



Figures and figure supplements

The universally-conserved transcription factor RfaH is recruited to a hairpin structure of the non-template DNA strand

Philipp K Zuber *et al*

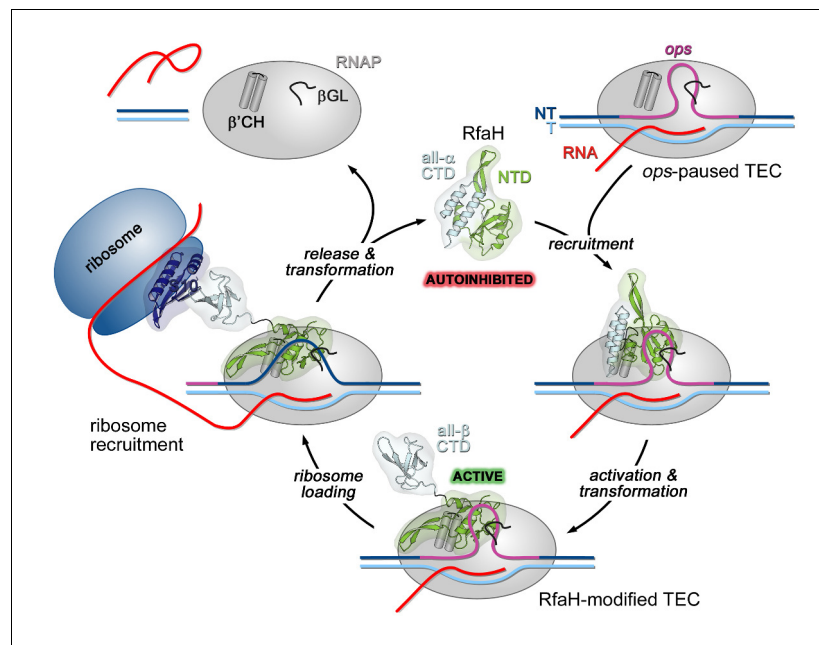


Figure 1. Life cycle of RfaH. Available experimental data demonstrate RfaH recruitment to the *ops*-paused RNAP in vitro (Artsimovitch and Landick, 2002) and in vivo (Belogurov et al., 2009) via a hairpin in the NT DNA (this work). Belogurov et al. (2007) showed that destabilization of the interdomain interface was required for RfaH switch from the autoinhibited into the active state, and proposed that the RfaH-CTD refolds into a β -barrel upon release. The RfaH-CTD refolding and interactions with S10 were demonstrated by NMR spectroscopy, and functional evidence in support of their role in ribosome recruitment in vivo was reported (Burmann et al., 2012). A hypothesis that the autoinhibited state is regained after RfaH is released from TEC at a terminator has been proposed (Tomar et al., 2013) and awaits testing. The details of RfaH:RNAP contacts that mediate initial recruitment at *ops*, the molecular mechanism of ribosome recruitment, and hypothetical coupling of transcription and translation by RfaH (Burmann et al., 2012) remain to be investigated. β' CH, β' clamp helices; β GL, β gate loop.

DOI: <https://doi.org/10.7554/eLife.36349.002>

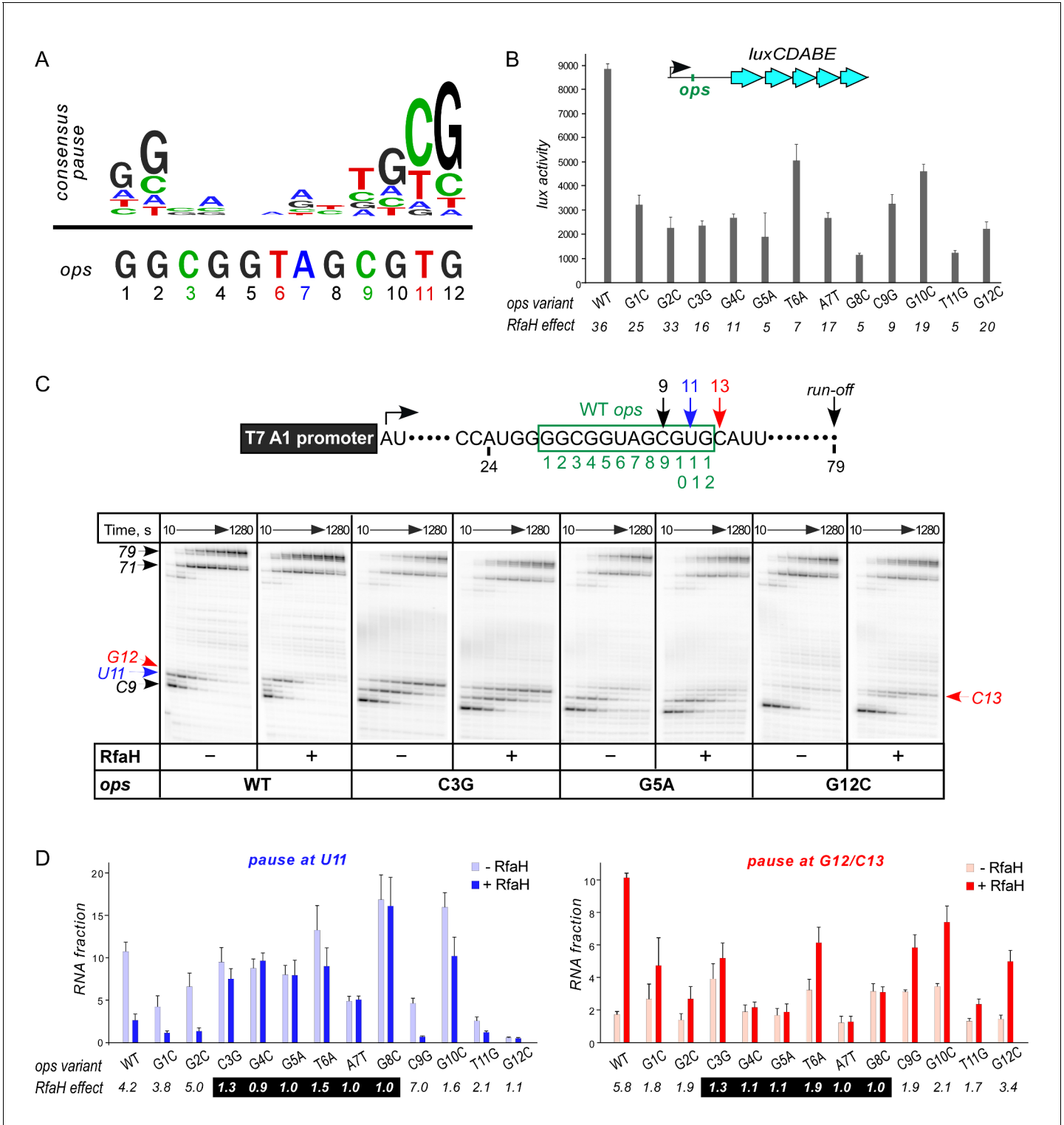


Figure 2. Contribution of individual *ops* bases to RNAP pausing and RfaH recruitment. (A) Consensus pause and *E. coli ops* sequences. (B) Expression of *luxCDABE* reporter fused to *ops* mutants in the absence and presence of RfaH determined in three independent experiments, each with three biological replicates (see source file), is presented as average ± standard deviation. Only the data obtained in the presence of RfaH are plotted; the levels of expression in the absence of RfaH are very low. RfaH effect, the ratio of *lux* activities observed with and without RfaH, is shown below each mutant. (C) In vitro analysis of *ops* mutants. Transcript generated from the T7A1 promoter on a linear DNA template is shown on top; the transcription start site (bent arrow), *ops* element (green box), and transcript end are indicated. Halted A24 TECs were formed as described in Materials and Methods

Figure 2 continued on next page

Figure 2 continued

on templates with single substitutions in the *ops* element. Elongation was restarted upon addition of NTPs and rifapentin in the absence or presence of 50 nM RfaH. Aliquots were withdrawn at 10, 20, 40, 80, 160, 320, 640, and 1280 s and analyzed on 8% denaturing gels. Positions of the paused and run-off transcripts are indicated with arrows. Pause sites within the *ops* region are numbered relative to the *ops* consensus sequence and color-coded. Results with WT, C3G, G5A, and G12C *ops* variants are shown, for all other variants see **Figure 2—figure supplement 1**. (D) Analysis of RfaH effects in vitro (from (C)). The assays were performed in triplicates. RfaH effects at U11 reflect the antipausing modification of RNAP by RfaH. RfaH effects at G12/C13 reflect RfaH binding to the NT DNA strand, which hinders RNAP escape from *ops*. Fractions of U11 RNA (left) and G12 + C13 RNAs (right) at 20 s in the absence or the presence of RfaH, presented as average \pm standard deviation from three independent experiments. RfaH effects (determined as a ratio of RNA fractions with vs. without RfaH) are shown below the variant. The core *ops* region is indicated by a black box.

DOI: <https://doi.org/10.7554/eLife.36349.003>

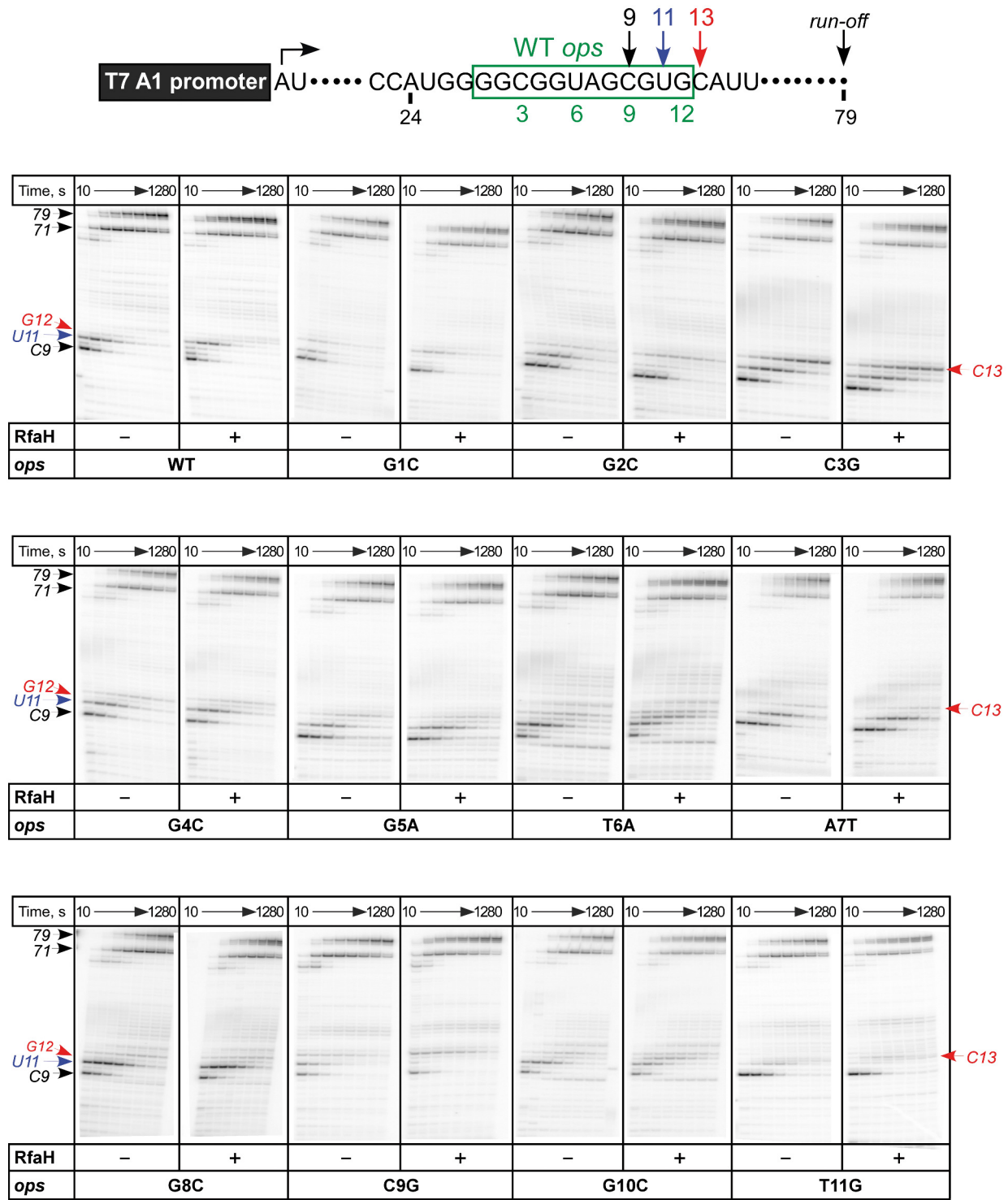


Figure 2—figure supplement 1. In vitro analysis of ops mutants. Transcript generated from the T7A1 promoter on a linear DNA template is shown on top; the transcription start site (bent arrow), ops element (green box), and transcript end are indicated. Halted A24 TECs were formed as described in Methods on templates with single substitutions in the ops element. Elongation was restarted upon addition of NTPs and rifapentin in the absence or presence of 50 nM RfaH. Aliquots were withdrawn at 10, 20, 40, 80, 160, 320, 640, and 1280 s and analyzed on 8% denaturing gels. Positions of the

Figure 2—figure supplement 1 continued on next page

Figure 2—figure supplement 1 continued

paused and run-off transcripts are indicated with arrows. Pause sites within the *ops* region are numbered relative to the *ops* consensus sequence and color-coded.

DOI: <https://doi.org/10.7554/eLife.36349.004>

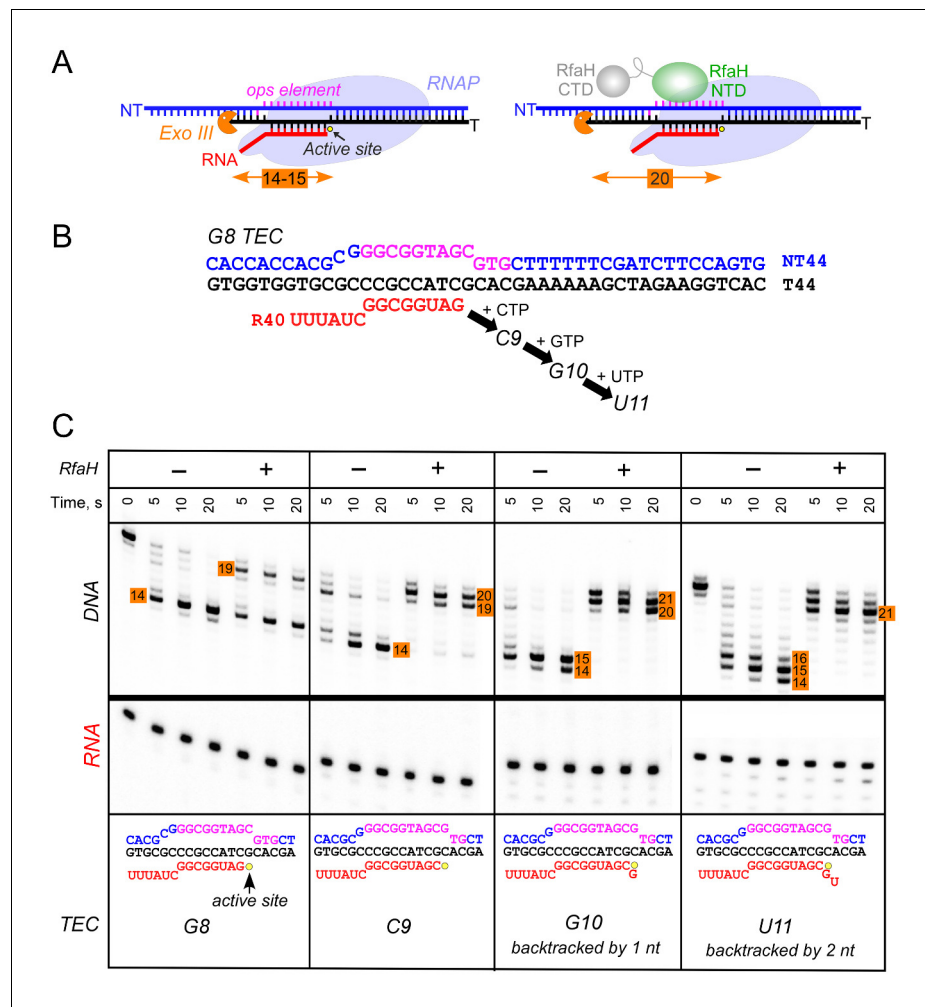


Figure 3. RfaH recruitment to RNAP transcribing through the *ops* element. (A) Schematic of Exo III footprinting of free and RfaH-bound TECs. Numbers indicate the upstream footprint boundaries relative to the RNA 3' end. (B) The G8 TEC was assembled on the scaffold, with RNA and template (T) DNA strands labeled with [γ^{32} P]-ATP and T4 polynucleotide kinase (PNK), and walked in one-nucleotide steps to C9, G10, and U11 positions in the presence of the matching NTP substrates. (C) RfaH was added to 50 nM, where indicated. Following the addition of Exo III, the reactions were quenched at indicated times (0 represents an untreated DNA control) and analyzed on a 12% urea-acrylamide (19:1) gel in 0.5X TBE. Numbers indicate the distance from the RNA 3' end. Hypothetical TEC structures are shown below. G8 and C9 complexes are predominantly post-translocated, as indicated by 14 bp protection of the upstream DNA. In G10 TEC, the pre-translocated state (15 bp protection) is observed, and in U11 an additional backtracked state (16 bp protection). Exo III may counteract backtracking; the sensitivity of the nascent RNA in G10 and U11 TECs to GreB-assisted cleavage (Nedialkov et al., 2018) was used to infer the translocation states shown in the schematics.

DOI: <https://doi.org/10.7554/eLife.36349.007>

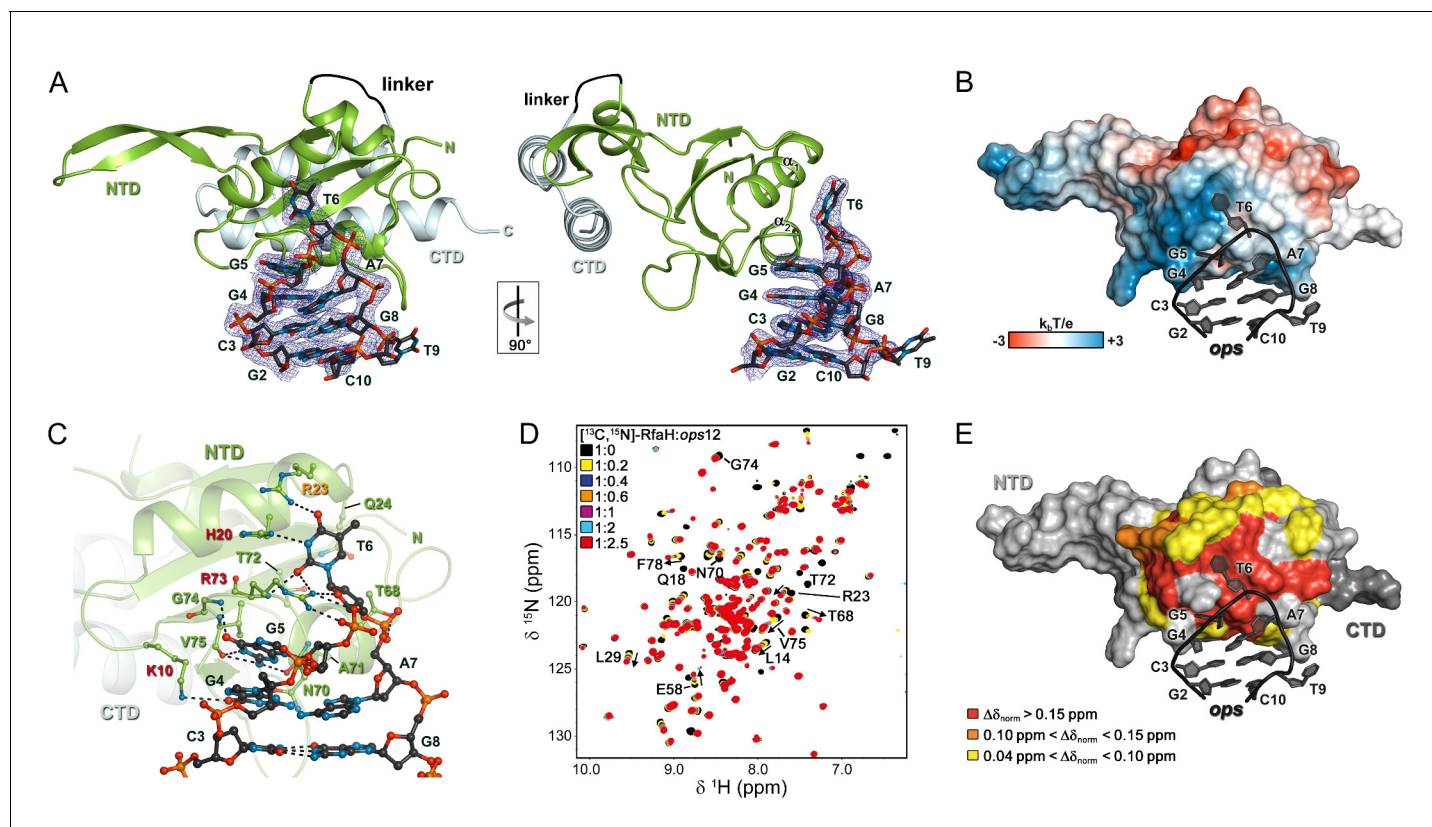


Figure 4. Specific recognition of ops by RfaH. (A) Crystal structure of the RfaH:ops9 complex with the 2F_o - F_c electron density map contoured at 1 σ. (B) Structure of RfaH:ops9 complex with RfaH shown in surface representation, colored according to its electrostatic potential and ops9 as sticks. (C) Details of RfaH:ops9 interactions. Hydrogen bonds are shown as black dashed lines. RfaH residues that interact with ops are labeled in green. Alanine substitutions of RfaH residues that make base-specific contacts to G5 and T6 via their side chains and that compromise RfaH recruitment (Belogurov et al., 2010) are highlighted in red (strongly defective) and orange (moderately defective). (D) RfaH:ops interactions in solution. [1H, 15N]-HSQC spectra of 110 μM [13C, 15N]-RfaH titrated with 803 μM ops12 DNA. Arrows indicate changes of chemical shifts. Selected signals are labeled. (E) Mapping of normalized chemical shift perturbations observed in (D) on the RfaH:ops9 structure.

DOI: <https://doi.org/10.7554/eLife.36349.008>

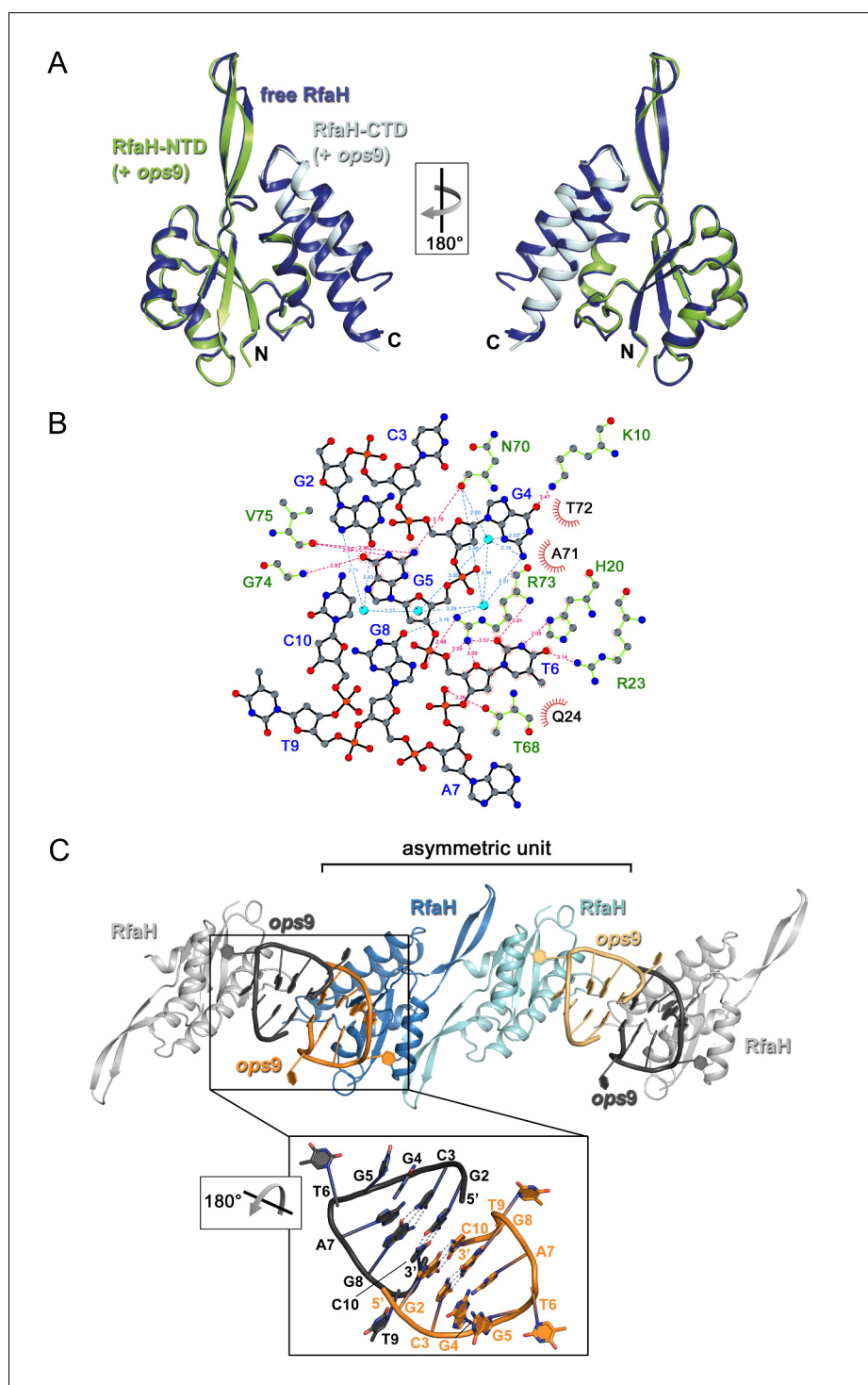


Figure 4—figure supplement 1. Analysis of RfaH:ops interactions. (A) Superposition of free RfaH (PDB ID 2OUG, blue) and RfaH in the RfaH:ops9 complex (RfaH-NTD, green; RfaH-CTD, cyan), both in ribbon representation. The superposition is based on C α atoms, yielding a root mean square deviation value of 0.72 Å. (B) Two-dimensional view of ops9:RfaH interactions. Water molecules are shown as cyan spheres. Bases are labeled in blue, amino acids in green. Hydrogen bonds between amino acids and DNA are shown as pink dashed lines, water-mediated hydrogen bonds as blue dashed lines. Van-der-Waals interactions are indicated in red. (C) Crystal packing of RfaH:ops9 crystals (space group P1). The content of one asymmetric unit is shown in color, molecules from neighboring asymmetric units are in grey (all molecules in ribbon representation). The panel shows an enlargement of the

Figure 4—figure supplement 1 continued on next page

Figure 4—figure supplement 1 continued

arrangement of DNA molecules from neighboring asymmetric units. Bases are shown as sticks (oxygen, red; nitrogen, blue; carbon, according to backbone color) and are labeled. Hydrogen bonds are indicated by dashed lines.

DOI: <https://doi.org/10.7554/eLife.36349.009>

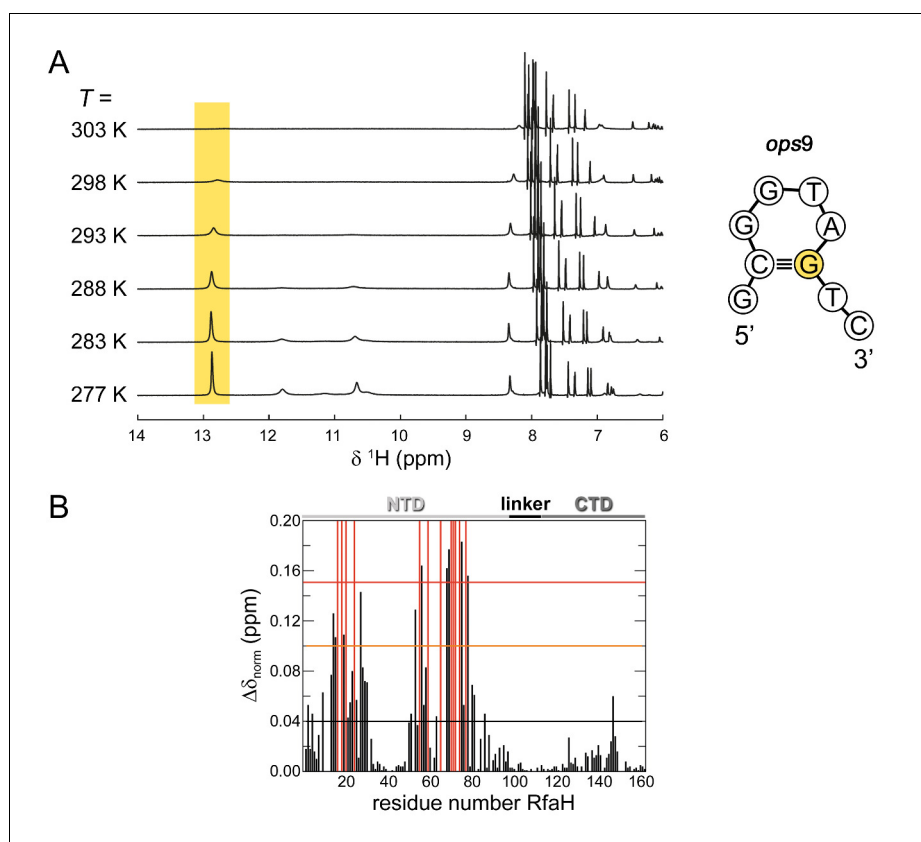


Figure 4—figure supplement 2. Secondary structure of isolated *ops9* and RfaH:*ops9* interaction in solution. (A) Downfield region of the 1D ^1H NMR spectra of *ops9* DNA measured at the temperatures indicated. The region characteristic for imino protons involved in Watson-Crick base pairs is highlighted in yellow. The inferred secondary structure of *ops9* DNA is shown to the right. (B) RfaH:*ops12* interaction in solution. ^1H , ^{15}N -HSQC-derived normalized chemical shift perturbations vs. residue position in RfaH. The chemical shift perturbations were obtained from a titration of ^{15}N -RfaH with *ops12* (for spectra see **Figure 3D**). Horizontal lines: significance levels of $\Delta\delta_{\text{norm}} = 0.04$ ppm, black; $=0.1$ ppm, orange; $=0.15$ ppm, red. Disappearing signals are highlighted in red. The domain arrangement of RfaH is indicated above.

DOI: <https://doi.org/10.7554/eLife.36349.010>

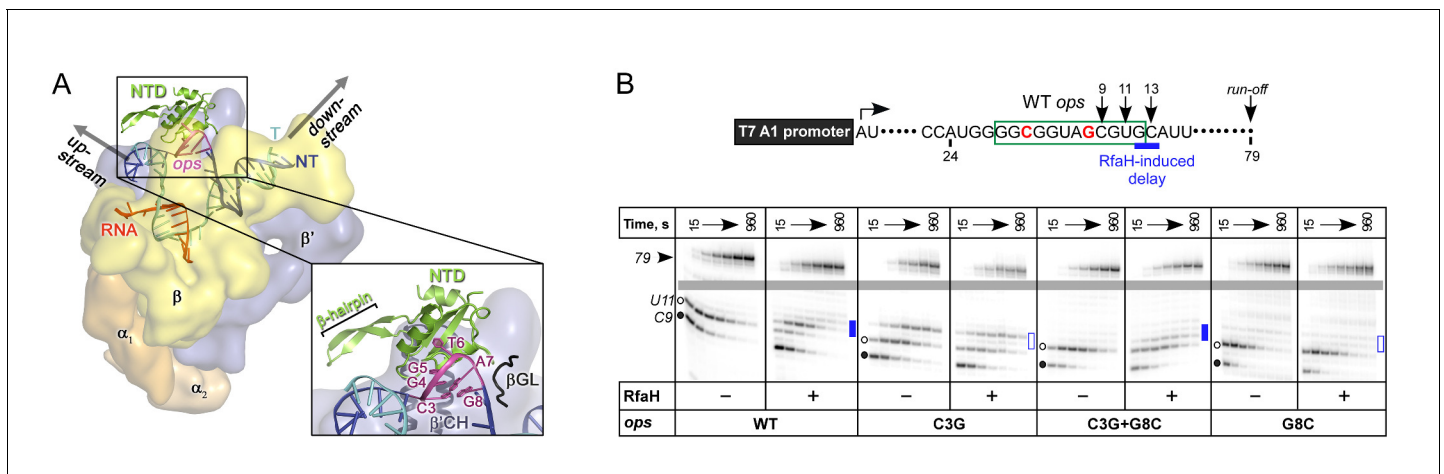


Figure 5. The role of NT DNA hairpin. (A) Model of RfaH-NTD bound to the ops-paused TEC. Surface-accessible NT DNA bases are shown as sticks. (B) The double C3G + G8C substitution partially restores RfaH-dependent recruitment. The assay was done as in **Figure 2**. The position of an RfaH-induced delay in RNAP escape is shown with a blue bar, solid if delay is enhanced.

DOI: <https://doi.org/10.7554/eLife.36349.013>

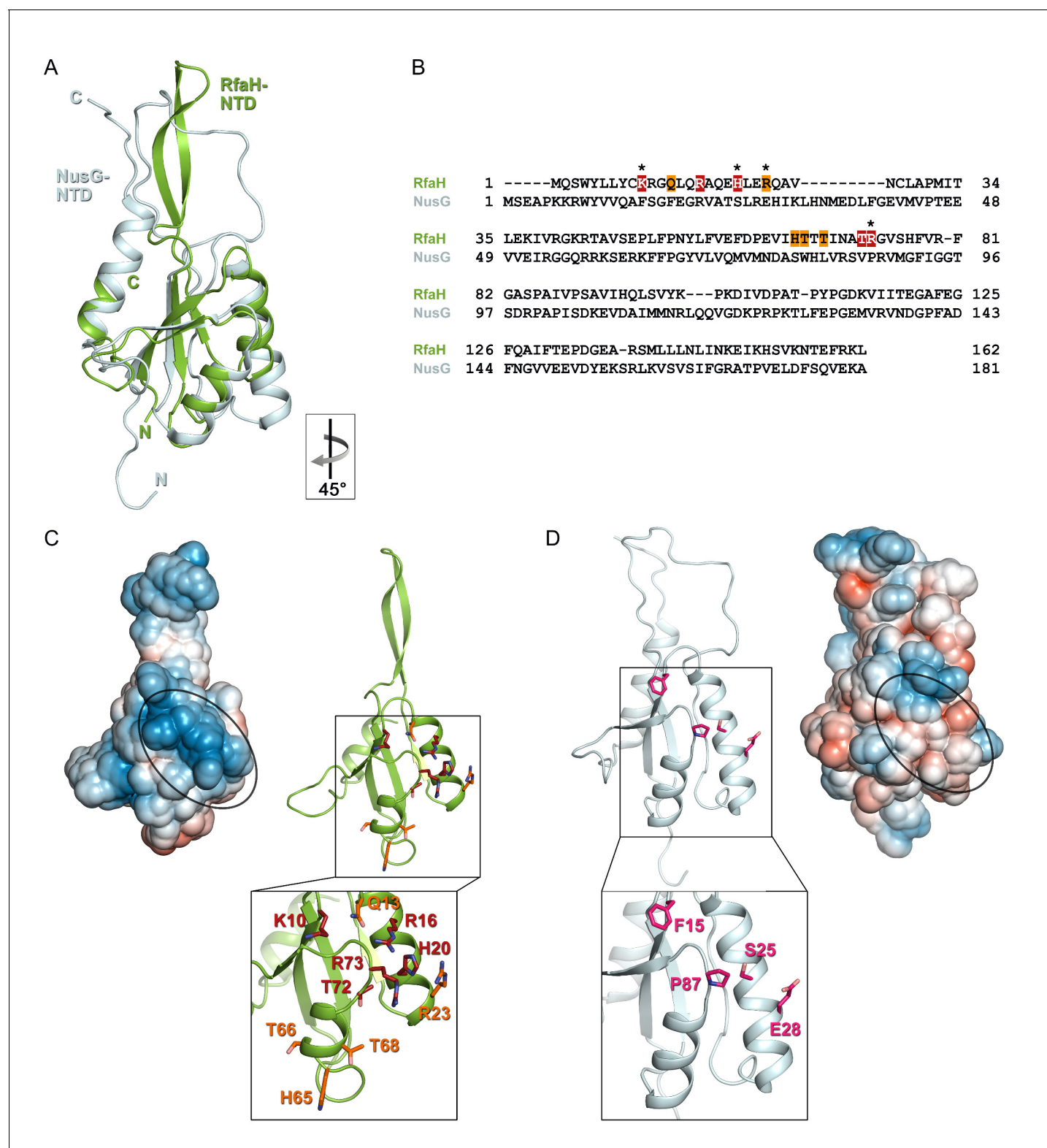


Figure 6. Specificity of RfaH for *ops*. Superposition based on backbone atoms of NusG-NTD (PDB ID 2K06, light blue) and RfaH-NTD (taken from the RfaH:ops9 structure, green; root mean square deviation: 4.3 Å). Both proteins in ribbon representation. (B) Structure-based sequence alignment of NusG and RfaH. RfaH residues whose substitutions for Ala compromise RfaH recruitment (Belogurov et al., 2010) are highlighted in red (strongly defective) and orange (moderately defective). RfaH residues that make base-specific interactions with *ops* via their side chains are marked by an asterisk. (C) Structure of RfaH-NTD in (left) surface representation colored according to its electrostatic potential (from $-3k_B T/e$, red, to $+3k_B T/e$, blue)

Figure 6 continued on next page

Figure 6 continued

and (right) ribbon representation with residues highlighted in (B) shown as sticks (C atoms, red or orange; N atoms, blue; O atoms, light red). (D) Structure of NusG-NTD (PDB ID 2K06) in (left) surface representation colored according to its electrostatic potential and (right) ribbon representation. Residues corresponding to the amino acids of RfaH highlighted in (B) are shown as sticks (C atoms, pink; N atoms, blue; O atoms, light red).

DOI: <https://doi.org/10.7554/eLife.36349.014>