Full wwPDB NMR Structure Validation Report (i)

Jun 16, 2024 – 11:06 AM EDT

PDB ID	:	2LO6
BMRB ID	:	17173
Title	:	Structure of Nrd1 CID bound to phosphorylated RNAP II CTD
Authors	:	Kubicek, K.; Cerna, H.; Pasulka, J.; Holub, P.; Hrossova, D.; Loehr, F.; Hofr, C.; Vanacova, S.; Stefl, R.
Deposited on	:	2012-01-17

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 (i) 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 (1)) were used in the production of this report:

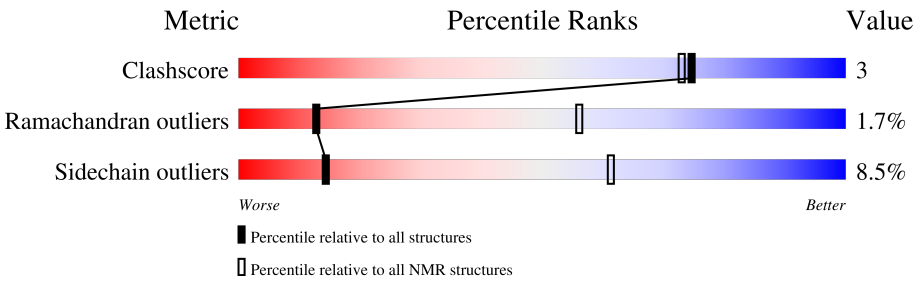
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
SOLUTION NMR

The overall completeness of chemical shifts assignment is 65%.

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 <=5%

Mol	Chain	Length	Quality of chain
1	A	161	61% 5% • 29% •
2	B	14	43% 57%

2 Ensemble composition and analysis (i)

This entry contains 20 models. Model 6 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:6-A:18, A:24-A:81, A:93-A:129, B:160-B:165 (114)	0.49	6

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 2 single-model clusters were found.

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

3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 2635 atoms, of which 1309 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Protein NRD1.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	154	2434	760	1219	209	240	6	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	154	LEU	-	EXPRESSION TAG	UNP P53617
A	155	GLU	-	EXPRESSION TAG	UNP P53617
A	156	HIS	-	EXPRESSION TAG	UNP P53617
A	157	HIS	-	EXPRESSION TAG	UNP P53617
A	158	HIS	-	EXPRESSION TAG	UNP P53617
A	159	HIS	-	EXPRESSION TAG	UNP P53617
A	160	HIS	-	EXPRESSION TAG	UNP P53617
A	161	HIS	-	EXPRESSION TAG	UNP P53617

- Molecule 2 is a protein called DNA-directed RNA polymerase II subunit RPB1.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	P	
2	B	14	201	64	90	14	31	2	0

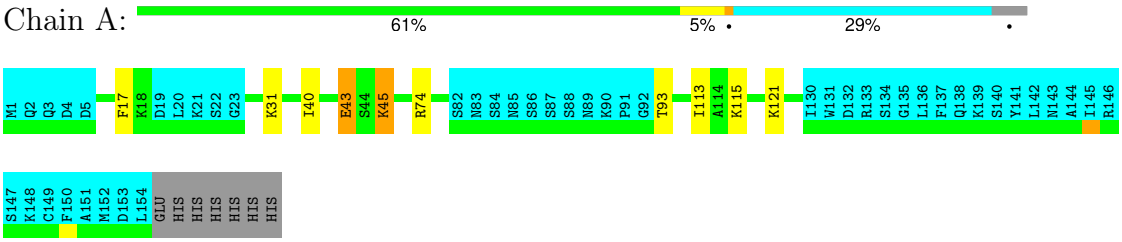
4 Residue-property plots (i)

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.

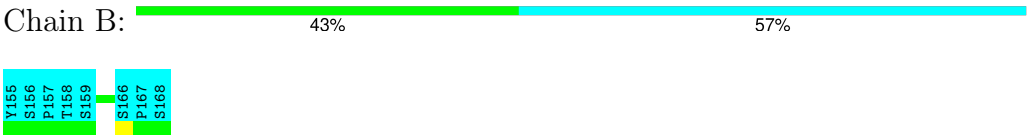
• Molecule 1: Protein NRD1

Chain A:



• Molecule 2: DNA-directed RNA polymerase II subunit RPB1

Chain B:



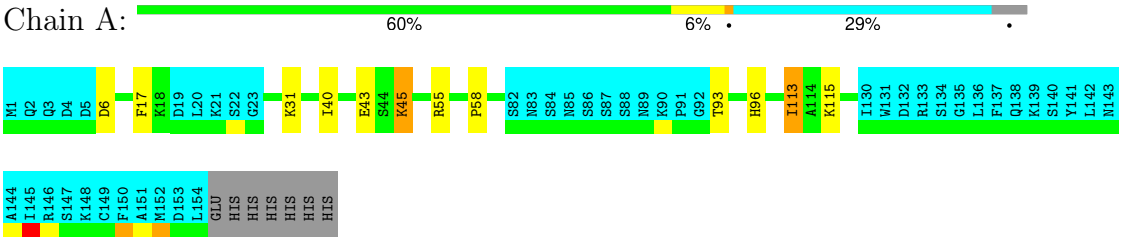
4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1

• Molecule 1: Protein NRD1

Chain A:



• Molecule 2: DNA-directed RNA polymerase II subunit RPB1

Chain B: 43% 57%

Y155 S156 P157 T158 S159 S166 P167 S168

4.2.2 Score per residue for model 2

- Molecule 1: Protein NRD1

Chain A: 59% 6% • 29% •

M1 Q2 Q3 D4 D5 F17 K18 D19 L20 K21 S22 G23 K31 I40 D41 I42 E43 S44 K45 I46 I47 R55 S82 N83 S84 N85 S86 S87 S88 N89 K90 P91 G92 T93 A97 I113 A114 K115 K121 I130 W131 D132 R133 S134 G135 L136 F137 Q138 K139 S140 Y141

L142 N143 A144 I145 R146 S147 K148 C149 F150 A151 M152 D153 L154 GLU HIS HIS HIS HIS HIS HIS

- Molecule 2: DNA-directed RNA polymerase II subunit RPB1

Chain B: 36% 7% 57%

Y155 S156 P157 T158 S159 P160 S166 P167 S168

4.2.3 Score per residue for model 3

- Molecule 1: Protein NRD1

Chain A: 57% 9% • 29% •

M1 Q2 Q3 D4 D5 D6 F17 K18 D19 L20 K21 S22 G23 K31 L32 T33 I40 D41 I42 E43 S44 K45 C57 H61 D70 S71 S82 N83 S84 N85 S86 S87 S88 N89 K90 P91 G92 T93 I113 A114 K115 K121 I130 W131 D132 R133 S134 G135 L136 F137

Q138 K139 S140 Y141 L142 N143 A144 I145 R146 S147 K148 C149 F150 A151 M152 D153 L154 GLU HIS HIS HIS HIS HIS HIS

- Molecule 2: DNA-directed RNA polymerase II subunit RPB1

Chain B: 36% 7% 57%

Y155 S156 P157 T158 S159 P160 S166 P167 S168

4.2.4 Score per residue for model 4

- Molecule 1: Protein NRD1

Chain A: 61% 5% • • 29% •

M1 Q2 Q3 D4 D5 F17 K18 D19 L20 K21 S22 G23 K31 I40 E43 S44 K45 R74 S82 N83 S84 N85 S86 S87 S88 N89 K90 P91 G92 T93 I113 A114 K115 K121 I130 W131 D132 R133 S134 G135 L136 F137 Q138 K139 S140 Y141 L142 N143 A144 I145 R146

S147 K148 C149 F150 A151 M152 D153 L154 GLU HIS HIS HIS HIS HIS HIS HIS

- Molecule 2: DNA-directed RNA polymerase II subunit RPB1

Chain B: 36% 7% 57%

Y155 S156 P157 T158 S159 Y162 S166 P167 S168

4.2.5 Score per residue for model 5

- Molecule 1: Protein NRD1

Chain A: 58% 6% • • 29% •

M1 Q2 Q3 D4 D5 F17 K18 D19 L20 K21 S22 G23 K31 L32 T33 I40 D41 I42 E43 S44 K45 S71 R74 S82 N83 S84 N85 S86 S87 S88 N89 K90 P91 G92 T93 L109 I113 A114 K115 Q118 I130 W131 D132 R133 S134 G135 L136 F137 Q138 K139

S140 Y141 L142 N143 A144 T145 R146 S147 K148 C149 F150 A151 M152 D153 L154 GLU HIS HIS HIS HIS HIS HIS HIS

- Molecule 2: DNA-directed RNA polymerase II subunit RPB1

Chain B: 36% 7% 57%

Y155 S156 P157 T158 S159 Y162 S166 P167 S168

4.2.6 Score per residue for model 6 (medoid)

- Molecule 1: Protein NRD1

Chain A: 61% 5% • 29% •

M1 Q2 Q3 D4 D5 F17 K18 D19 L20 K21 S22 G23 K31 I40 E43 S44 K45 S82 N83 S84 N85 S86 S87 S88 N89 K90 P91 G92 T93 L109 I113 A114 K115 K121 I130 W131 D132 R133 S134 G135 L136 F137 Q138 K139 S140 Y141 L142 N143 A144 I145 R146

S147 K148 C149 F150 A151 M152 D153 L154 GLU HIS HIS HIS HIS HIS HIS HIS

- Molecule 2: DNA-directed RNA polymerase II subunit RPB1

Chain B: 36% 7% 57%

Y155 S156 P157 T158 S159 Y162 S166 P167 S168

4.2.7 Score per residue for model 7

- Molecule 1: Protein NRD1

Chain A: 61% 29%

M1 Q2 Q3 D4 D5 F17 K18 D19 L20 K21 S22 G23 K31 I40 E43 S44 K45 R74 S82 N83 S84 N85 S86 S87 S88 N89 K90 P91 G92 T93 I113 A114 K115 I130 W131 D132 R133 S134 G135 L136 F137 Q138 K139 S140 Y141 L142 N143 A144 I145 R146 S147 K148

C149 F150 A151 M152 D153 L154 GLU HIS HIS HIS HIS HIS HIS

- Molecule 2: DNA-directed RNA polymerase II subunit RPB1

Chain B: 36% 7% 57%

Y155 S156 P157 T158 S159 Y162 S166 P167 S168

4.2.8 Score per residue for model 8

- Molecule 1: Protein NRD1

Chain A: 60% 7% 29%

M1 Q2 Q3 D4 D5 F17 K18 D19 L20 K21 S22 G23 K31 I40 E43 S44 K45 R55 P58 S82 N83 S84 N85 S86 S87 S88 N89 K90 P91 G92 T93 L109 I113 Q118 K121 I130 W131 D132 R133 S134 G135 L136 F137 Q138 K139 S140 Y141 L142

N143 A144 I145 R146 S147 K148 C149 F150 A151 M152 D153 L154 GLU HIS HIS HIS HIS HIS HIS

- Molecule 2: DNA-directed RNA polymerase II subunit RPB1

Chain B: 43% 57%

Y155 S156 P157 T158 S159 S166 P167 S168

4.2.9 Score per residue for model 9

- Molecule 1: Protein NRD1

Chain A: 60% 6% • 29% •

M1 Q2 Q3 D4 D5 F17 K18 D19 L20 K21 S22 G23 K31 I40 D41 I42 E43 S44 K45 S82 N83 S84 N85 S86 S87 S88 N89 K90 P91 G92 T93 K115 S116 N117 H120 R125 I130 W131 D132 R133 S134 G135 L136 F137 Q138 K139 S140 Y141 L142 N143 A144 I145

R146 S147 K148 C149 F150 A151 M152 D153 L154 GLU HIS HIS HIS HIS HIS HIS

• Molecule 2: DNA-directed RNA polymerase II subunit RPB1

Chain B: 43% 57%

Y155 S156 P157 T158 S159 S166 P167 S168

4.2.10 Score per residue for model 10

• Molecule 1: Protein NRD1

Chain A: 59% 7% •• 29% •

M1 Q2 Q3 D4 D5 D6 F17 K18 D19 L20 K21 S22 G23 K31 L32 T33 I42 E43 S44 K45 S71 R74 S82 N83 S84 N85 S86 S87 S88 N89 K90 P91 G92 T93 I113 A114 K115 K121 I130 W131 D132 R133 S134 G135 L136 F137 Q138 K139 S140 Y141 L142

N143 A144 I145 R146 S147 K148 C149 F150 A151 M152 D153 L154 GLU HIS HIS HIS HIS HIS HIS

• Molecule 2: DNA-directed RNA polymerase II subunit RPB1

Chain B: 36% 7% 57%

Y155 S156 P157 T158 S159 Y162 S166 P167 S168

4.2.11 Score per residue for model 11

• Molecule 1: Protein NRD1

Chain A: 60% 6% •• 29% •

M1 Q2 Q3 D4 D5 F17 K18 D19 L20 K21 S22 G23 K31 I40 E43 S44 K45 P58 R74 S82 N83 S84 N85 S86 S87 S88 N89 K90 P91 G92 T93 I113 A114 K115 K121 I130 W131 D132 R133 S134 G135 L136 F137 Q138 K139 S140 Y141 L142 N143 A144

I145 R146 S147 K148 C149 F150 A151 M152 D153 L154 GLU HIS HIS HIS HIS HIS HIS

• Molecule 2: DNA-directed RNA polymerase II subunit RPB1

Chain B: 36% 7% 57%

Y155 S156 P157 T158 S159 Y162 S166 P167 S168

4.2.12 Score per residue for model 12

- Molecule 1: Protein NRD1

Chain A: 60% 6% 29%

M1 Q2 Q3 D4 D5 L14 D19 L20 K21 S22 G23 K31 I40 E43 S44 K45 Y53 S82 N83 S84 N85 S86 S87 S88 N89 K90 P91 G92 T93 S110 K115 K121 M126 I130 W131 D132 R133 S134 G135 L136 F137 Q138 K139 S140 Y141 L142 N143 A144

I145 R146 S147 K148 C149 F150 A151 M152 D153 L154 GLU HIS HIS HIS HIS HIS HIS

- Molecule 2: DNA-directed RNA polymerase II subunit RPB1

Chain B: 36% 7% 57%

Y155 S156 P157 T158 S159 P160 S166 P167 S168

4.2.13 Score per residue for model 13

- Molecule 1: Protein NRD1

Chain A: 61% 29%

M1 Q2 Q3 D4 D5 F17 K18 D19 L20 K21 S22 G23 K31 I40 E43 S44 K45 R74 S82 N83 S84 N85 S86 S87 S88 N89 K90 P91 G92 T93 K115 K121 I130 W131 D132 R133 S134 G135 L136 F137 Q138 K139 S140 Y141 L142 N143 A144 I145 R146 S147 K148

C149 F150 A151 M152 D153 L154 GLU HIS HIS HIS HIS HIS HIS

- Molecule 2: DNA-directed RNA polymerase II subunit RPB1

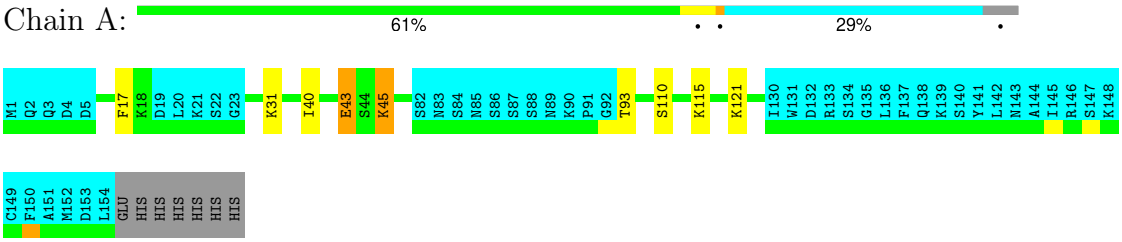
Chain B: 29% 14% 57%

Y155 S156 P157 T158 S159 Y162 T165 S166 P167 S168

4.2.14 Score per residue for model 14

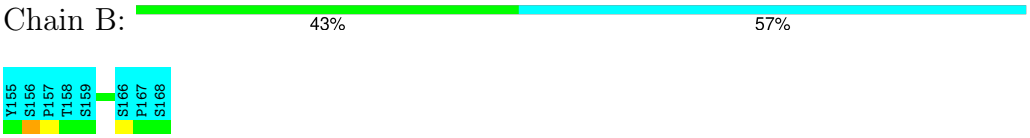
- Molecule 1: Protein NRD1

Chain A:



• Molecule 2: DNA-directed RNA polymerase II subunit RPB1

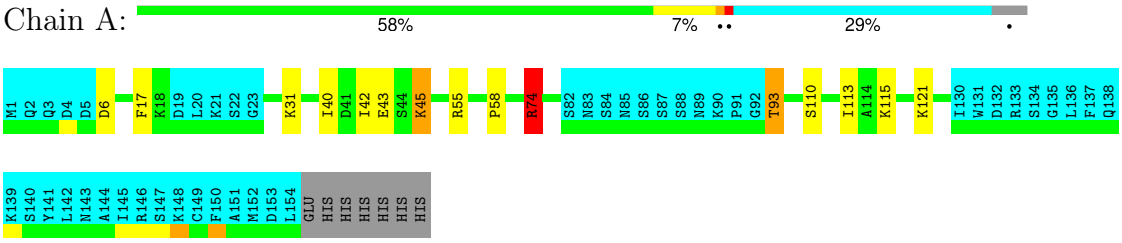
Chain B:



4.2.15 Score per residue for model 15

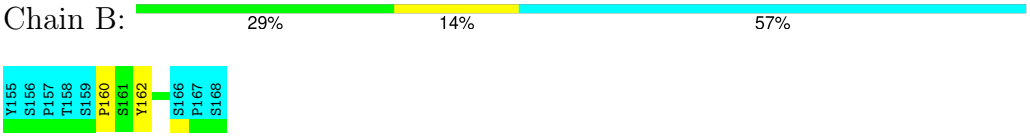
• Molecule 1: Protein NRD1

Chain A:



• Molecule 2: DNA-directed RNA polymerase II subunit RPB1

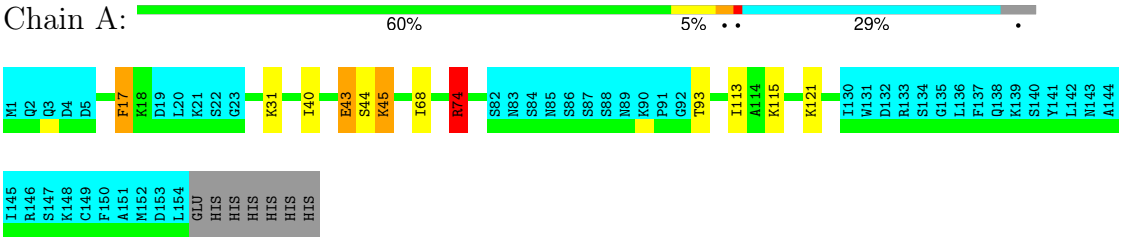
Chain B:



4.2.16 Score per residue for model 16

• Molecule 1: Protein NRD1

Chain A:



• Molecule 2: DNA-directed RNA polymerase II subunit RPB1

Chain B: 36% 7% 57%

Y155 S156 P157 T158 S159 Y162 S166 P167 S168

4.2.17 Score per residue for model 17

- Molecule 1: Protein NRD1

Chain A: 61% 29%

M1 Q2 Q3 D4 D5 F17 K18 D19 L20 K21 S22 G23 K31 I40 E43 S44 K45 R74 S82 N83 S84 N85 S86 S87 S88 N89 K90 P91 G92 T93 I113 K121 I130 W131 D132 R133 S134 G135 L136 F137 Q138 K139 S140 Y141 L142 N143 A144 I145 R146 S147 K148

C149 F150 A151 M152 D153 L154 GLU HIS HIS HIS HIS HIS HIS

- Molecule 2: DNA-directed RNA polymerase II subunit RPB1

Chain B: 36% 7% 57%

Y155 S156 P157 T158 S159 P160 S166 P167 S168

4.2.18 Score per residue for model 18

- Molecule 1: Protein NRD1

Chain A: 60% 6% 29%

M1 Q2 Q3 D4 D5 F17 K18 D19 L20 K21 S22 G23 K31 I42 E43 S44 K45 P58 R74 S82 N83 S84 N85 S86 S87 S88 N89 K90 P91 G92 T93 L108 I113 A114 K115 K121 I130 W131 D132 R133 S134 G135 L136 F137 Q138 K139 S140 Y141 L142 N143

A144 I145 R146 S147 K148 C149 F150 A151 M152 D153 L154 GLU HIS HIS HIS HIS HIS HIS

- Molecule 2: DNA-directed RNA polymerase II subunit RPB1

Chain B: 36% 7% 57%

Y155 S156 P157 T158 S159 Y162 S166 P167 S168

4.2.19 Score per residue for model 19

- Molecule 1: Protein NRD1

Chain A: 57% 9% • 29% •

M1 Q2 Q3 D4 D5 F17 K18 D19 L20 K21 S22 G23 K31 L32 T33 I40 E43 S44 K45 P58 H61 K62 S71 R74 S82 N83 S84 N85 S86 S87 S88 N89 K90 P91 G92 T93 I113 A114 K115 K121 I130 W131 D132 R133 S134 G135 L136 F137

Q138 K139 S140 Y141 L142 N143 A144 I145 R146 S147 K148 C149 F150 A151 M152 D153 L154 GLU HIS HIS HIS HIS HIS HIS

• Molecule 2: DNA-directed RNA polymerase II subunit RPB1

Chain B: 43% 57%

Y155 S156 P157 T158 S159 S166 P167 S168

4.2.20 Score per residue for model 20

• Molecule 1: Protein NRD1

Chain A: 61% 5% • 29% •

M1 Q2 Q3 D4 D5 F17 K18 D19 L20 K21 S22 G23 K31 I40 D41 I42 E43 S44 K45 S82 N83 S84 N85 S86 S87 S88 N89 K90 P91 G92 T93 S110 I113 K121 I130 W131 D132 R133 S134 G135 L136 F137 Q138 K139 S140 Y141 L142 N143 A144 I145 R146 S147

K148 C149 F150 A151 M152 D153 L154 GLU HIS HIS HIS HIS HIS HIS HIS

• Molecule 2: DNA-directed RNA polymerase II subunit RPB1

Chain B: 36% 7% 57%

Y155 S156 P157 T158 S159 P160 S166 P167 S168

5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *torsion angle dynamics, simulated annealing.*

Of the 40 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 solution	
AMBER	refinement	
CING	geometry optimization	
CYANA	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)	working_cs.cif
Number of chemical shift lists	1
Total number of shifts	1167
Number of shifts mapped to atoms	1157
Number of unparsed shifts	0
Number of shifts with mapping errors	10
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	65%

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: SEP

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	0.45±0.00	0±0/869 (0.0± 0.0%)	0.74±0.02	0±0/1172 (0.0± 0.0%)
2	B	0.72±0.02	0±0/47 (0.0± 0.0%)	0.80±0.04	0±0/64 (0.0± 0.0%)
All	All	0.46	0/18320 (0.0%)	0.74	10/24720 (0.0%)

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(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	A	74	ARG	NE-CZ-NH1	6.60	123.60	120.30	7	9
1	A	74	ARG	NE-CZ-NH2	-5.26	117.67	120.30	19	1

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	857	874	874	5±2
2	B	45	40	40	0±0
All	All	18040	18280	18280	116

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:45:LYS:H	1:A:45:LYS:CD	0.60	2.10	6	19
1:A:74:ARG:CG	1:A:74:ARG:HH11	0.56	2.13	11	9
1:A:45:LYS:CD	1:A:45:LYS:N	0.50	2.74	2	14
1:A:109:LEU:O	1:A:113:ILE:HD12	0.49	2.08	5	1
1:A:74:ARG:HH11	1:A:74:ARG:HG3	0.48	1.69	5	8
1:A:45:LYS:HD2	1:A:45:LYS:H	0.47	1.69	4	6
1:A:57:CYS:HB3	1:A:61:HIS:CG	0.47	2.44	3	1
1:A:45:LYS:H	1:A:45:LYS:HD3	0.47	1.69	10	14
2:B:162:TYR:C	2:B:162:TYR:CD1	0.47	2.87	16	6
1:A:43:GLU:N	1:A:43:GLU:CD	0.45	2.70	18	6
1:A:40:ILE:HG22	1:A:93:THR:HG22	0.45	1.89	2	2
1:A:43:GLU:CD	1:A:43:GLU:H	0.45	2.16	18	4
1:A:43:GLU:CD	1:A:44:SER:H	0.44	2.16	19	3
1:A:74:ARG:CG	1:A:74:ARG:NH1	0.43	2.82	4	6
2:B:162:TYR:CD1	2:B:162:TYR:C	0.42	2.92	11	4
1:A:47:ILE:CG1	1:A:97:ALA:HB1	0.42	2.45	2	1
1:A:17:PHE:CE2	1:A:68:ILE:HD11	0.41	2.50	16	1
1:A:42:ILE:HA	1:A:45:LYS:HE2	0.41	1.92	18	3
1:A:43:GLU:H	1:A:43:GLU:CD	0.41	2.18	9	1
1:A:33:THR:HG21	1:A:71:SER:HB2	0.41	1.90	3	4
1:A:14:LEU:HG	1:A:53:TYR:CD2	0.40	2.51	12	1
1:A:43:GLU:CD	1:A:43:GLU:N	0.40	2.75	9	1
1:A:117:ASN:HD21	1:A:120:HIS:CD2	0.40	2.34	9	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	108/161 (67%)	100±1 (92±1%)	7±1 (6±1%)	2±1 (1±1%)	14	59
2	B	6/14 (43%)	5±1 (91±13%)	0±1 (4±9%)	0±0 (5±8%)	4	25
All	All	2280/3500 (65%)	2099 (92%)	143 (6%)	38 (2%)	13	56

All 6 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	40	ILE	18
1	A	58	PRO	6
2	B	160	PRO	6
1	A	6	ASP	4
1	A	93	THR	3
1	A	62	LYS	1

6.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all NMR entries. The Analysed column shows the number of residues for which the sidechain conformation was analysed and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	97/145 (67%)	88±1 (91±1%)	9±1 (9±1%)	13 60
2	B	6/12 (50%)	6±0 (99±4%)	0±0 (1±4%)	82 97
All	All	2060/3140 (66%)	1885 (92%)	175 (8%)	14 61

All 19 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	31	LYS	20
1	A	43	GLU	20
1	A	45	LYS	20
1	A	93	THR	20
1	A	17	PHE	19
1	A	115	LYS	17
1	A	113	ILE	16
1	A	121	LYS	16
1	A	74	ARG	9
1	A	42	ILE	5
1	A	55	ARG	4
1	A	118	GLN	2
1	A	96	HIS	1
1	A	70	ASP	1
1	A	125	ARG	1
1	A	126	MET	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
2	B	165	THR	1
1	A	108	LEU	1
1	A	61	HIS	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)

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	SEP	B	159	2	8,9,10	1.02±0.01	0±0 (0±0%)
2	SEP	B	166	2	8,9,10	1.01±0.02	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	SEP	B	159	2	7,12,14	1.41±0.29	0±0 (4±6%)
2	SEP	B	166	2	7,12,14	1.52±0.10	1±0 (12±4%)

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	SEP	B	159	2	-	0±0,6,8,10	-
2	SEP	B	166	2	-	0±0,6,8,10	-

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(^{o})	Ideal(^{o})	Models	
								Worst	Total
2	B	159	SEP	OG-CB-CA	5.27	113.28	108.14	9	6
2	B	166	SEP	OG-CB-CA	3.59	111.64	108.14	15	18

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 (i)

The completeness of assignment taking into account all chemical shift lists is 65% for the well-defined parts and 52% 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 (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	1167
Number of shifts mapped to atoms	1157
Number of unparsed shifts	0
Number of shifts with mapping errors	10
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	3

The following assigned chemical shifts were not mapped to the molecules present in the coordinate file.

- No matching atom found in the structure. All 10 occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	155	GLU	H	8.223	.	.
1	A	155	GLU	HA	4.14	.	.
1	A	155	GLU	HB2	1.75	.	.
1	A	155	GLU	HB3	1.75	.	.
1	A	155	GLU	HG2	2.062	.	.
1	A	155	GLU	HG3	1.966	.	.
1	A	155	GLU	CA	52.81	.	.
1	A	155	GLU	CB	27.55	.	.
1	A	155	GLU	CG	33.7	.	.
1	A	155	GLU	N	120.302	.	.

7.1.2 Chemical shift referencing (i)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	123	1.73 ± 0.14	Should be checked
$^{13}C_{\beta}$	115	2.75 ± 0.10	Should be checked
$^{13}C'$	0	—	None (insufficient data)
^{15}N	116	0.50 ± 0.35	None needed (< 0.5 ppm)

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 65%, i.e. 1019 atoms were assigned a chemical shift out of a possible 1570. 0 out of 15 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	405/568 (71%)	203/229 (89%)	103/228 (45%)	99/111 (89%)
Sidechain	614/899 (68%)	415/587 (71%)	199/283 (70%)	0/29 (0%)
Aromatic	0/103 (0%)	0/51 (0%)	0/48 (0%)	0/4 (0%)
Overall	1019/1570 (65%)	618/867 (71%)	302/559 (54%)	99/144 (69%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 52%, i.e. 1157 atoms were assigned a chemical shift out of a possible 2233. 0 out of 19 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

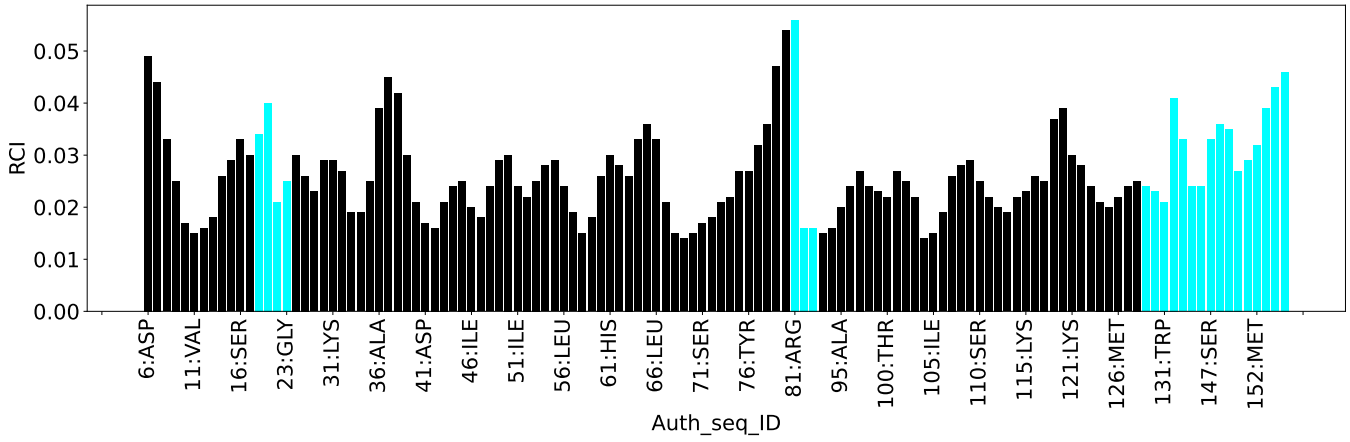
	Total	^{1}H	^{13}C	^{15}N
Backbone	473/825 (57%)	236/333 (71%)	122/332 (37%)	115/160 (72%)
Sidechain	684/1255 (55%)	461/815 (57%)	223/394 (57%)	0/46 (0%)
Aromatic	0/153 (0%)	0/75 (0%)	0/73 (0%)	0/5 (0%)
Overall	1157/2233 (52%)	697/1223 (57%)	345/799 (43%)	115/211 (55%)

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.

List Id	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	41	ASP	CG	6.71	149.18 – 208.82	-28.9
1	A	128	LEU	CD1	14.34	16.71 – 32.55	-6.5
1	A	10	PHE	HB2	1.18	1.20 – 4.80	-5.0

7.1.5 Random Coil Index (RCI) plots (i)

The image below reports *random coil index* values for the protein chains in the structure. The height of each bar gives a probability of a given residue to be disordered, as predicted from the available chemical shifts and the amino acid sequence. A value above 0.2 is an indication of significant predicted disorder. The colour of the bar shows whether the residue is in the well-defined core (black) or in the ill-defined residue ranges (cyan), as described in section 2 on ensemble composition.If well-defined core and ill-defined regions are not identified then it is shown as gray bars.

Random coil index (RCI) for chain A: