W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

Full wwPDB NMR Structure Validation Report (i)

Jun 11, 2024 – 09:23 PM EDT

PDB ID	:	1TLH
Title	:	T4 AsiA bound to sigma70 region 4
Authors	:	Lambert, L.J.; Wei, Y.; Schirf, V.; Demeler, B.; Werner, M.H.
Deposited on	:	2004-06-09

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

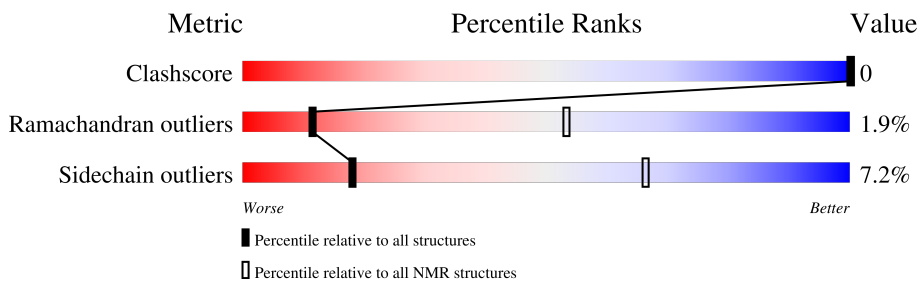|		
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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 was not calculated.

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	Percentile Ranks	Value
Clashscore		0
Ramachandran outliers		1.9%
Sidechain outliers		7.2%

Worse   *Better*

▌ Percentile relative to all structures

⧠ Percentile relative to all NMR structures

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	90	88% ... 10% ..
2	B	81	58% ... 11% ... 15% ... 16%

WORLDWIDE PDB
PROTEIN DATA BANK

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: *fewest violations*.

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:3-A:90, B:546-B:601 (144)	1.08	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 5 clusters and 3 single-model clusters were found.

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

3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 2605 atoms, of which 1312 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called 10 kDa anti-sigma factor.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	89	1473	466	737	127	142	1	0

- Molecule 2 is a protein called RNA polymerase sigma factor rpoD.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
2	B	68	1132	346	575	107	102	2	0

4 Residue-property plots ⓘ

4.1 Average score per residue in the NMR ensemble

These plots are provided for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic is the same as shown in the summary in section 1 of this report. The second graphic shows the sequence where residues are colour-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. Stretches of 2 or more consecutive residues without any outliers are shown as green connectors. Residues which are classified as ill-defined in the NMR ensemble, are shown in cyan with an underline colour-coded according to the previous scheme. Residues which were present in the experimental sample, but not modelled in the final structure are shown in grey.

• Molecule 1: 10 kDa anti-sigma factor

Chain A:

• Molecule 2: RNA polymerase sigma factor rpoD

Chain B:
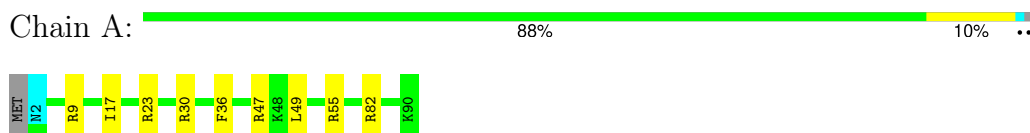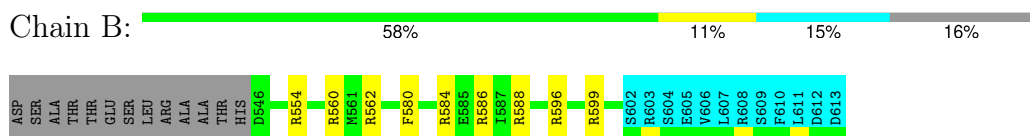
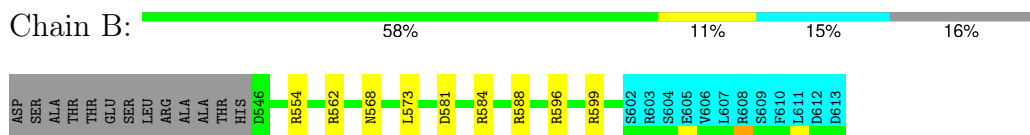## 4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1
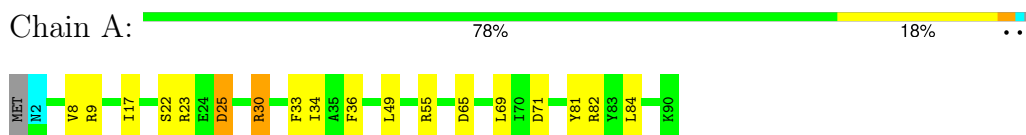
• Molecule 1: 10 kDa anti-sigma factor
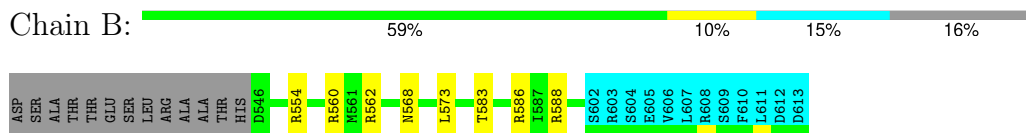
Chain A:

• Molecule 2: RNA polymerase sigma factor rpoD

Chain B:
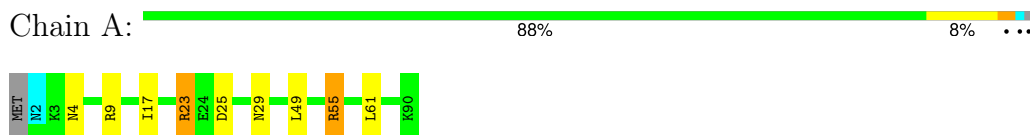
4.2.2 Score per residue for model 2

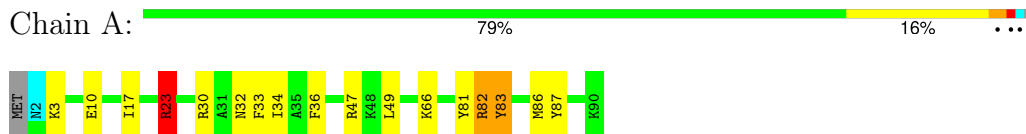• Molecule 1: 10 kDa anti-sigma factor
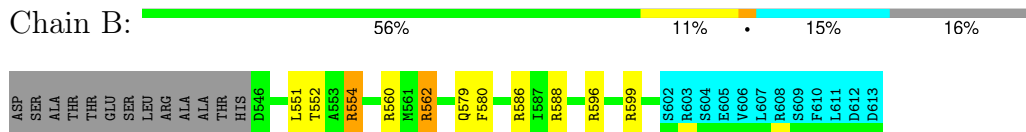
Chain A:

• Molecule 2: RNA polymerase sigma factor rpoD

Chain B:

4.2.3 Score per residue for model 3

• Molecule 1: 10 kDa anti-sigma factor

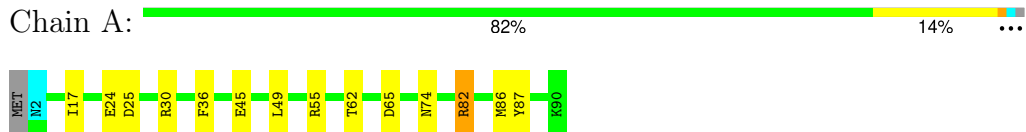Chain A:
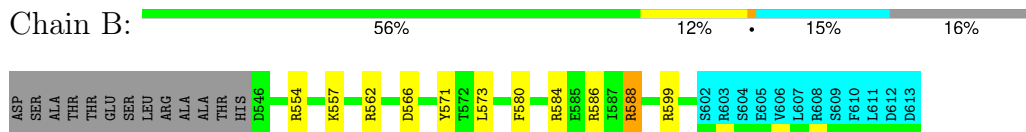

• Molecule 2: RNA polymerase sigma factor rpoD

Chain B:

4.2.4 Score per residue for model 4

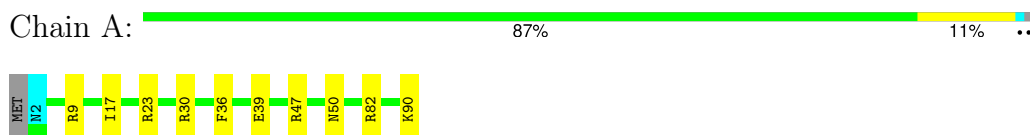• Molecule 1: 10 kDa anti-sigma factor

Chain A:

• Molecule 2: RNA polymerase sigma factor rpoD

Chain B:

4.2.5 Score per residue for model 5

• Molecule 1: 10 kDa anti-sigma factor

Chain A:

• Molecule 2: RNA polymerase sigma factor rpoD

Chain B:
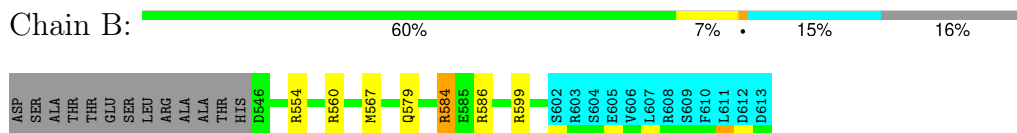
4.2.6 Score per residue for model 6

• Molecule 1: 10 kDa anti-sigma factor

Chain A:

• Molecule 2: RNA polymerase sigma factor rpoD

Chain B:
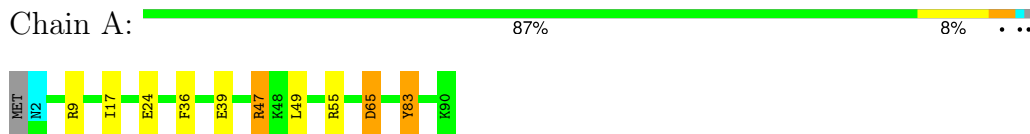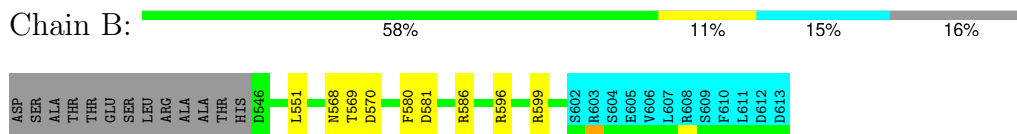
4.2.7 Score per residue for model 7

• Molecule 1: 10 kDa anti-sigma factor

Chain A:

• Molecule 2: RNA polymerase sigma factor rpoD

Chain B:
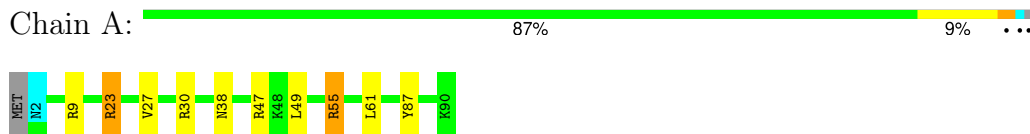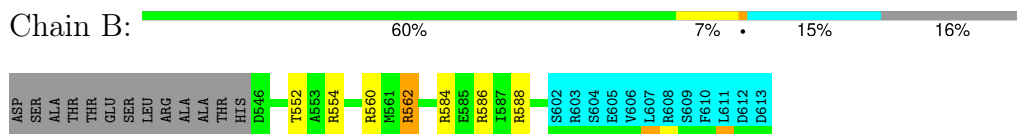
4.2.8　Score per residue for model 8

• Molecule 1: 10 kDa anti-sigma factor

Chain A:

• Molecule 2: RNA polymerase sigma factor rpoD

Chain B:
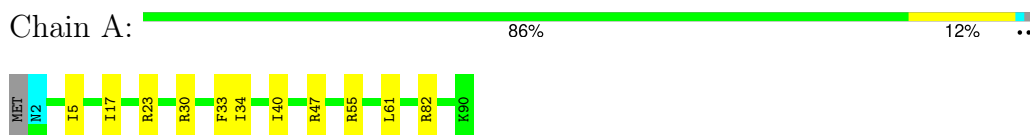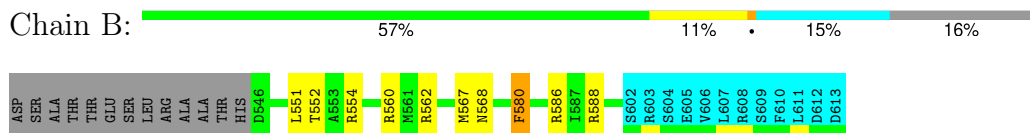
4.2.9　Score per residue for model 9 (medoid)

• Molecule 1: 10 kDa anti-sigma factor

Chain A:

• Molecule 2: RNA polymerase sigma factor rpoD

Chain B:
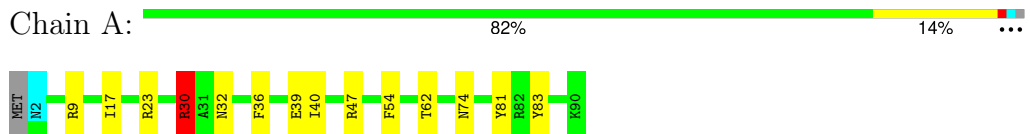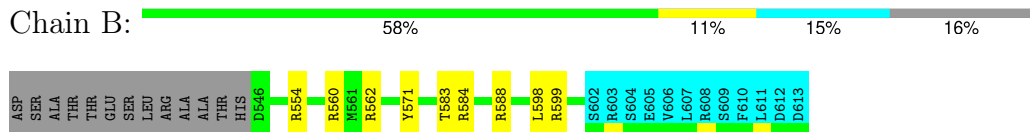
4.2.10　Score per residue for model 10

• Molecule 1: 10 kDa anti-sigma factor

Chain A:

• Molecule 2: RNA polymerase sigma factor rpoD

Chain B:

4.2.11 Score per residue for model 11

• Molecule 1: 10 kDa anti-sigma factor

Chain A:

• Molecule 2: RNA polymerase sigma factor rpoD

Chain B:
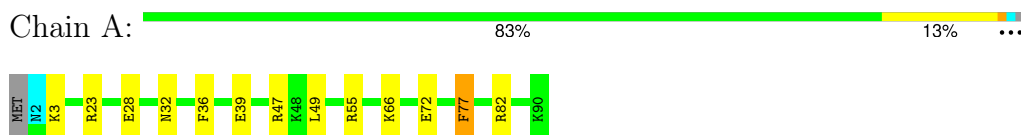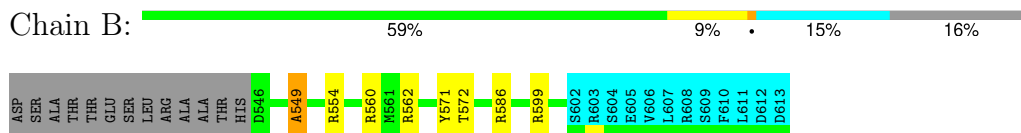
4.2.12 Score per residue for model 12

• Molecule 1: 10 kDa anti-sigma factor

Chain A:

• Molecule 2: RNA polymerase sigma factor rpoD

Chain B:
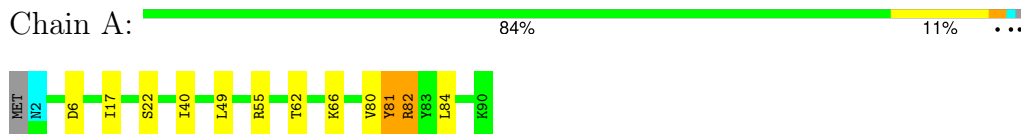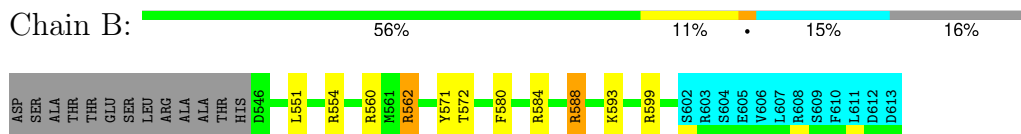
4.2.13 Score per residue for model 13

• Molecule 1: 10 kDa anti-sigma factor

Chain A:

• Molecule 2: RNA polymerase sigma factor rpoD

Chain B:
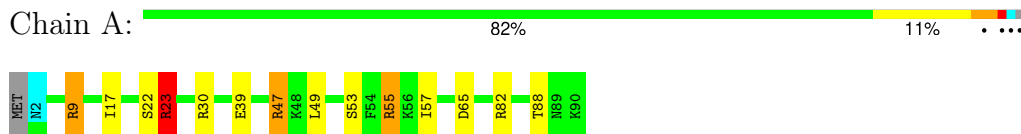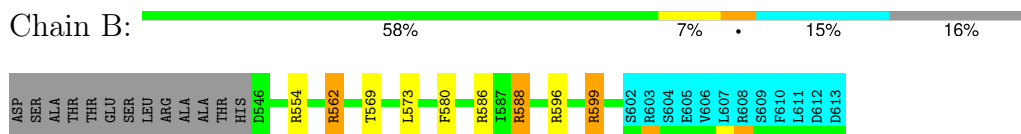
4.2.14 Score per residue for model 14

• Molecule 1: 10 kDa anti-sigma factor

Chain A:

• Molecule 2: RNA polymerase sigma factor rpoD

Chain B:
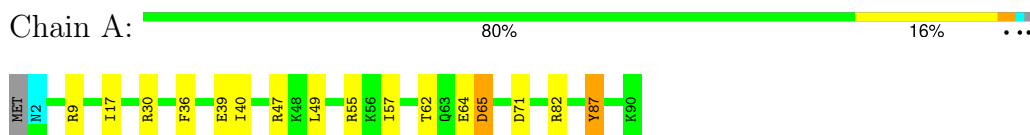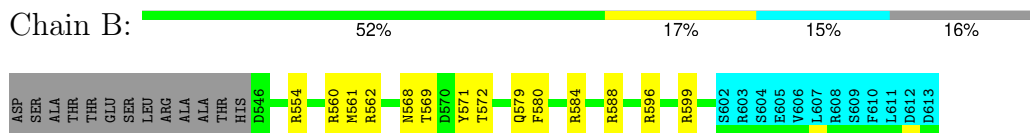
4.2.15 Score per residue for model 15

• Molecule 1: 10 kDa anti-sigma factor

Chain A:

• Molecule 2: RNA polymerase sigma factor rpoD

Chain B:
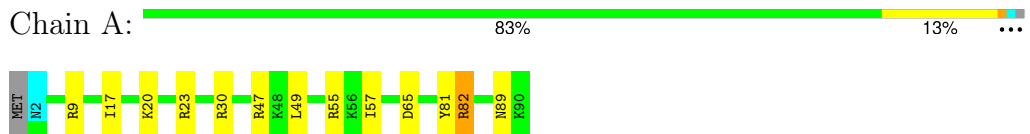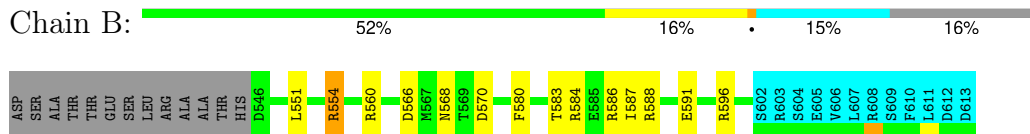
4.2.16 Score per residue for model 16

• Molecule 1: 10 kDa anti-sigma factor

Chain A:

• Molecule 2: RNA polymerase sigma factor rpoD

Chain B:
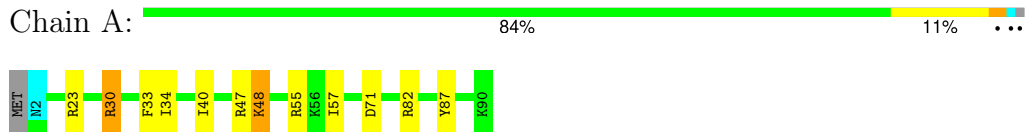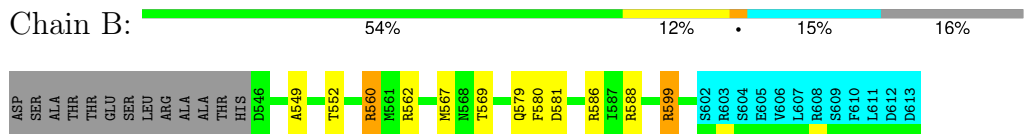
4.2.17   Score per residue for model 17

• Molecule 1: 10 kDa anti-sigma factor

Chain A:

• Molecule 2: RNA polymerase sigma factor rpoD

Chain B:
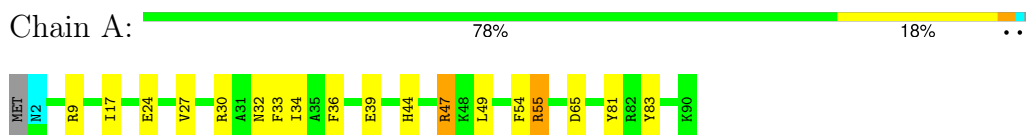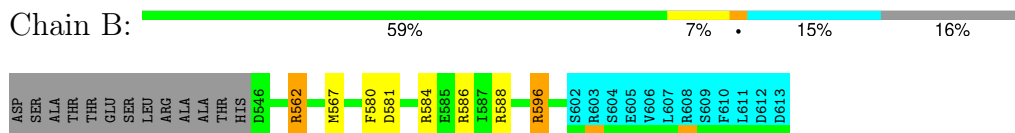
4.2.18   Score per residue for model 18

• Molecule 1: 10 kDa anti-sigma factor

Chain A:

• Molecule 2: RNA polymerase sigma factor rpoD

Chain B:
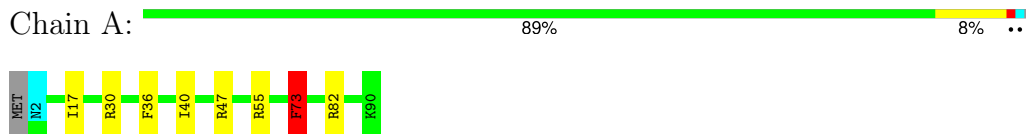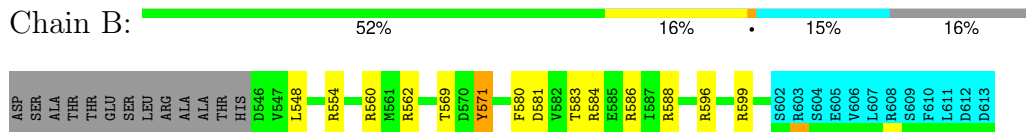
4.2.19   Score per residue for model 19

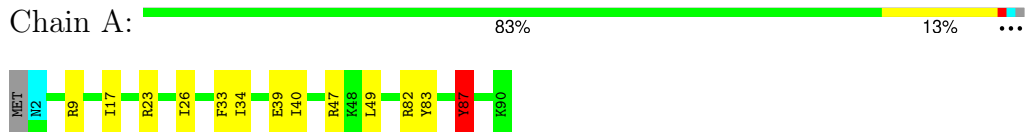• Molecule 1: 10 kDa anti-sigma factor

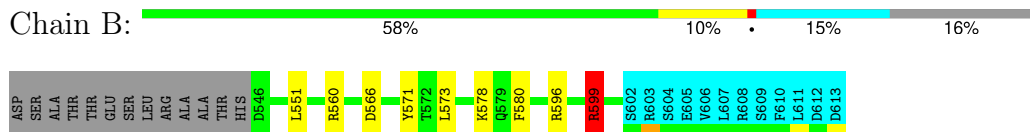Chain A:

• Molecule 2: RNA polymerase sigma factor rpoD

Chain B:
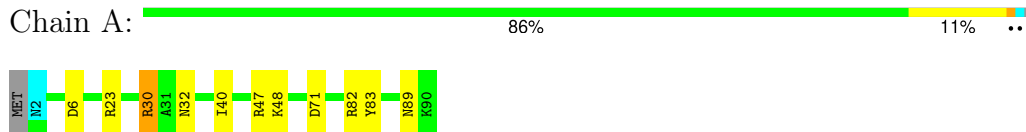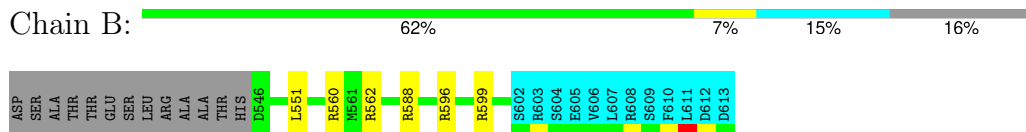
4.2.20 Score per residue for model 20

• Molecule 1: 10 kDa anti-sigma factor

Chain A:

• Molecule 2: RNA polymerase sigma factor rpoD

Chain B:
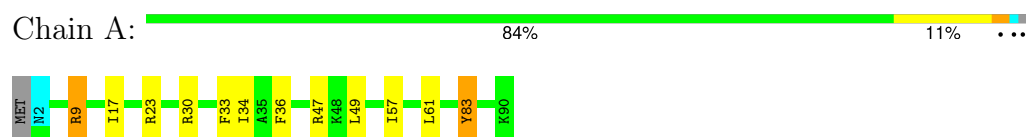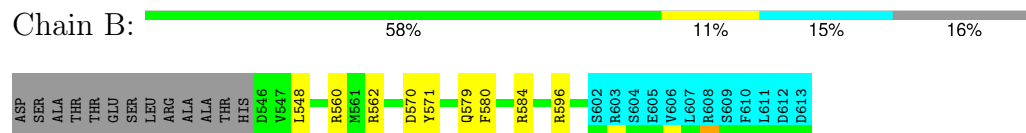
5 Refinement protocol and experimental data overview ⓘ

The models were refined using the following method: *dynamical simulated annealing/restrained energy minimization.*

Of the 100 calculated structures, 20 were deposited, based on the following criterion: *structures with the least restraint violations.*

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

Software name	Classification	Version
Xplor-NIH	refinement	2.9.3

No chemical shift data was provided.

6 Model quality ⓘ

6.1 Standard geometry ⓘ

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	0.76±0.06	0±0/737 (0.0± 0.0%)	1.20±0.06	6±2/989 (0.6± 0.2%)
2	B	0.75±0.09	0±0/464 (0.0± 0.0%)	1.31±0.11	5±2/621 (0.9± 0.3%)
All	All	0.76	0/24020 (0.0%)	1.25	217/32200 (0.7%)

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	2.5±1.0
2	B	0.0±0.0	2.2±1.0
All	All	0	93

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	584	ARG	NE-CZ-NH2	-13.72	113.44	120.30	7	4
2	B	554	ARG	NE-CZ-NH1	11.65	126.12	120.30	7	11
1	A	23	ARG	NE-CZ-NH2	-10.19	115.20	120.30	7	7
2	B	584	ARG	NE-CZ-NH1	9.90	125.25	120.30	7	8
1	A	23	ARG	NE-CZ-NH1	9.84	125.22	120.30	11	6
1	A	55	ARG	NE-CZ-NH1	8.64	124.62	120.30	15	11
1	A	82	ARG	NE-CZ-NH1	8.55	124.58	120.30	19	12
2	B	588	ARG	NE-CZ-NH1	8.46	124.53	120.30	7	8
2	B	586	ARG	NE-CZ-NH1	8.39	124.50	120.30	15	11
2	B	560	ARG	NE-CZ-NH1	8.20	124.40	120.30	3	7
1	A	87	TYR	CB-CG-CD2	-7.97	116.22	121.00	14	5
1	A	87	TYR	CB-CG-CD1	7.91	125.74	121.00	14	3
2	B	562	ARG	NE-CZ-NH1	7.83	124.21	120.30	16	11

Continued on next page...

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models Worst	Total
2	B	596	ARG	NE-CZ-NH2	-7.76	116.42	120.30	10	3
1	A	47	ARG	NE-CZ-NH1	7.67	124.14	120.30	11	10
2	B	599	ARG	NE-CZ-NH1	7.51	124.05	120.30	11	11
1	A	9	ARG	NE-CZ-NH1	7.50	124.05	120.30	5	11
2	B	596	ARG	NE-CZ-NH1	7.48	124.04	120.30	18	10
1	A	47	ARG	NE-CZ-NH2	-7.48	116.56	120.30	19	6
2	B	599	ARG	NE-CZ-NH2	-7.39	116.60	120.30	5	7
1	A	30	ARG	NE-CZ-NH1	7.35	123.98	120.30	16	8
2	B	562	ARG	NE-CZ-NH2	-7.30	116.65	120.30	17	4
1	A	55	ARG	NE-CZ-NH2	-7.30	116.65	120.30	16	5
1	A	65	ASP	N-CA-CB	-7.23	97.58	110.60	6	4
2	B	586	ARG	NE-CZ-NH2	-7.20	116.70	120.30	16	2
1	A	82	ARG	NE-CZ-NH2	-6.92	116.84	120.30	10	5
1	A	30	ARG	NE-CZ-NH2	-6.79	116.91	120.30	16	4
1	A	83	TYR	CB-CG-CD2	-6.77	116.94	121.00	9	2
2	B	554	ARG	NE-CZ-NH2	-6.73	116.93	120.30	7	3
1	A	83	TYR	CB-CG-CD1	6.62	124.97	121.00	9	1
2	B	560	ARG	NE-CZ-NH2	-6.47	117.06	120.30	7	2
1	A	77	PHE	CB-CG-CD1	6.40	125.28	120.80	11	1
2	B	566	ASP	CB-CG-OD1	6.36	124.03	118.30	15	1
2	B	588	ARG	NE-CZ-NH2	-6.25	117.18	120.30	12	3
1	A	9	ARG	NE-CZ-NH2	-6.22	117.19	120.30	10	2
1	A	73	PHE	CB-CG-CD1	6.06	125.04	120.80	18	1
1	A	81	TYR	CB-CG-CD2	-6.01	117.39	121.00	2	1
2	B	549	ALA	N-CA-CB	-5.94	101.78	110.10	11	1
1	A	77	PHE	CB-CG-CD2	-5.94	116.64	120.80	11	1
1	A	22	SER	CB-CA-C	-5.86	98.97	110.10	13	1
1	A	73	PHE	CB-CG-CD2	-5.44	116.99	120.80	18	1
1	A	81	TYR	CB-CG-CD1	5.36	124.21	121.00	2	1
1	A	25	ASP	N-CA-CB	-5.06	101.49	110.60	2	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	30	ARG	Sidechain	8
2	B	562	ARG	Sidechain	8
1	A	33	PHE	Peptide	7
1	A	23	ARG	Sidechain	6
2	B	560	ARG	Sidechain	6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group	Models (Total)
2	B	571	TYR	Sidechain	6
1	A	47	ARG	Sidechain	6
1	A	82	ARG	Sidechain	5
2	B	554	ARG	Sidechain	5
1	A	83	TYR	Sidechain	5
1	A	55	ARG	Sidechain	4
2	B	580	PHE	Peptide	4
2	B	588	ARG	Sidechain	4
2	B	584	ARG	Sidechain	2
2	B	596	ARG	Sidechain	2
2	B	599	ARG	Sidechain	2
1	A	9	ARG	Sidechain	2
2	B	548	LEU	Peptide	2
2	B	586	ARG	Sidechain	1
2	B	598	LEU	Peptide	1
1	A	87	TYR	Sidechain	1
1	A	28	GLU	Peptide	1
1	A	49	LEU	Peptide	1
1	A	80	VAL	Peptide	1
1	A	64	GLU	Peptide	1
1	A	81	TYR	Sidechain	1
2	B	587	ILE	Peptide	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	728	731	731	0±0
2	B	459	480	479	0±0
All	All	23740	24220	24200	2

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 0.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:5:ILE:HD13	1:A:61:LEU:HD21	0.43	1.91	8	1
1:A:48:LYS:H	1:A:48:LYS:HD3	0.42	1.74	16	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	87/90 (97%)	76±2 (88±2%)	9±2 (10±3%)	2±1 (2±1%)	9	48
2	B	55/81 (68%)	49±2 (89±3%)	6±2 (10±4%)	1±1 (1±2%)	16	63
All	All	2840/3420 (83%)	2500 (88%)	286 (10%)	54 (2%)	11	53

All 19 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	17	ILE	16
1	A	34	ILE	7
1	A	23	ARG	4
2	B	581	ASP	4
2	B	568	ASN	3
1	A	24	GLU	3
1	A	65	ASP	3
2	B	549	ALA	2
2	B	569	THR	2
2	B	573	LEU	1
1	A	25	ASP	1
2	B	599	ARG	1
1	A	72	GLU	1
1	A	81	TYR	1
1	A	89	ASN	1
2	B	570	ASP	1
1	A	27	VAL	1
1	A	49	LEU	1
1	A	73	PHE	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	81/83 (98%)	75±2 (93±3%)	6±2 (7±3%)	18	66
2	B	48/70 (69%)	44±2 (93±4%)	4±2 (7±4%)	17	65
All	All	2580/3060 (84%)	2393 (93%)	187 (7%)	18	66

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	49	LEU	12
1	A	36	PHE	11
2	B	580	PHE	10
1	A	39	GLU	8
1	A	40	ILE	8
2	B	551	LEU	7
2	B	588	ARG	5
1	A	32	ASN	5
2	B	579	GLN	5
1	A	57	ILE	5
1	A	71	ASP	4
2	B	573	LEU	4
2	B	583	THR	4
1	A	81	TYR	4
2	B	552	THR	4
1	A	62	THR	4
2	B	567	MET	4
1	A	61	LEU	3
2	B	568	ASN	3
1	A	66	LYS	3
1	A	83	TYR	3
1	A	87	TYR	3
2	B	569	THR	3
2	B	571	TYR	3
2	B	572	THR	3
1	A	25	ASP	2
1	A	22	SER	2
1	A	65	ASP	2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	84	LEU	2
1	A	3	LYS	2
1	A	86	MET	2
1	A	74	ASN	2
2	B	566	ASP	2
2	B	570	ASP	2
1	A	30	ARG	2
1	A	54	PHE	2
2	B	560	ARG	2
2	B	599	ARG	2
1	A	6	ASP	2
1	A	48	LYS	2
1	A	4	ASN	1
1	A	29	ASN	1
1	A	8	VAL	1
1	A	69	LEU	1
1	A	10	GLU	1
1	A	23	ARG	1
1	A	45	GLU	1
2	B	557	LYS	1
2	B	584	ARG	1
1	A	50	ASN	1
1	A	90	LYS	1
1	A	27	VAL	1
1	A	38	ASN	1
1	A	26	ILE	1
2	B	578	LYS	1
1	A	77	PHE	1
2	B	586	ARG	1
2	B	593	LYS	1
1	A	9	ARG	1
1	A	53	SER	1
1	A	88	THR	1
2	B	561	MET	1
1	A	20	LYS	1
2	B	591	GLU	1
1	A	44	HIS	1
1	A	47	ARG	1
1	A	73	PHE	1
2	B	581	ASP	1
1	A	89	ASN	1

6.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.6 Ligand geometry ⓘ

There are no ligands in this entry.

6.7 Other polymers ⓘ

There are no such molecules in this entry.

6.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

7 Chemical shift validation (i)

No chemical shift data were provided