Supplementary Table 1. Molecular replacement solution in $P4_12_12$ space group for ttEC.

	Rotation search	Translation search		
Resolution range (Å)	20.0 – 4.0	7.0 – 4.0		
Search radius (Å)	45.0			
Rotation function			**Correlation coefficients**	**R-factor (%)**
1st peak (σ units)	14.7	**1st peak**	33.5	45.4
		2nd peak	23.4	48.5
2nd peak (σ units)	9.7	**1st peak**	10.5	52.0

Supplementary Table 2. Effect of zonal scaling on the EC refinement.

	R-free	R-free
Zonal scaling	-	+
Space group	$P4_12_12$ (no twinning)	$P4_12_12$ (no twinning)
Number of molecules	1	1
Resolution (Å)	20 – 2.5	20 – 2.5
Initial	0.498	0.441
Bulk solvent correction	0.530 (was not used below)	0.431
Rigid body (domains)	0.512 (no convergence)	0.408
B-factor refinement	- (no convergence)	0.389
Space group	N/A	$P4_1$ (twinning)
Number of molecules	N/A	2
Rigid body (two molecules)	N/A	0.335
Rigid body (domains)	N/A	0.328

Further progress obtained through the B-factor and positional refinements is not reflected in the Table.

Supplementary Table 3. Data collection and refinement statistics.

Data collection	
Space group	$P4_12_12$
Unit cell parameters (Å)	$a = b = 156.2$, $c = 499.2$
Resolution (Å)	20.0-2.5 (2.59 – 2.50)*
Reflections (Total/Unique)	516297/190044
$I/\sigma(I)$	7.1 (2.5)
R_{merge} (%)	10.0 (42.3)
Completeness (%)	90.2 (81.2)
Refinement	
Space group¶	$P4_1$
Twinning (%)	50.0
Twinning operator	h,-k,-l
Resolution (Å)	20.0–2.5 (2.59 – 2.50)
Reflections used	368781
R_{factor} (%)	23.7 (33.0)
R_{free} (%)	26.8 (33.4)
Overall B-factor/RMSD ($Å^2$)	73.5/1.6
Cross-validated sigma-A coordinate error (Å)	0.42
Number of protein atoms	46990
Number of nucleic acid atoms	2168
Number of water molecules	3563
Number of Zn^{2+} ions	4
Number of Mg^{2+} ions	2
Model quality	
RMSD bond length (Å)	0.016
RMSD bond angles (°)	1.88
RMSD improper angles (°)	1.42
Ramachadran plot	
	Number of residues (%)
Most favorable regions	84.7
Allowed region	14.8
Generously allowed regions	0.5
Disallowed region	0.0

$R_{merge}=\Sigma_{hkl}\Sigma_{j} \mid I_{j}(hkl) - <I(hkl)> \mid /\Sigma_{hkl}\Sigma_{j} <I(hkl)>$, where $I_{j}(hkl)$ and $<I(hkl)>$ are the

intensity of measurement j and the mean intensity for the reflection with indices

hkl, respectively. $R_{factor, free} = \Sigma_{hkl} ||F_{calc}(hkl)| - |F_{obs}(hkl)||/\Sigma_{hkl}|F_{obs}|$, where the crystallographic R-factor is calculated including and excluding reflections in the refinement. The free reflections constituted 5% of the total number of reflections. RMSD – root mean square deviation. $I/\sigma(I)$ – ratio of mean intensity to a mean standard deviation of intensity. *The data for the highest resolution shell are shown in brackets. ¶The refinement was first carried out in the P41212 space group. However, although the procedure of zonal scaling provided a substantially better match between the experimental and model structure factor amplitude, we were still unable to obtain an R-factor below 39% at 2.5Å resolution, whereas the EC map remained quite noisy albeit showing the clear ED in the protein and "omit", nucleic acid regions. Inability to improve the R-factor and to provide a high quality ED suggested that the data are likely affected by merohedral twinning, as was also observed in other projects in our lab[2-7]. Indeed, the calculations of the intensity statistics proposed by Yeates[8] and implemented in the CNS program[9] indicated a presence of the perfect merohedral twinning, thereby suggesting that a proper space group of the crystals is $P4_1$, rather than $P4_12_12$. At the same time, the fact that we still have been able to obtain an interpretable ED using the phases calculated in the $P4_12_12$ space group led us to conclude that, as observed previously,[2-7] merohedral twinning mimicking the $P4_12_12$ space group is likely coupled with the non-crystallographic symmetry that also closely resembles the $P4_12_12$ crystallographic symmetry operators.

Therefore to obtain the high quality of the ED we have carried out the twinning refinement in the $P4_1$ space group using the CNS program[9]. For this,

before the refinement we have expanded the crystallographic data processed in the $P4_12_12$ space group to that of $P4_1$ and have generated the two additional ttEC molecules related by the corresponding crystallographic symmetry operator. The rigid body twinning refinement using the two molecules in the $P4_1$ space group converged from the initial value of 39% to ~33% at 2.5Å resolution. The resulting $|2F_{obs} - F_{calc}|$ ED map was of substantially better quality than the one obtained for the $P4_12_12$ space group and allowed us to improve the model and to easily refine the structure to the crystallographic standards corresponding to the 2.5Å resolution data.

References

1.   Kashkina, E. et al. Elongation complexes of Thermus thermophilus RNA polymerase that possess distinct translocation conformations. Nucleic Acids Res 34, 4036-4045 (2006).

2.   Vassylyev, D. G. et al. Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 A resolution. Nature 417, 712-719 (2002).

3.   Artsimovitch, I. et al. Allosteric modulation of the RNA polymerase catalytic reaction is an essential component of transcription control by rifamycins. Cell 122, 351-363 (2005).

4.   Vassylyeva, M. N. et al. Purification, crystallization and initial crystallographic analysis of RNA polymerase holoenzyme from Thermus thermophilus. Acta Crystallogr D Biol Crystallogr 58, 1497-1500 (2002).

5. Temiakov, D. et al. Structural basis of transcription inhibition by antibiotic
    streptolydigin. Mol Cell 19, 655-666 (2005).

6. Perederina, A. et al. Regulation through the secondary channel--structural
    framework for ppGpp-DksA synergism during transcription. Cell 118, 297-
    309 (2004).

7. Symersky, J. et al. Regulation through the RNA polymerase secondary
    channel. Structural and functional variability of the coiled-coil transcription
    factors. J Biol Chem 281, 1309-1312 (2006).

8. Yeates, T. D. Detecting and overcoming crystal twinning. Methods
    Enzymol 276, 344-358 (1997).

9. Brunger, A. T. et al. Crystallography & NMR system: A new software suite
    for macromolecular structure determination. Acta Crystallogr D Biol
   Crystallogr 54, 905-921 (1998).
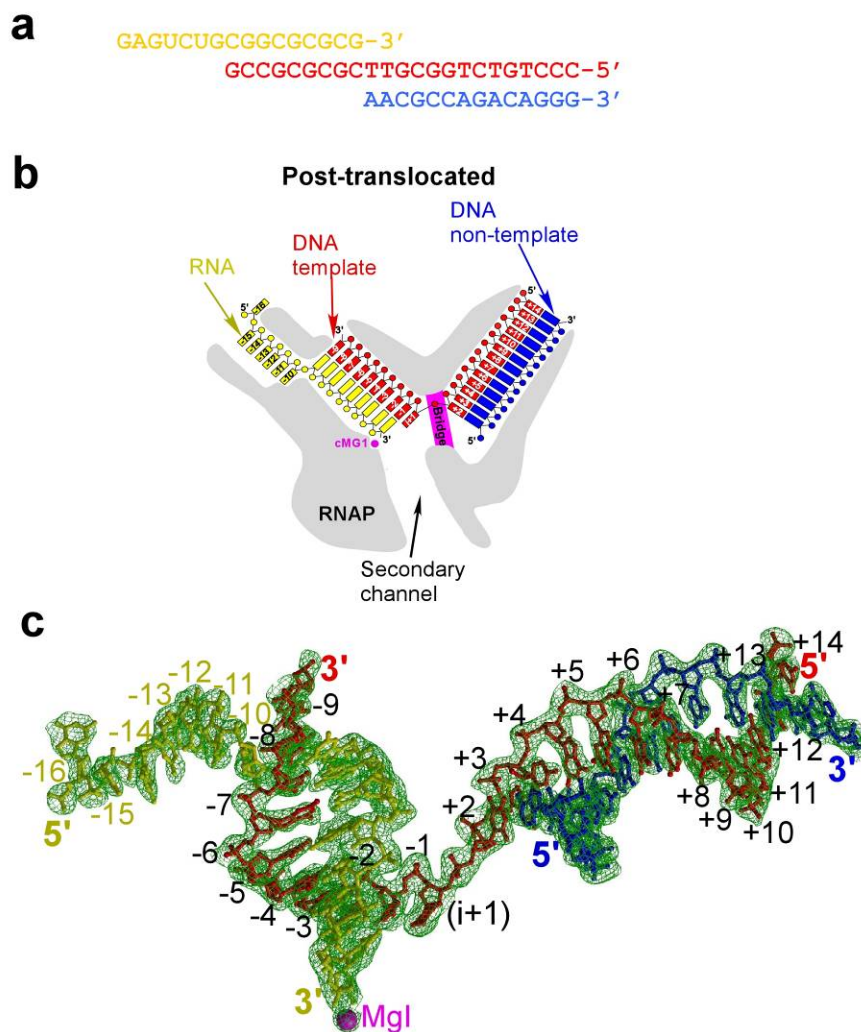
Supplementary Figure 1 The experimental details of the ttEC structure. a, The sequence of the nucleic acid scaffold used. **b,** Two-dimensional schematic projection of the ttEC structure. The (i+1) non-template nt was not resolved in the ED map and therefore is not shown. **c,** The final slow annealing (F_{obs}-F_{calc}) omit ED (green, 2.4σ level) for the nucleic acids in the ttEC.
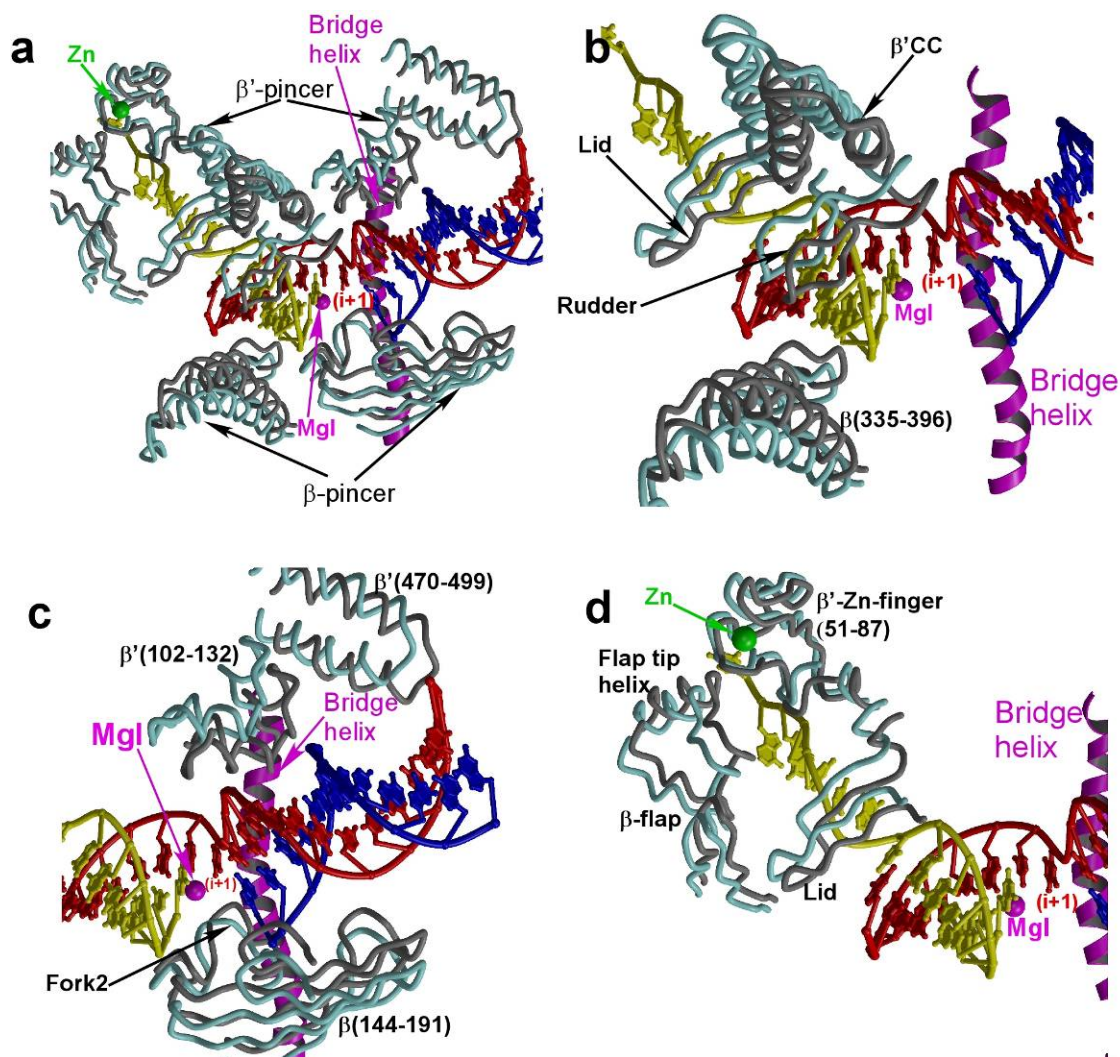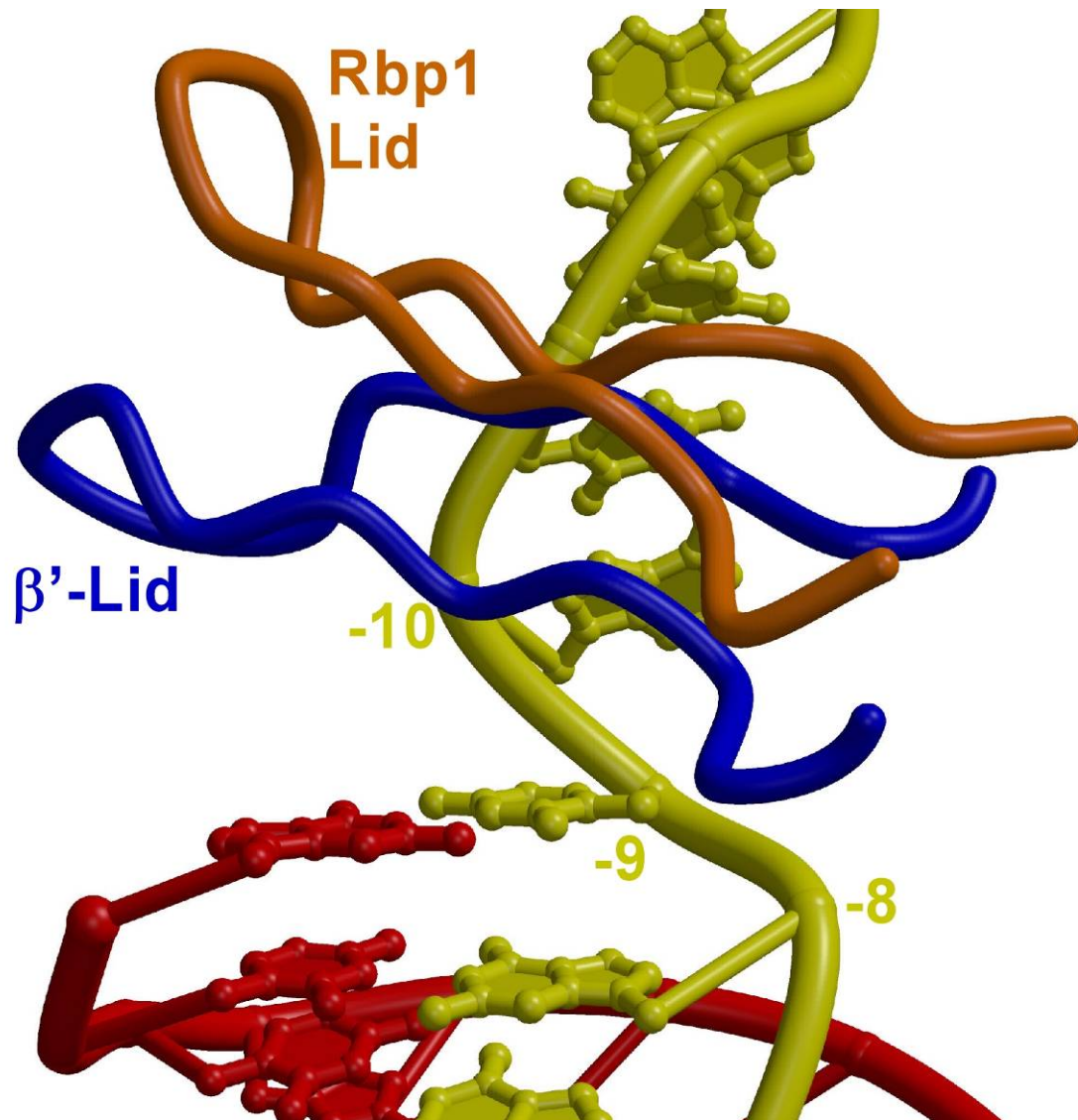
Supplementary Figure 2 Closing of the RNAP claws in the ttEC relative to the RNAP holoenzyme. a, Superposition of the structural elements forming the claws in the ttEC (gray) and holoenzyme (cyan). **b,c,d,** Close-up views of the claws in the vicinities to the RNA/DNA hybrid **(a)**, dwDNA **(b)**, and displaced single-stranded RNA transcript **(c)**. The view corresponding to that of the lower panel in **Fig. 1** is used in all the panels.
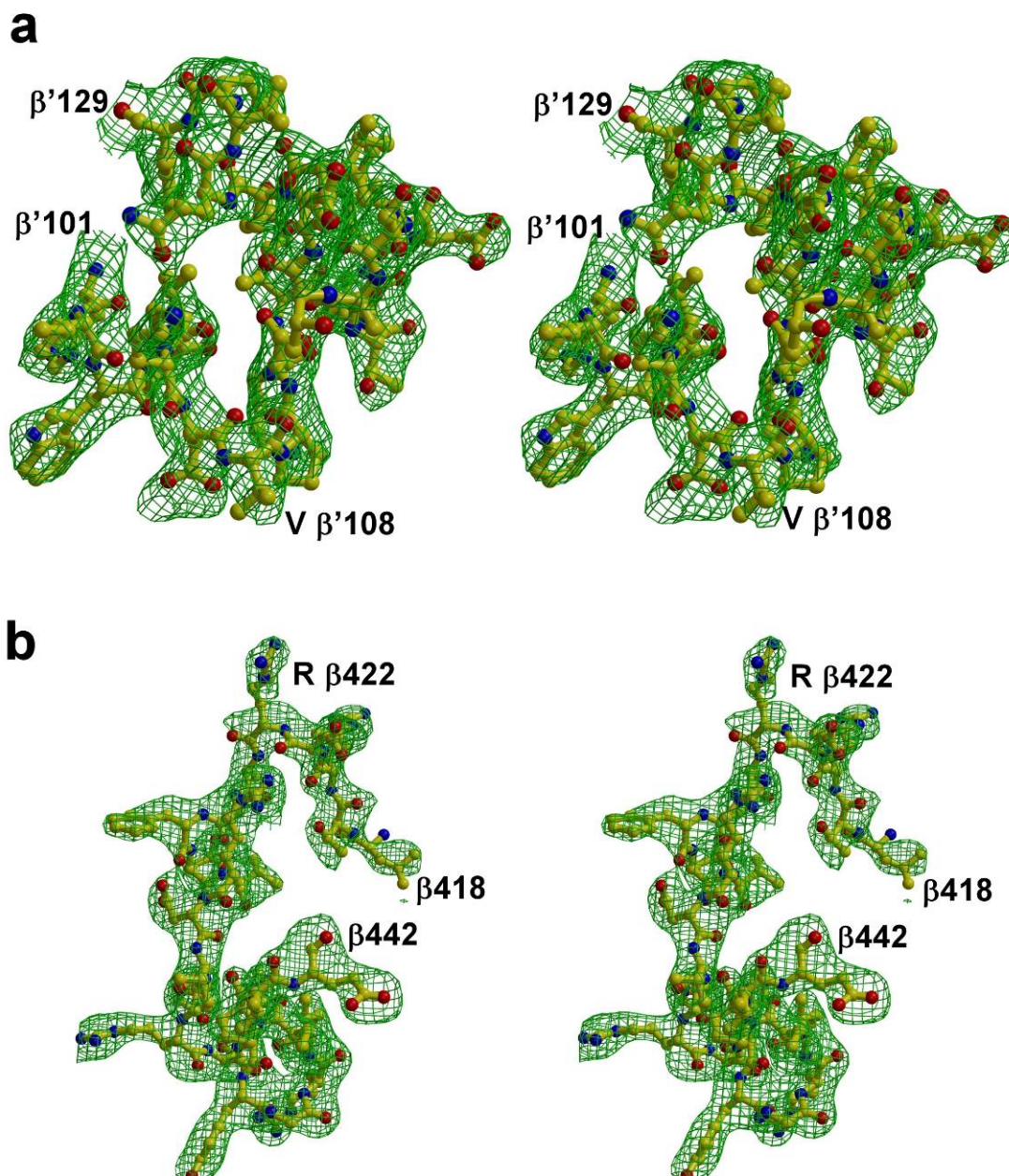
Supplementary Figure 3 The quality of the ttEC structure. a, b, The final slow annealing (F_{obs}-F_{calc}) omit ED (green, 2.4σ level) for the Val β'108 loop **(a)** and fork2 **(b)** in the ttEC.
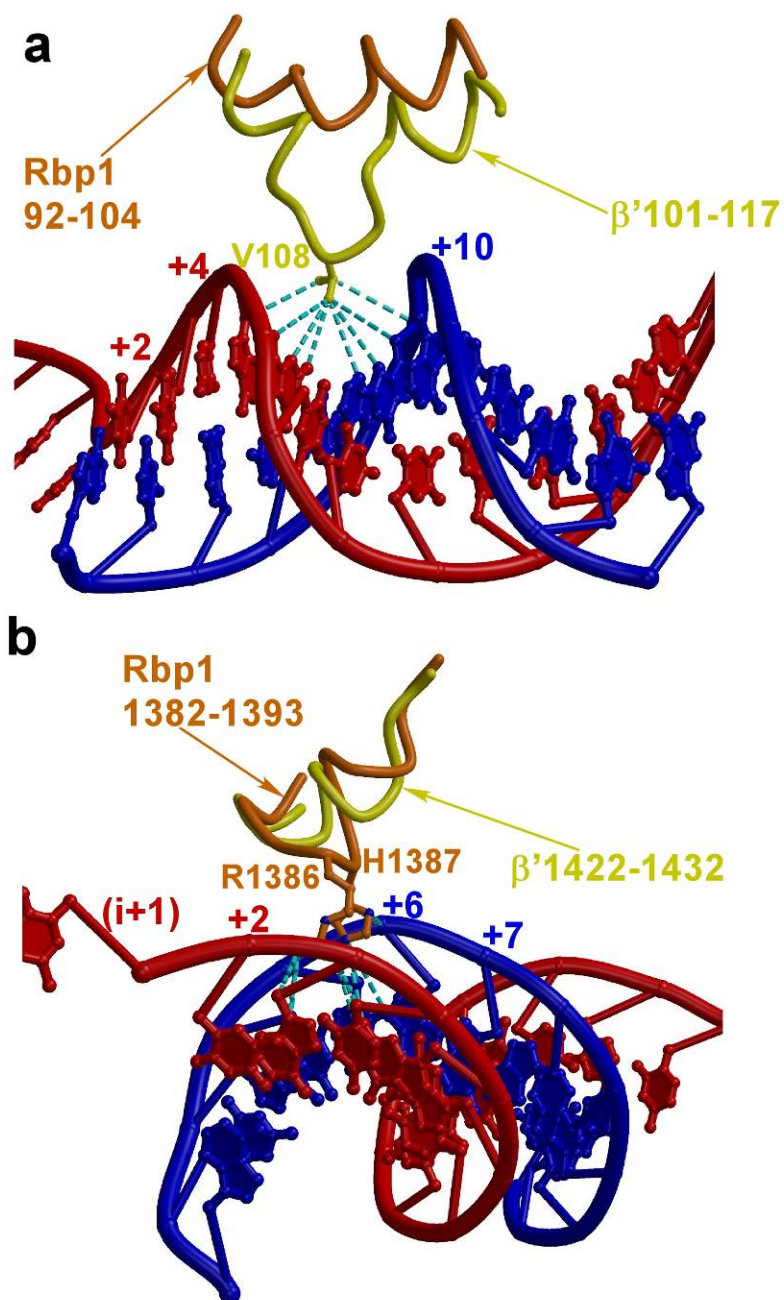
Supplementary Figure 4 Interactions of the bacterial and eukaryotic RNAPs with the dwDNA minor groove. a, b, Structural counterparts of the bacterial **(a)** and eukaryotic **(b)** structural elements penetrating the minor groove of the dwDNA are shown for comparison. The view in **(b)** corresponds to the ~90° rotation around the vertical axis of that in **Fig. 2c**.

Supplementary Figure 5 Superposition of the bacterial and eukaryotic lid loops.
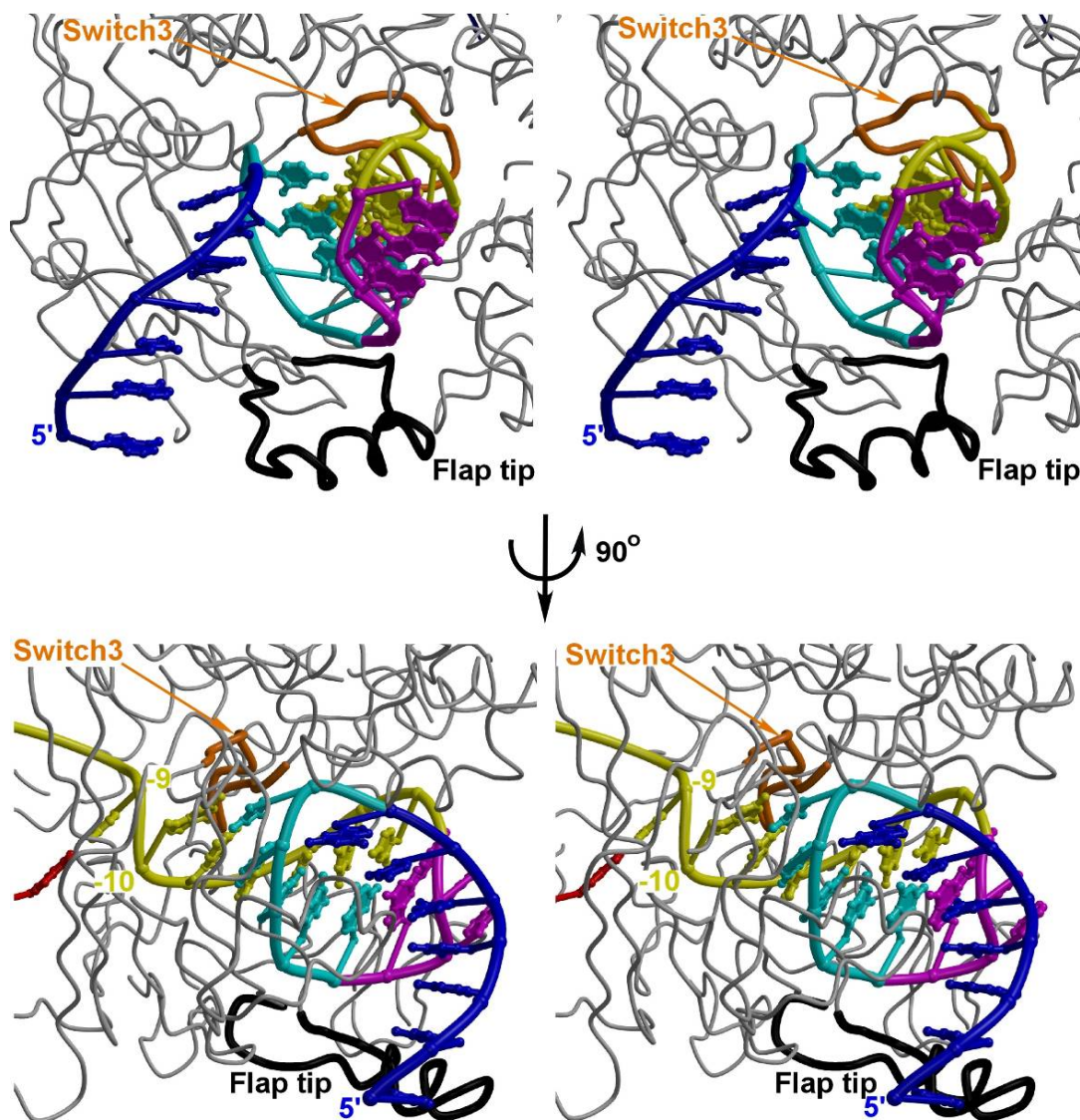
Supplementary Figure 6 Modeling of the RNA hairpin in the ttEC RNA-exit channel. The two distinct stereo views of the model are shown (the upper panel shows the view through the RNA-exit channel). The RNA duplex (5 bp) beginning at the register -11 is shown in cyan and yellow (for its 5'- and 3'-portions, respectively), while the loop intervening between the duplex strands (4 nt) is colored in magenta. The modeling of the 5' single-stranded RNA segment (blue) following the hairpin duplex clearly indicates that the RNA-exit channel can readily accommodate three RNA stands. The natural RNA hairpin (PDB ID 1M5P) was used as a template for modeling.
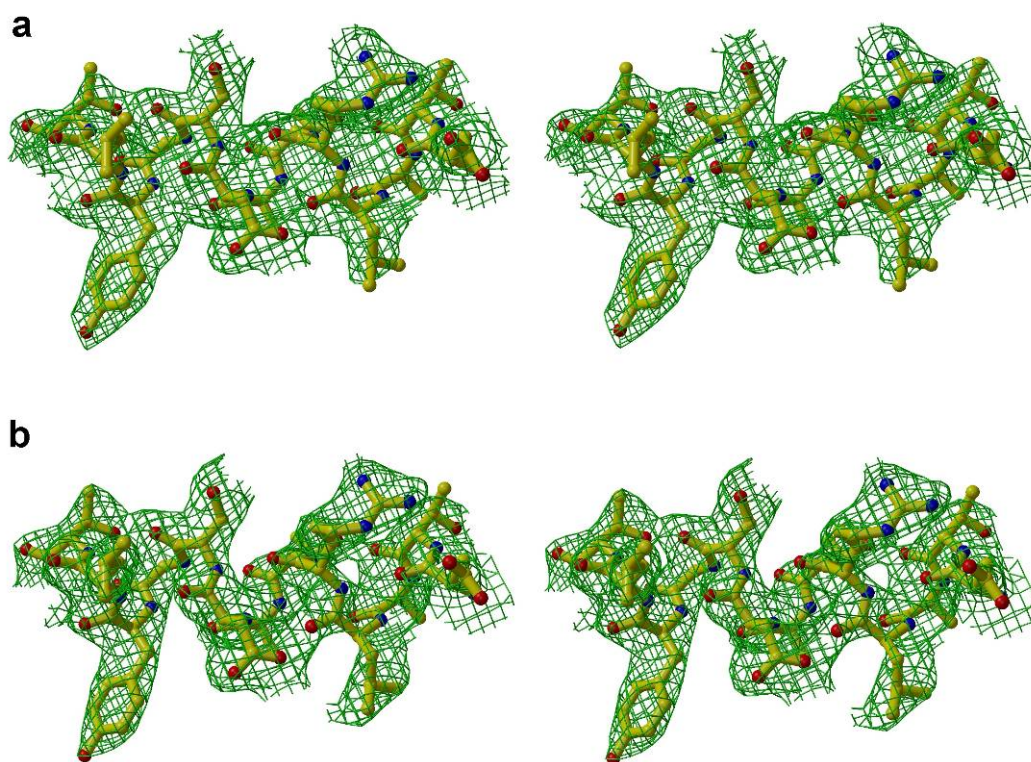
Supplementary Figure7 Effect of the zonal scaling on the ($2F_{obs} - F_{calc}$) ED in the model region. a, b, Stereo view of the region of the ($2F_{obs} - F_{calc}$) ED map (1σ level) in the vicinity to the bridge helix before **(a)** and after **(b)** the zonal scaling.
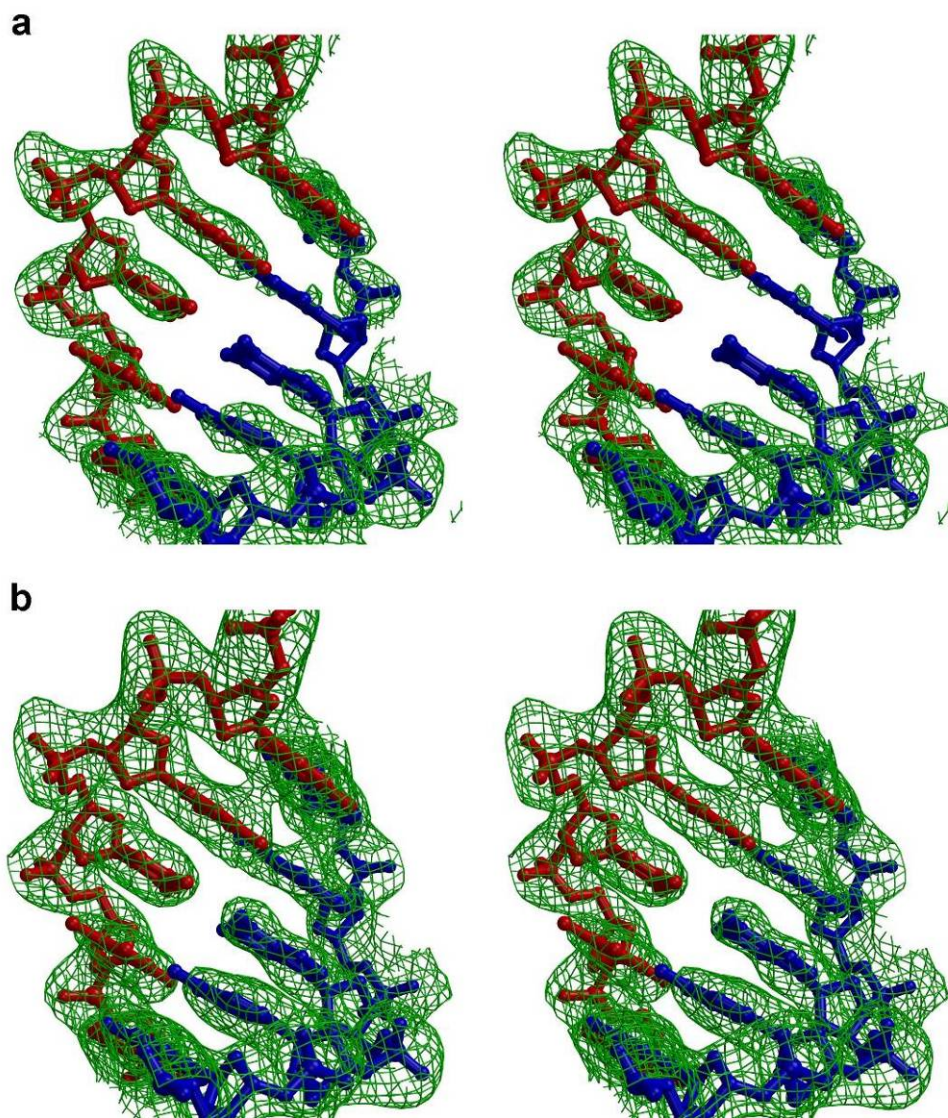
Supplementary Figure 8 Effect of the zonal scaling on the $(2F_{obs} - F_{calc})$ ED in the omit region. a, b, Stereo view of the region of the $(2F_{obs} - F_{calc})$ omit ED map (1σ level) in the vicinity to the upstream edge of the dwDNA before **(a)** and after **(b)** the zonal scaling.
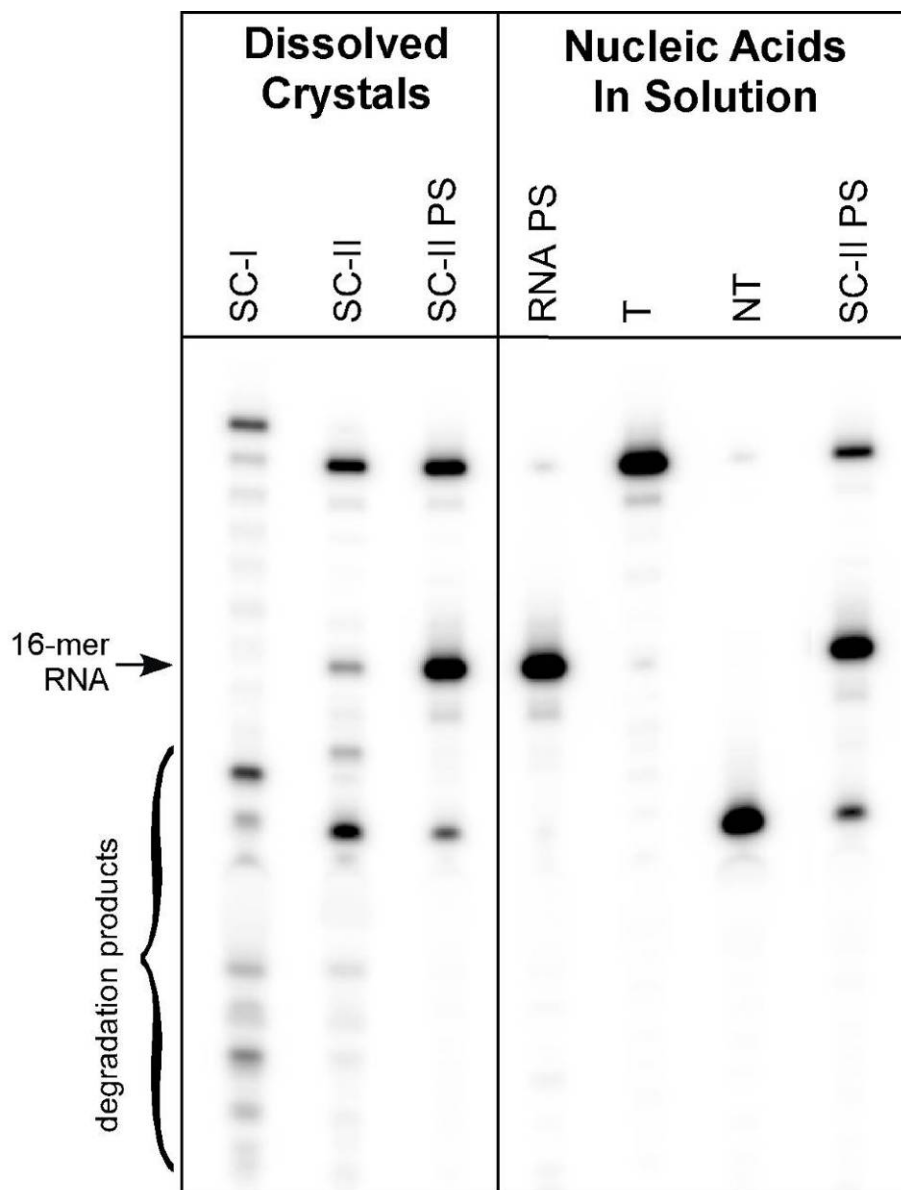
Supplementary Figure 9 Sensitivity of the RNA from different nucleic acid scaffolds used for crystallization to the nucleolytic digestion. Nucleic acid strands from the dissolved ttEC crystals, in which the RNAP core enzyme was assembled on the originally reported scaffold[1] (SC-I), and the scaffolds with the sequence used in this work (**Fig. 1a**) in which the two 3'-RNA nt were linked by the phosphodiester (SC-II) and phosphorothioate (SC-II PS) bonds, were labeled with ^{32}P γ-[ATP] and T4 polynucleotide kinase (NEB). The SC-I sample indicates complete degradation of the RNA in the crystals that were growing for ~3.5 months, while the SC-II and SC-II/thio samples obtained from the crystals that grew to the sizes suitable for diffraction experiments in ~4 weeks exhibit lesser and no RNA degradation, respectively.