Supplementary Information for

Title: Structural basis for transcription activation by the nitrate-responsive regulator NarL

Authors: Dmytro Kompaniiets[1,†], Lina He[2,3,†], Dong Wang[1], Wei Zhou[2], Yang Yang[4], Yangbo Hu[2,5,*], Bin Liu[1,*]

Affiliations:

[1]Section of Transcription & Gene Regulation, The Hormel Institute, University of Minnesota, Austin, MN 55912, USA

[2]State Key Laboratory of Virology, Wuhan Institute of Virology, Center for Biosafety Mega-Science, Chinese Academy of Sciences, Wuhan 430071, China

[3]University of Chinese Academy of Sciences, Beijing 100049, China

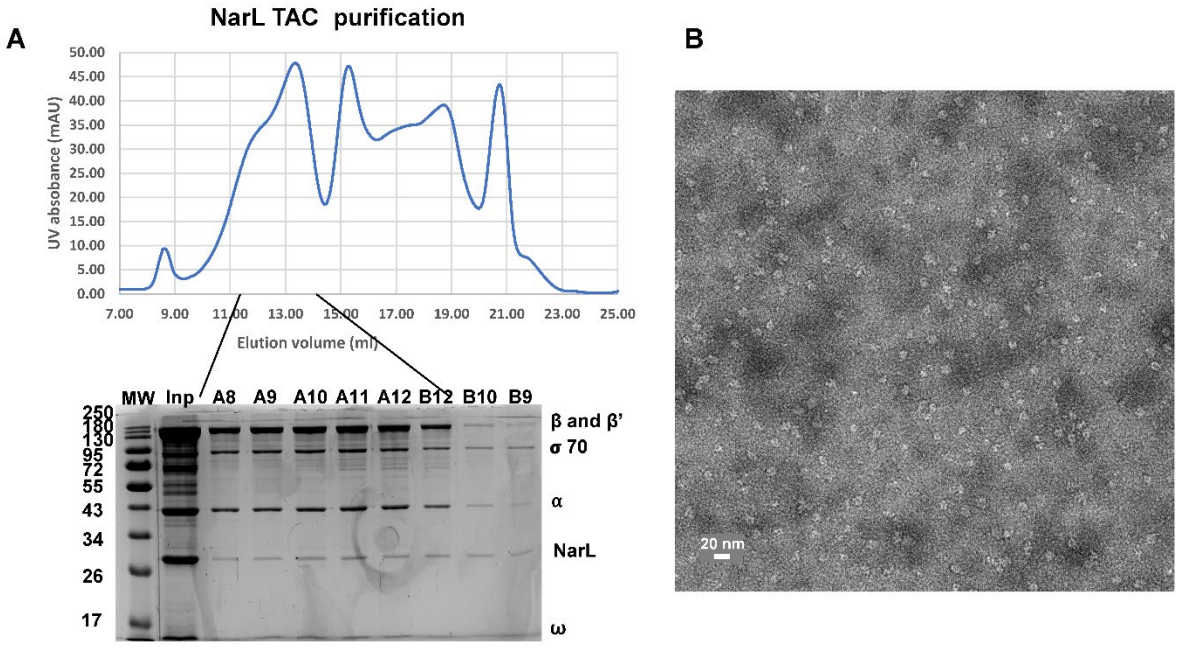
[4]Roy J. Carver Department of Biochemistry, Biophysics and Molecular Biology, Iowa State University, Ames, IA 50011, USA

[5]Hubei JiangXia Laboratory, Wuhan 430071, China

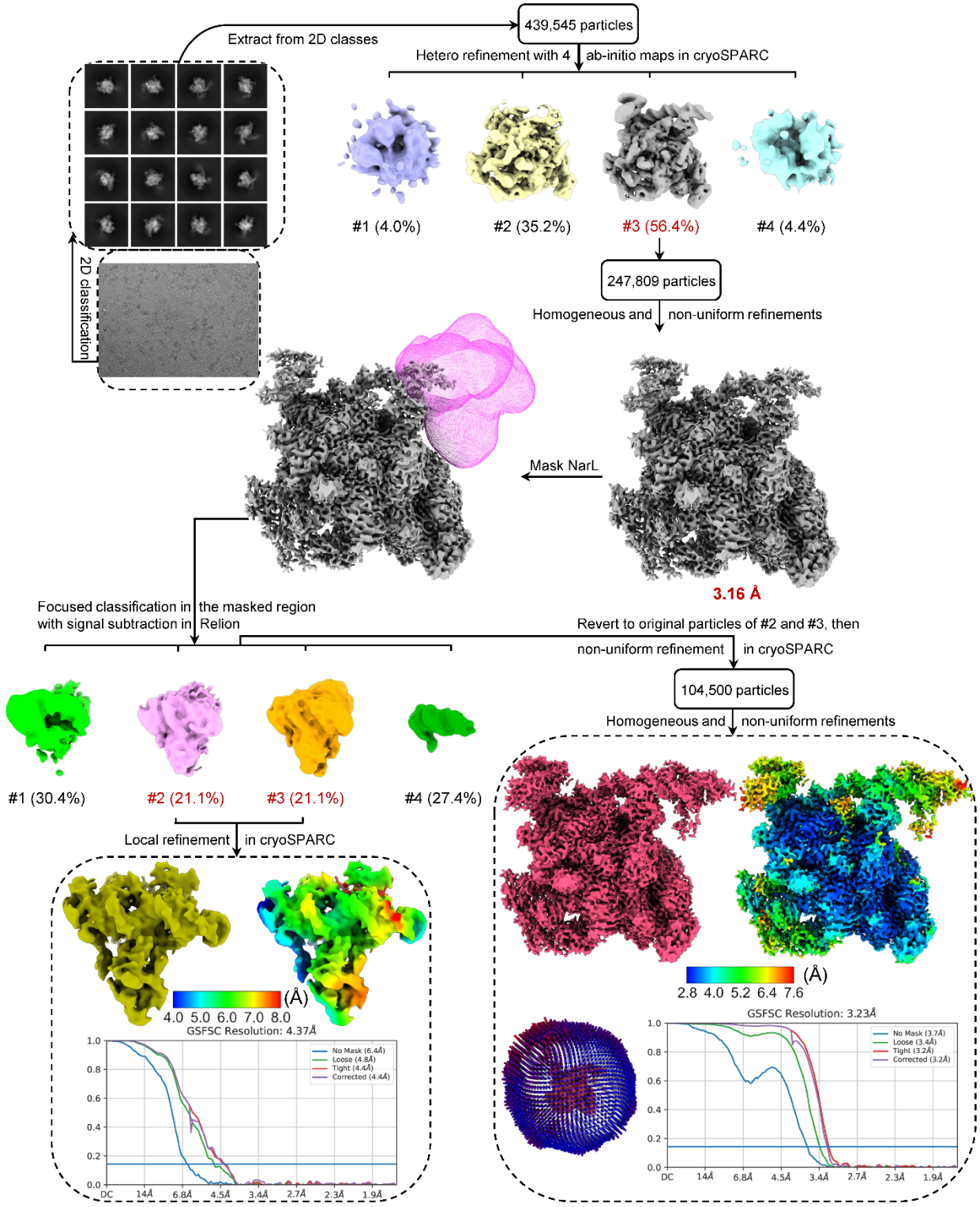
† The authors wish it to be known that, in their opinion, the first two authors should be regarded as Joint First Authors.

* To whom correspondence should be addressed. Tel: +1 507 437 9646; Email: liu00794@umn.edu

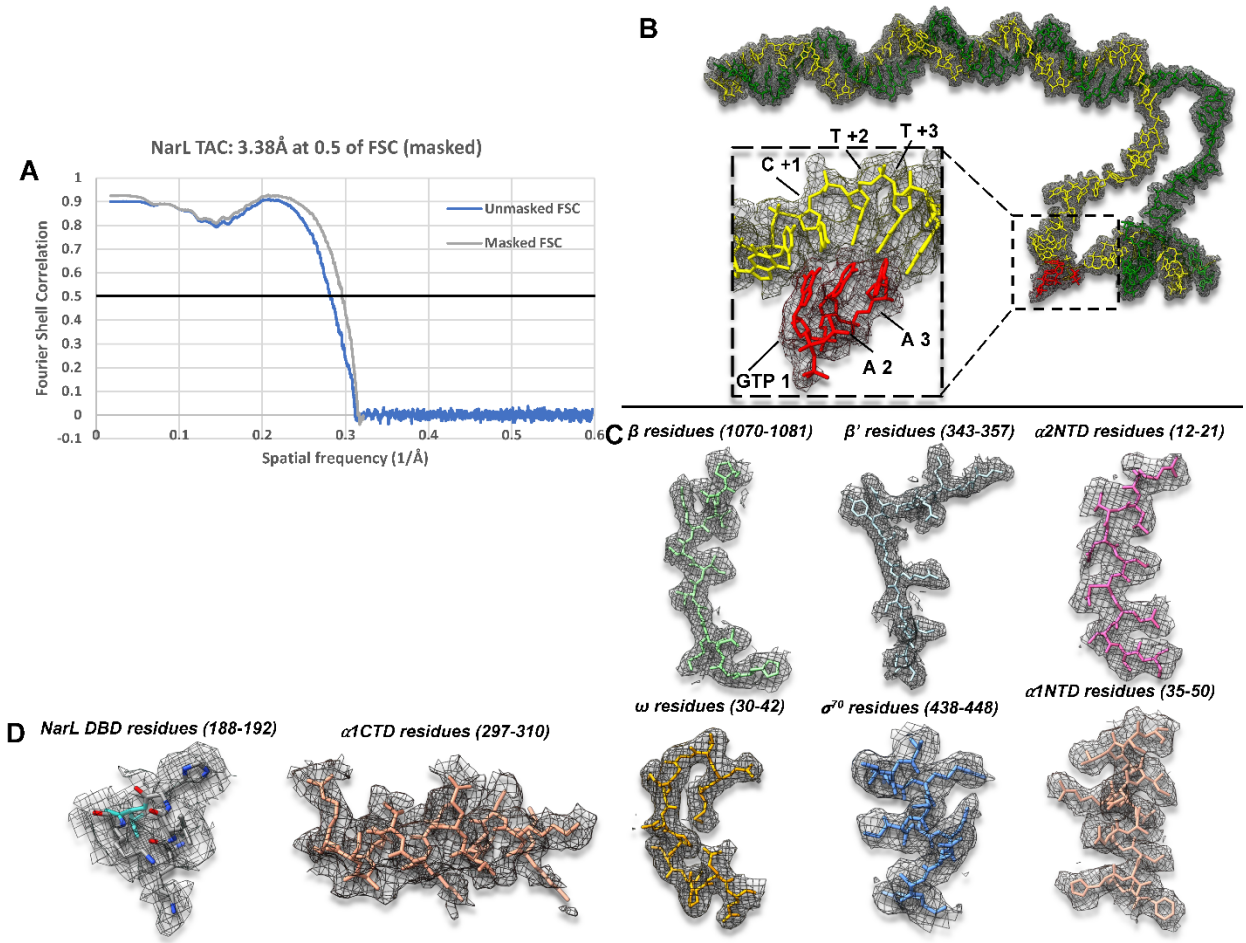
Correspondence may also be addressed to Yangbo Hu. Tel: +86 27 8719 9354; Email: ybhu@wh.iov.cn



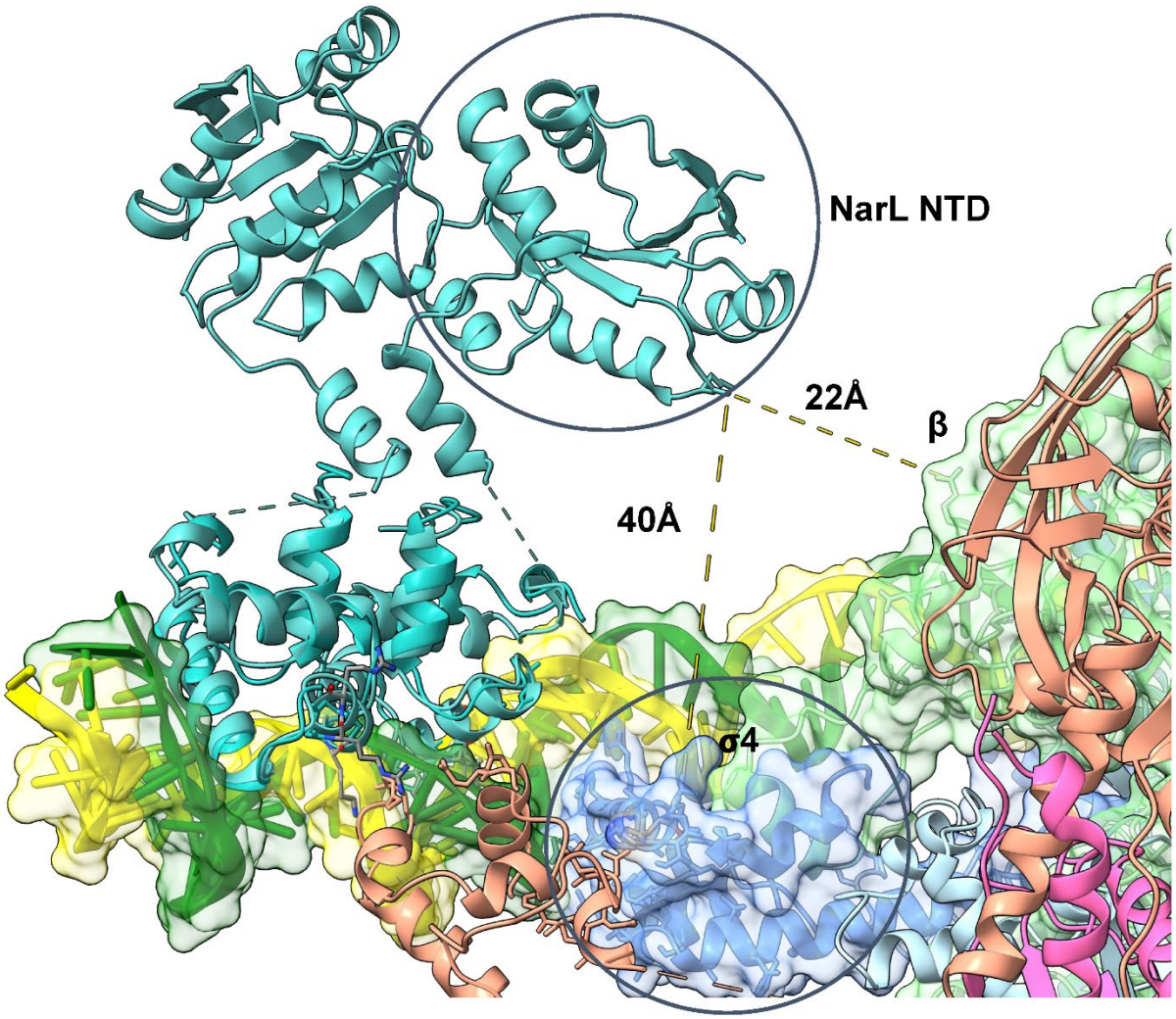
Supplementary Figure S1. Isolation of NarL-TAC. (A) Size exclusion chromatography (top) and SDS-PAGE analysis (bottom) of NarL-TAC. MW - molecular weight markers; Inp – input; A8 to B12 NarL-TAC peak elution fractions. Harvested peaks contain all the components. **(B)** Negative staining visualization of the NarL-TAC. Negative staining EM was performed to confirm sample quality, absence of aggregation and estimate general particle size.


439,545 particles
Extract from 2D classes
Hetero refinement with 4 ab-initio maps in cryoSPARC
#1 (4.0%)
#2 (35.2%)
#3 (56.4%)
#4 (4.4%)
2D classification
247,809 particles
Homogeneous and non-uniform refinements
Mask NarL
3.16 Å
Focused classification in the masked region with signal subtraction in Relion
Revert to original particles of #2 and #3, then non-uniform refinement in cryoSPARC
104,500 particles
Homogeneous and non-uniform refinements
#1 (30.4%)
#2 (21.1%)
#3 (21.1%)
#4 (27.4%)
Local refinement in cryoSPARC
(Å)
4.0 5.0 6.0 7.0 8.0
GSFSC Resolution: 4.37Å
No Mask (6.4Å)
Loose (4.8Å)
Tight (4.4Å)
Corrected (4.4Å)
(Å)
2.8 4.0 5.2 6.4 7.6
GSFSC Resolution: 3.23Å
No Mask (3.7Å)
Loose (3.4Å)
Tight (3.2Å)
Corrected (3.2Å)

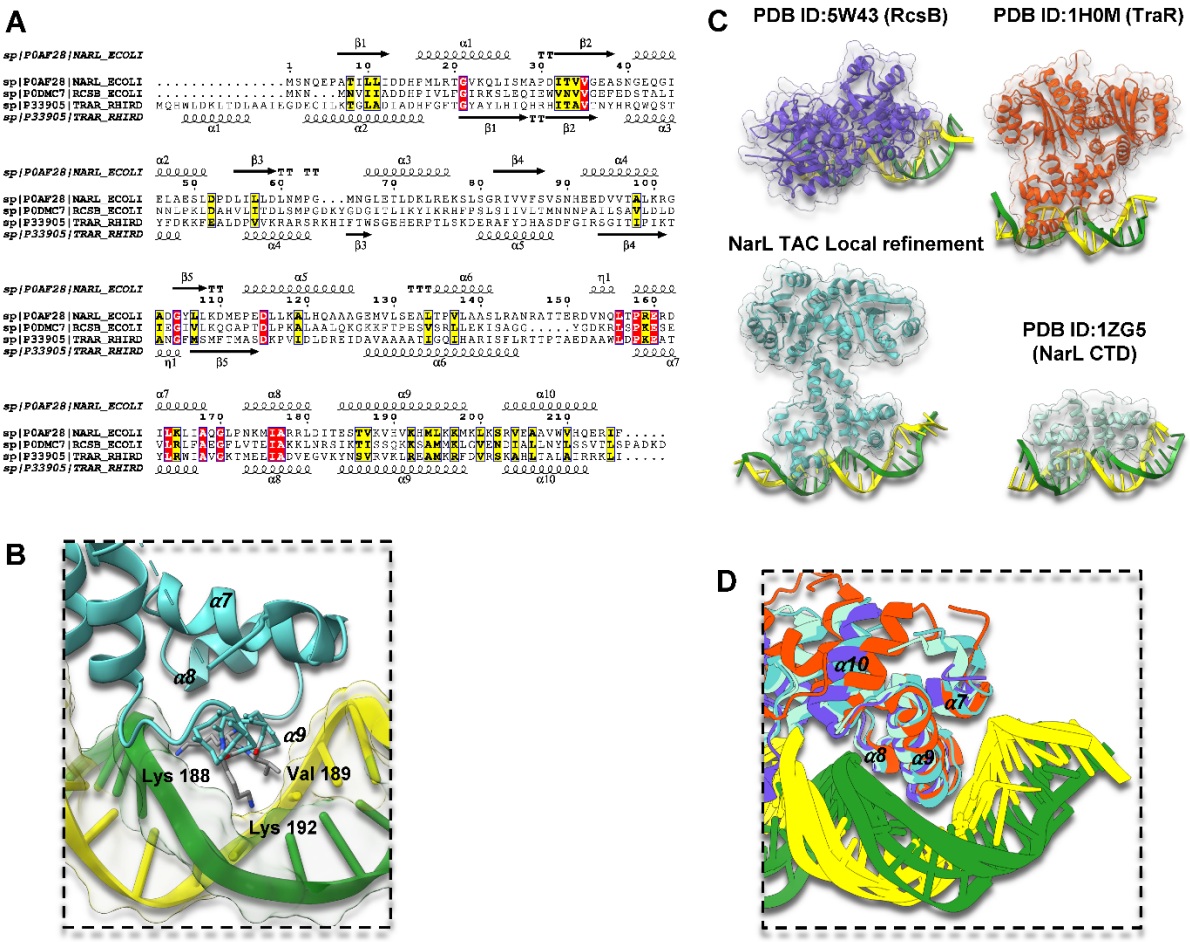
Supplementary Figure S2. Flow chart of cryo-EM image processing. Representative raw cryo-EM image and 2D classes are presented. 3D refinements using all the particles from good 3D classes generated a 3.16 Å map. Further masked 3D classification and local refinement generated a 4.37 Å map, clearly visualizing the density for the NarL/DNA region. Particles with good signal of NarL/DNA region were reverted to original particles and refined in cryoSPARC, producing a final 3.23 Å map for the NarL-TAC. Angular distribution plot, the final maps, half-map FSC curves, and accompanying local resolution illustrations are enclosed in the dashed black boxes.



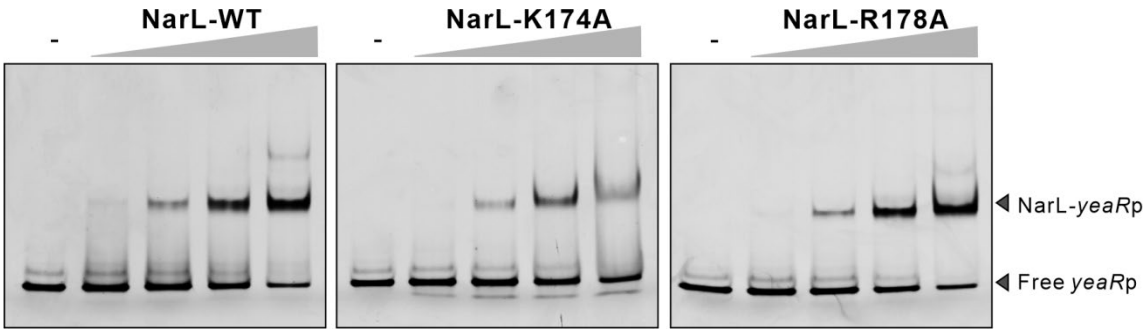
Supplementary Figure S3. Cryo-EM validations of NarL-TAC. (A) Model-to-map FSC curves. **(B)** Cryo-EM density of the promoter DNA with the highlighted RNA/DNA hybrid, confirming *de novo* RNA synthesis. Split maps were used here (color-zone 2.5 Å). **(C)** Cryo-EM densities of the represented regions of NarL-TAC. **(D)** Cryo-EM densities of the NarL/DNA/αCTD interaction regions.



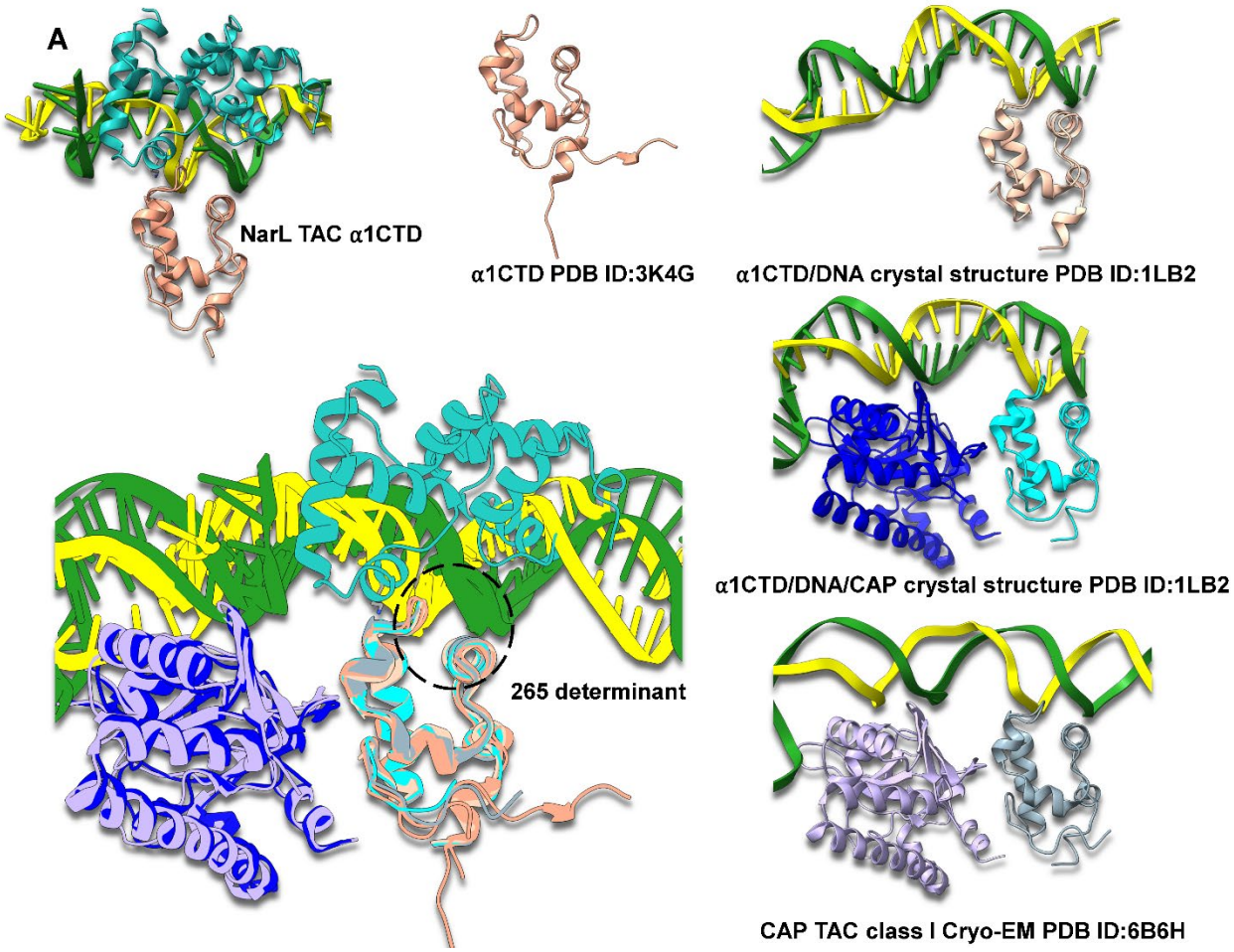
Supplementary Figure S4. The docked model of NarL NTD in the NarL-TAC complex. Docked NarL NTD is shown in the intact complex. Two distances between NarL NTD and σ4, and between NarL NTD and β subunit are labeled. The color scheme is same as Figure 1.

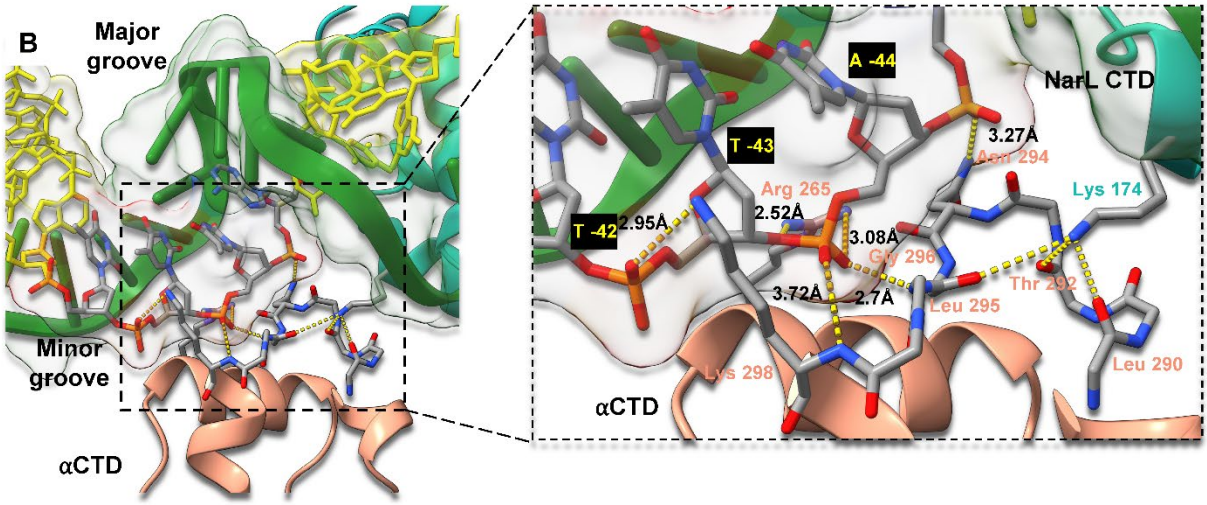


Supplementary Figure S5. Protein/DNA interaction within the FixJ/NarL family. (A) Sequence comparison of NarL, RscB, and TraR. Red box and white character – Identical residues, yellow box, and black bold character – Similar residues. **(B)** NarL CTD α9 DNA recognition helix and residues involved in sequence-specific recognition of DNA. **(C)** Protein/DNA interfaces of RscB (Blue), TraR (Red), NarL CTD (light green), and NarL full length in this Narl-TAC (light sea green). **(D)** Superposition of the four structures suggests a high structural similarity between DBD domains and the conserved DNA binding mode.

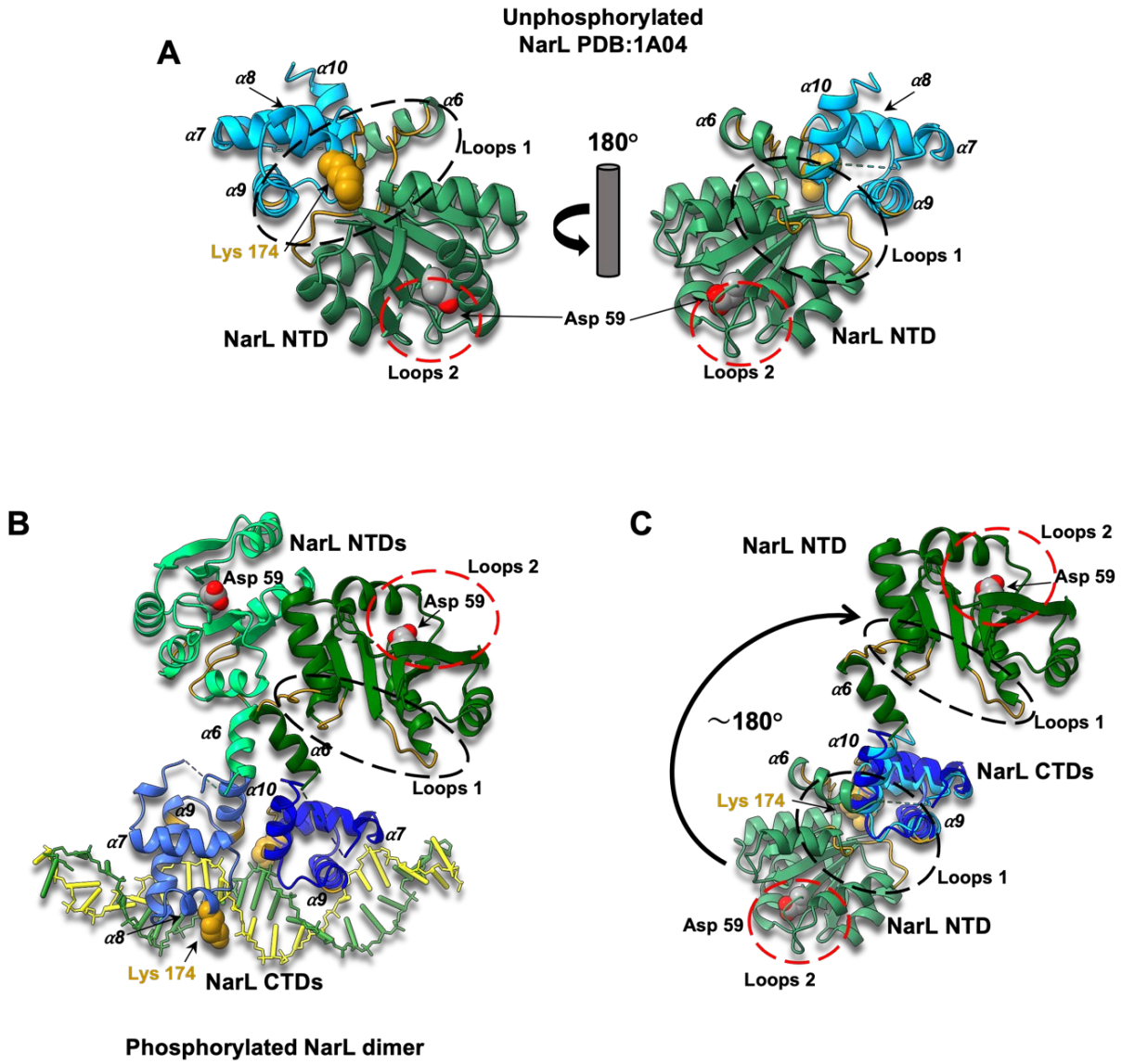


Supplementary Figure S6. Interactions between NarL and *yeaR* promoter as tested by EMSA assay. Wild-type (WT), Lys174 and Arg178 mutated NarL proteins were used at final concentrations of 0, 200, 400, 800, 1600 nM respectively in these assays.


A
NarL TAC α1CTD
α1CTD PDB ID:3K4G
α1CTD/DNA crystal structure PDB ID:1LB2
α1CTD/DNA/CAP crystal structure PDB ID:1LB2
265 determinant
CAP TAC class I Cryo-EM PDB ID:6B6H


B
Major groove
Minor groove
αCTD
A -44
T -43
T -42
NarL CTD
3.27Å
Asn 294
Arg 265
2.52Å
2.95Å
3.08Å
Lys 174
Gly 296
Thr 292
3.72Å
2.7Å
Leu 295
Lys 298
Leu 290
αCTD

Supplementary Figure S7. Comparison of αCTD/DNA interfaces from different complexes. **(A)** The α1CTD/DNA interface of this NarL-TAC; crystal structures of αCTD alone (PDB ID:3K4G); crystal structure of αCTD/DNA from CAP/αCTD/DNA structure (PDB ID:1LB2), in which the αCTD subunit doesn't interact with CAP; crystal structure of CAP/αCTD/DNA structure (PDB ID:1LB2); the interface from the cryo-EM structure of class I CAP TAC (PDB ID:6B6H); overlay of the different structures with the R265 region encircled. **(B)** View of DNA-αCTD interface in this complex. The interacted DNA and αCTD residues are shown in stick representations.



Supplementary Figure S8. Comparison of unphosphorylated NarL and phosphorylated NarL. (A) Overview of the crystal structure of NarL (PDB ID: 1A04). Phosphorylation site (Asp59) and Lys174 are represented as balls; loops on the interdomain interface are labeled dark goldenrod. **(B)** The NarL/DNA region of this NarL-TAC. **(C)** Superposition of unphosphorylated NarL (PDB ID: 1A04) and a monomer of phosphorylated NarL from this structure. The movement direction of NarL NTD is indicated by the arrow. Positions of Loops 1 and Loops 2 are highlighted

by black and red broken circles, respectively. Phosphorylation of Asp59 triggers structural changes in the loops around it (Loops 2). Such changes will further trigger changes in the interdomain interface loops (Loops 1), disrupting this interface and leading to NarL NTD movement.

Supplementary Table S1. Strains, plasmids, and oligos used in this study.

Name	Application, characters, or sequences	Source
Strains		
E. coli BL21(DE3)	Protein expression	Novagen
E. coli DH5α	Cloning construction	Novagen
E. coli ΔlacZ	*E. coli* strain MG1655 with deletion of *lacZ* gene.	(67)
E. coli narL-K174A	*E. coli ΔlacZ* expressing K174A mutated NarL protein on the genome	This study
E. coli narL-R178A	*E. coli ΔlacZ* expressing R178A mutated NarL protein on the genome	This study
Plasmids		
pVS10-RNAP	Plasmid expressing *E. coli* RNAP core enzyme, β′ subunit with C-terminal His_6-tag.	Addgene
pVS10-dαCTD	Plasmid expressing *E. coli* RNAP core enzyme with deletion of αCTD.	This study
pCas	Plasmid expressing spCas9 protein	(68)
pTargetF	Plasmid for expressing sgRNA	(68)
pET21a-NarL	pET21a expressing wild-type NarL protein with C-terminal His_6-tag	This study
pET21a-NarL-K174A	pET21a expressing K174A mutated NarL protein with C-terminal His_6-tag	This study
pET21a-NarL-R178A	pET21a expressing R178A mutated NarL protein with C-terminal His_6-tag	This study
pZT-*yeaR*p	Plasmid containing the *yeaR::lacZ* fusion	This study
Oligos (5'-3')		
Nontemplate DNA	GGTGCTAGTAACCAATAAATGGTATTTAAAATGCAAATTATCAGGCGTACCCTCTTTGCGAATTCGCGGCAGCGG	IDT, USA
Template DNA	CCGCTGCCGCGAATTCCGTTTCAGGGTACGCCTGATAATTTGCATTTTAAATACCATTTATTGGTTACTAGCACC	IDT, USA
narL-for	GGAATTCCATATGAGTAATCAGGAACCGGC	IDT, USA
narL-rev	TAACCGCTCGAGGAAAATGCGCTCCTGATGC	IDT, USA
narL target region 1	CGGGCATATTGAGATCTAAC	Sangon Biotech, China

narL target region 1	TTCGGTGATATCCAGGCGGC	Sangon Biotech, China
narL-K174A-U-F	CGTGGTGACGGTGGCGCAAAACG	Sangon Biotech, China
narL-K174A-U-R	GGCGAGCTATCATCGCATTCGGCAA ACCCTGGGCAATCAG	Sangon Biotech, China
narL-K174A-D-F	GGGTTTGCCGAATGCGATGATAGCT CGCCGGCTGGATATCACCGAAAGCA CAGT	Sangon Biotech, China
narL-K174A-D-R	CCTGTGCCATTAGTCACTGTGAC	Sangon Biotech, China
narL-R178A-U-F	CGTGGTGACGGTGGCGCAAAACG	Sangon Biotech, China
narL-R178A-U-R	GGGCAGCTATCATCTTATTCGGCAA ACCCTGGGCAATCAG	Sangon Biotech, China
narL-R178A-D-F	GGGTTTGCCGAATAAGATGATAGCT GCCCGGCTGGATATCACCGAAAGCA CAGT	Sangon Biotech, China
narL-R178A-D-R	CCTGTGCCATTAGTCACTGTGAC	Sangon Biotech, China
yeaRp-F	CTAGTCACGTCACGGCCGAGCTTCC TGATTATATCTGGTGC	Sangon Biotech, China
yeaRp-R	CCCTCCTTCGGTGAATGCGCT	Sangon Biotech, China
yeaR-qPCR-F	ACAGGGTATTGAACTGGCGG	Sangon Biotech, China
yeaR-qPCR-R	TGAATACCACAATGCGCCCT	Sangon Biotech, China
16s-qPCR-F	CGGTAATACGGAGGGTGCAA	Sangon Biotech, China
16s-qPCR-R	CGAGACTCAAGCTTGCCAGT	Sangon Biotech, China

Supplementary Table S2. Cryo-EM data collection, refinement, and validation statistics.

	NarLTAC	Local refinement on the NarL region
Data collection/processing		
Microscope	Krios	Krios
Voltage (kV)	300	300
Camera	K3	K3
Camera mode	Counting, CDS	Counting, CDS
Defocus range (μm)	−0.75 ~ −2.5	−0.75 ~ −2.5
Electron exposure (e^-/Å^2)	42	42
Magnification	130,000	130,000
Magnified pixel size (Å)	0.664	0.664
Reconstruction		
Symmetry	C1	C1
Particles refined	104,500	104,500
Resolution (Å)	3.2	4.4
Map sharpening B-factor (Å^2)	−92.9	−223.2
Access code	EMD-41856	EMD-41856
Model Statistics		
Number of residues (modeled)	3998	
Map CC	0.89	
MolProbity score	1.92	
All-atom Clashscore	10.77	
Cβ deviations	0	
Rotamer outliers	0.0%	
Ramachandran plot		
Outliers	0.13%	
Allowed	5.33%	
Favored	94.54%	
RMS deviations		
Bond length	0.011	
Bond angles	0.945	
Access code	8U3B	

Supplementary Table S3. Raw data for RT-qPCR assay.

Data Set	Target	Sample	Mean Cq	Cq SEM	Expression	Grouping	Calculation Method	Relative Quantification
1-SYBR	16S	WT-1	10.07	0.08172	Reference Gene			
1-SYBR	16S	WT-2	9.91	0.13110				
1-SYBR	16S	WT-3	9.94	0.12243				
1-SYBR	16S	WT-1 KNO3	9.77	0.18404				
1-SYBR	16S	WT-2 KNO3	9.40	0.00687				
1-SYBR	16S	WT-3 KNO3	9.39	0.02759				
1-SYBR	16S	K174A-1	9.88	0.14005				
1-SYBR	16S	K174A-2	9.70	0.00408				
1-SYBR	16S	K174A-3	9.55	0.00261				
1-SYBR	16S	K174A-1 KNO3	9.62	0.06700				
1-SYBR	16S	K174A-2 KNO3	11.14	0.07434				
1-SYBR	16S	K174A-3 KNO3	9.24	0.01340				
1-SYBR	16S	R178A-1	9.96	0.09901				
1-SYBR	16S	R178A-2	10.02	0.01367				
1-SYBR	16S	R178A-3	9.99	0.06803				
1-SYBR	16S	R178A-1 KNO3	9.76	0.00488				
1-SYBR	16S	R178A-2 KNO3	9.66	0.06765				
1-SYBR	16S	R178A-3 KNO3	10.23	0.03360				
1-SYBR	yeaR	WT-1	30.90	0.02713	0.00648	Control-1	$Expression_{control-1}/Expression_{control-1}$	1
1-SYBR	yeaR	WT-2	31.24	0.30401	0.00456	Control-2	$Expression_{control-2}/Expression_{control-2}$	1
1-SYBR	yeaR	WT-3	31.45	0.35734	0.00405	Control-3	$Expression_{control-3}/Expression_{control-3}$	1
1-SYBR	yeaR	WT-1 KNO3	24.60	0.10284	0.41325	Sample-A-1	$Expression_{sample-A-1}/Expression_{control-1}$	63.794452
1-SYBR	yeaR	WT-2 KNO3	22.80	0.13485	0.49687	Sample-A-2	$Expression_{sample-A-2}/Expression_{control-2}$	109.05146
1-SYBR	yeaR	WT-3 KNO3	25.36	0.24506	0.18798	Sample-A-3	$Expression_{sample-A-3}/Expression_{control-3}$	46.381693
1-SYBR	yeaR	K174A-1	31.30	0.16642	0.00430	Sample-B-1	$Expression_{sample-B-1}/Expression_{control-1}$	0.6640412
1-SYBR	yeaR	K174A-2	32.64	0.09296	0.00150	Sample-B-2	$Expression_{sample-B-2}/Expression_{control-2}$	0.3301628
1-SYBR	yeaR	K174A-3	32.21	0.32519	0.00181	Sample-B-3	$Expression_{sample-B-3}/Expression_{control-3}$	0.4467674
1-SYBR	yeaR	K174A-1 KNO3	28.10	0.24500	0.03297	Sample-C-1	$Expression_{sample-C-1}/Expression_{control-1}$	5.089892
1-SYBR	yeaR	K174A-2 KNO3	31.45	0.28301	0.00929	Sample-C-2	$Expression_{sample-C-2}/Expression_{control-2}$	2.0392618

1-SYBR	yeaR	K174A-3 KNO3	28.85	0.03056	0.01502	Sample-C-3	$Expression_{sample\text{-}C\text{-}3}/Expression_{control\text{-}3}$	3.7049326
1-SYBR	yeaR	R178A-1	32.77	0.37086	0.00164	Sample-D-1	$Expression_{sample\text{-}D\text{-}1}/Expression_{control\text{-}1}$	0.2530666
1-SYBR	yeaR	R178A-2	31.88	0.17903	0.00317	Sample-D-2	$Expression_{sample\text{-}D\text{-}2}/Expression_{control\text{-}2}$	0.6947316
1-SYBR	yeaR	R178A-3	32.53	0.50138	0.00198	Sample-D-3	$Expression_{sample\text{-}D\text{-}3}/Expression_{control\text{-}3}$	0.4876684
1-SYBR	yeaR	R178A-1 KNO3	25.63	0.26149	0.20057	Sample-E-1	$Expression_{sample\text{-}E\text{-}1}/Expression_{control\text{-}1}$	30.962816
1-SYBR	yeaR	R178A-2 KNO3	25.06	0.06583	0.27876	Sample-E-2	$Expression_{sample\text{-}E\text{-}2}/Expression_{control\text{-}2}$	61.181317
1-SYBR	yeaR	R178A-3 KNO3	25.35	0.20979	0.33729	Sample-E-3	$Expression_{sample\text{-}E\text{-}3}/Expression_{control\text{-}3}$	83.22159