU	LEU	ILE	GLY	TRP	LYS	ILE	ALA	PRO	PRO	THR	THR	GLY	VAL	ARG	GLY	LEU	ASP	ILE	ALA	PHE	PRO	ILE	ALA	PRO	SER	GLY	GLN	VAL	ILE	PRO	MET	PRO	ARG	GLY	SER	ALA	LEU	SER	ASP	ILE	GLU	GLY	THR	THR	ALA	ARG	TYR	SER	PRO	GLN	ASP	LEU	LEU	ASN	GLN	THR	ILE	GLY	LEU	ASN	THR	VAL	GLY				

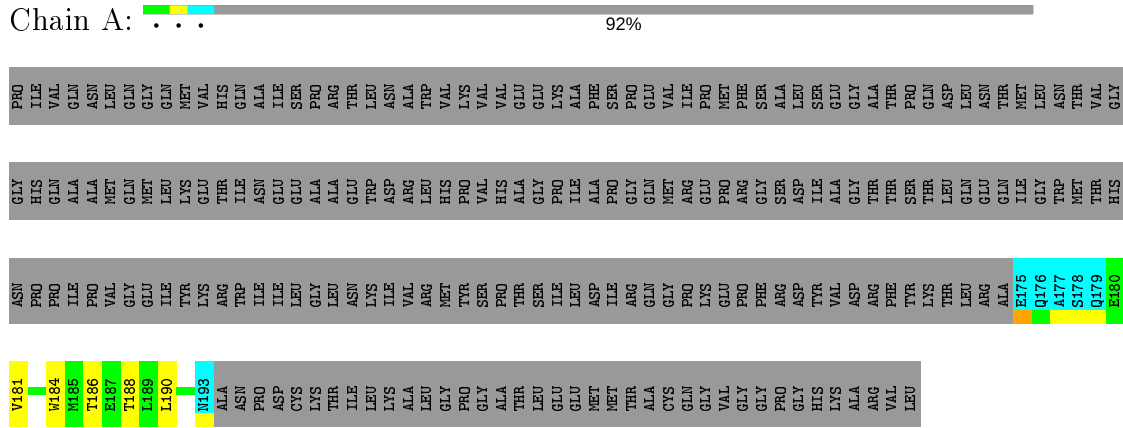
• Molecule 1: Capsid protein p24

Chain B:  92%

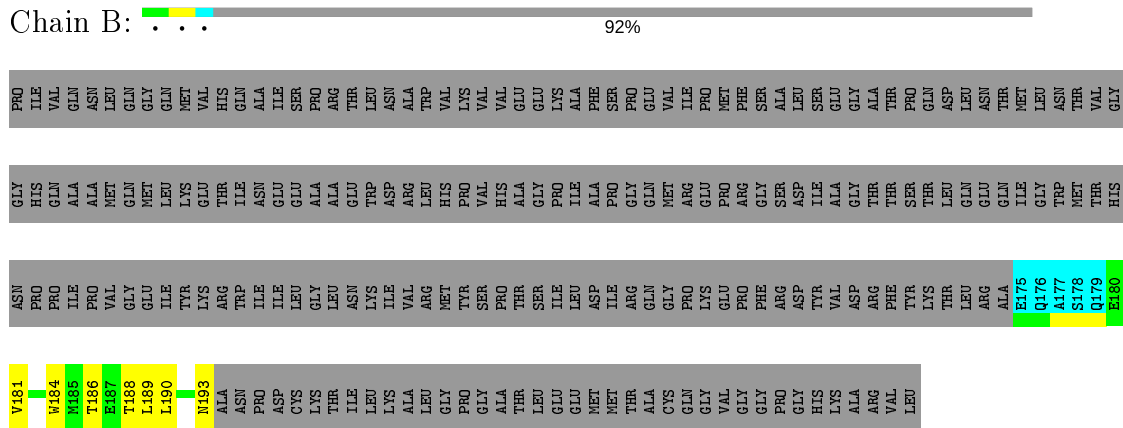
PRO	ILE	VAL	GLN	ASN	ALA	LEU	GLN	GLY	LEU	MET	LYS	VAL	HIS	GLN	ALA	ILE	GLU	SER	LEU	PRO	ALA	ARG	THR	LEU	ASN	ASP	LEU	TRP	GLY	ALA	ARG	LEU	VAL	PRO	LYS	ILE	ALA	PHE	PRO	SER	ILE	GLN	GLY	VAL	ILE	PRO	MET	PRO	ARG	GLY	SER	ALA	LEU	SER	ASP	ILE	GLU	GLY	THR	THR	ALA	ARG	TYR	SER	PRO	GLN	ASP	LEU	LEU	ASN	GLN	THR	ILE	GLY	LEU	ASN	THR	VAL	GLY								
GLY	HIS	GLN	ALA	ALA	ALA	MET	GLN	MET	GLY	LEU	LYS	THR	THR	ILE	ASN	GLN	ILE	GLU	LEU	GLY	ALA	ARG	TRP	GLY	GLU	THR	ASP	ILE	VAL	VAL	HIS	GLY	PRO	ALA	ILE	LEU	ASP	PRO	ILE	ALA	PHE	PRO	SER	GLY	GLN	VAL	ILE	PRO	MET	PRO	ARG	GLY	SER	ALA	LEU	SER	ASP	ILE	GLU	GLY	THR	THR	ALA	ARG	TYR	SER	PRO	GLN	ASP	LEU	LEU	ASN	GLN	THR	ILE	GLY	LEU	ASN	THR	VAL	GLY						
ASN	PRO	PRO	ILE	GLN	VAL	PRO	VAL	GLY	GLU	ILE	TYR	LYS	THR	ARG	TRP	ILE	LEU	LEU	GLY	ALA	ARG	LEU	ASN	GLY	TRP	LYS	ILE	ALA	PRO	PRO	THR	THR	GLY	VAL	ARG	GLY	LEU	ASP	ILE	ALA	PHE	PRO	ILE	ALA	PRO	SER	GLY	GLN	VAL	ILE	PRO	GLY	GLY	PRO	PHE	PRO	ARG	GLY	SER	ALA	LEU	SER	ASP	ILE	GLU	GLY	THR	THR	ALA	ARG	TYR	SER	PRO	GLN	ASP	LEU	LEU	ASN	GLN	THR	ALA	E175	Q176	A177	S178	Q179	E180
V181	K182	M183	M184	M185	T186	E187	T188	L189	L190	N193	ALA	ASN	PRO	PRO	ASP	CYS	LYS	THR	ILE	GLU	LEU	ILE	GLY	TRP	LYS	ILE	ALA	PRO	PRO	THR	THR	GLY	VAL	ARG	GLY	LEU	ASP	ILE	ALA	PHE	PRO	ILE	ALA	PRO	SER	GLY	GLN	VAL	ILE	PRO	MET	PRO	ARG	GLY	SER	ALA	LEU	SER	ASP	ILE	GLU	GLY	THR	THR	ALA	ARG	TYR	SER	PRO	GLN	ASP	LEU	LEU	ASN	GLN	THR	ILE	GLY	LEU	ASN	THR	VAL	GLY				

4.2.16 Score per residue for model 16

- Molecule 1: Capsid protein p24

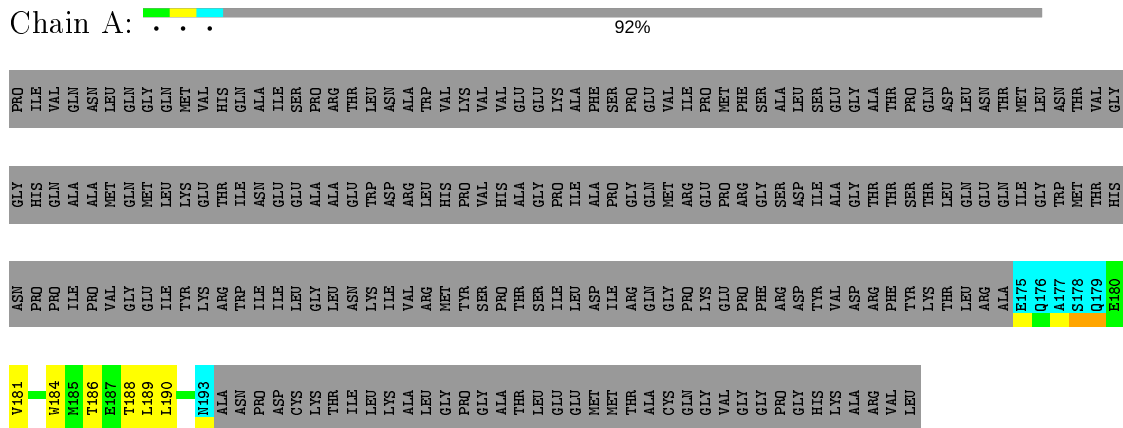


- Molecule 1: Capsid protein p24



4.2.17 Score per residue for model 17

- Molecule 1: Capsid protein p24



- Molecule 1: Capsid protein p24

Chain B: 92%

PRO
ILE
VAL
GLN
ASN
LEU
GLN
GLY
MET
VAL
HIS
GLN
ALA
SER
PRO
ARG
THR
LEU
ASN
VAL
GLU
ALA
TRP
VAL
LYS
VAL
PHE
SER
PRO
GLN
VAL
ILE
PHE
SER
GLY
MET
PRO
MET
PHE
SER
ALA
LEU
SER
SER
GLU
GLY
ALA
THR
THR
PRO
GLN
ASP
LEU
LEU
ASN
THR
MET
LEU
ASN
THR
VAL
GLY

GLY
HIS
GLN
ALA
ALA
MET
GLN
MET
LEU
LYS
GLU
THR
ILE
ASN
ILE
GLU
LEU
ALA
GLY
LEU
ALA
GLU
TRP
ASP
ARG
LEU
HIS
PRO
VAL
HIS
GLY
PRO
ILE
ILE
ALA
PHE
SER
GLY
GLN
MET
ARG
GLU
PRO
ARG
GLY
SER
SER
ASP
ILE
ALA
GLY
THR
THR
SER
THR
LEU
LEU
ARG
GLU
GLN
THR
ILE
GLY
TRP
MET
THR
HIS

ASN
PRO
ILE
PRO
ILE
VAL
PRO
VAL
GLY
GLU
ILE
TYR
LYS
ARG
TRP
ILE
ILE
LEU
GLY
LEU
ASN
LYS
ILE
VAL
ARG
MET
TYR
SER
PRO
THR
SER
ILE
LEU
ASP
ILE
ARG
GLY
PRO
LYS
GLU
PHE
GLY
ARG
ASP
TYR
VAL
ASP
PHE
THR
LYS
THR
LEU
ARG
ALA
E175
Q176
A177
S178
Q179
E180

V181
K182
M183
W184
M185
T186
E187
T188
L189
L190
M193
ALA
ASN
PRO
ASP
CYS
LYS
THR
ILE
LEU
LYS
ALA
LEU
GLY
PRO
GLY
THR
LEU
GLU
MET
THR
CYS
GLN
GLY
VAL
GLY
PRO
GLY
HIS
ALA
ARG
VAL
LEU

5 Refinement protocol and experimental data overview

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

Of the 100 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
X-PLOR NIH	structure calculation	
CNS	refinement	

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)	input_cs.cif
Number of chemical shift lists	1
Total number of shifts	791
Number of shifts mapped to atoms	791
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	12%

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.

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

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	110	112	112	5±1
1	B	118	118	118	5±1
All	All	4560	4600	4600	181

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 20.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:184:TRP:CD1	1:A:188:THR:HG21	0.68	2.24	6	3
1:B:184:TRP:CD1	1:B:188:THR:HG21	0.66	2.26	6	3
1:B:186:THR:HG23	1:B:187:GLU:N	0.63	2.09	7	2
1:A:186:THR:HG23	1:A:187:GLU:N	0.61	2.10	7	2
1:B:184:TRP:CD1	1:B:188:THR:OG1	0.58	2.52	18	17
1:A:184:TRP:CD1	1:A:188:THR:OG1	0.58	2.53	14	17
1:A:186:THR:CG2	1:A:187:GLU:N	0.57	2.68	11	2
1:B:186:THR:CG2	1:B:187:GLU:N	0.56	2.67	7	2
1:A:186:THR:O	1:A:190:LEU:CB	0.55	2.55	20	16
1:B:186:THR:O	1:B:190:LEU:N	0.55	2.40	3	4
1:B:186:THR:O	1:B:190:LEU:CB	0.53	2.56	20	16
1:A:186:THR:O	1:A:190:LEU:N	0.53	2.40	3	4
1:B:184:TRP:O	1:B:188:THR:CB	0.52	2.57	4	13

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:181:VAL:HG22	1:B:181:VAL:HG22	0.51	1.82	9	2
1:A:184:TRP:O	1:A:188:THR:CB	0.51	2.59	8	13
1:A:181:VAL:HG12	1:B:181:VAL:HG12	0.50	1.82	3	1
1:B:189:LEU:O	1:B:193:ASN:ND2	0.49	2.46	18	3
1:B:184:TRP:O	1:B:188:THR:OG1	0.48	2.32	4	7
1:B:190:LEU:HD13	1:B:190:LEU:O	0.47	2.10	4	2
1:B:184:TRP:NE1	1:B:188:THR:HG21	0.47	2.24	20	2
1:A:190:LEU:O	1:A:190:LEU:HD13	0.47	2.10	4	2
1:A:186:THR:O	1:A:190:LEU:HB2	0.47	2.10	8	6
1:B:186:THR:O	1:B:190:LEU:HB2	0.46	2.10	8	4
1:B:180:GLU:CD	1:B:180:GLU:N	0.46	2.68	15	1
1:A:184:TRP:NE1	1:A:188:THR:HG21	0.46	2.25	20	2
1:A:180:GLU:CD	1:A:180:GLU:N	0.46	2.69	15	1
1:A:184:TRP:O	1:A:188:THR:OG1	0.45	2.21	3	9
1:A:183:ASN:OD1	1:A:187:GLU:OE2	0.43	2.37	20	1
1:A:184:TRP:CZ3	1:B:184:TRP:CZ3	0.43	3.06	8	6
1:A:182:LYS:CB	1:A:182:LYS:NZ	0.43	2.81	9	1
1:B:183:ASN:OD1	1:B:187:GLU:OE2	0.43	2.37	20	1
1:B:190:LEU:O	1:B:193:ASN:OD1	0.43	2.35	9	3
1:A:186:THR:O	1:A:190:LEU:CG	0.43	2.67	7	2
1:B:185:MET:O	1:B:189:LEU:CG	0.42	2.67	6	3
1:B:186:THR:O	1:B:190:LEU:CG	0.42	2.68	14	2
1:B:193:ASN:OD1	1:B:193:ASN:C	0.41	2.58	6	1
1:B:182:LYS:NZ	1:B:182:LYS:CB	0.41	2.82	9	1
1:A:185:MET:O	1:A:189:LEU:CG	0.41	2.69	6	2
1:A:182:LYS:HB3	1:A:182:LYS:HZ3	0.41	1.75	9	1
1:B:189:LEU:O	1:B:193:ASN:CB	0.40	2.69	17	1

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	13/231 (6%)	13±0 (100±2%)	0±0 (0±2%)	0±0 (0±0%)	100	100
1	B	13/231 (6%)	13±0 (100±2%)	0±0 (0±2%)	0±0 (0±0%)	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	520/9240 (6%)	518 (100%)	2 (0%)	0 (0%)	100 100

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	13/194 (7%)	13±0 (98±3%)	0±0 (2±3%)	66 95
1	B	14/194 (7%)	14±0 (99±3%)	0±0 (1±3%)	68 95
All	All	540/7760 (7%)	532 (99%)	8 (1%)	66 95

All 6 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	B	190	LEU	2
1	A	190	LEU	2
1	B	182	LYS	1
1	A	180	GLU	1
1	B	180	GLU	1
1	A	182	LYS	1

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 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 [i](#)

The completeness of assignment taking into account all chemical shift lists is 12% for the well-defined parts and 12% for the entire structure.

7.1 Chemical shift list 1

File name: input_cs.cif

Chemical shift list name: *assigned_chem_shift_list_1*

7.1.1 Bookkeeping [i](#)

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	791
Number of shifts mapped to atoms	791
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	1

7.1.2 Chemical shift referencing [i](#)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction \pm precision, ppm	Suggested action
$^{13}\text{C}_\alpha$	184	-0.36 ± 0.77	None needed (< 0.5 ppm)
$^{13}\text{C}_\beta$	158	0.71 ± 0.76	None needed (imprecise)
$^{13}\text{C}'$	179	-0.39 ± 0.84	None needed (< 0.5 ppm)
^{15}N	188	2.74 ± 0.42	Should be applied

7.1.3 Completeness of resonance assignments [i](#)

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 12%, i.e. 42 atoms were assigned a chemical shift out of a possible 352. 0 out of 8 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^1H	^{13}C	^{15}N
Backbone	30/135 (22%)	0/54 (0%)	20/54 (37%)	10/27 (37%)
Sidechain	12/193 (6%)	0/110 (0%)	12/76 (16%)	0/7 (0%)

Continued on next page...

Continued from previous page...

	Total	¹H	¹³C	¹⁵N
Aromatic	0/24 (0%)	0/12 (0%)	0/10 (0%)	0/2 (0%)
Overall	42/352 (12%)	0/176 (0%)	32/140 (23%)	10/36 (28%)

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.

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 12%, i.e. 55 atoms were assigned a chemical shift out of a possible 478. 0 out of 8 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	¹H	¹³C	¹⁵N
Backbone	41/190 (22%)	0/76 (0%)	28/76 (37%)	13/38 (34%)
Sidechain	14/264 (5%)	0/152 (0%)	14/100 (14%)	0/12 (0%)
Aromatic	0/24 (0%)	0/12 (0%)	0/10 (0%)	0/2 (0%)
Overall	55/478 (12%)	0/240 (0%)	42/186 (23%)	13/52 (25%)

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](#)

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	1	PRO	N	49.48	167.35 – 100.35	-12.6

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:

