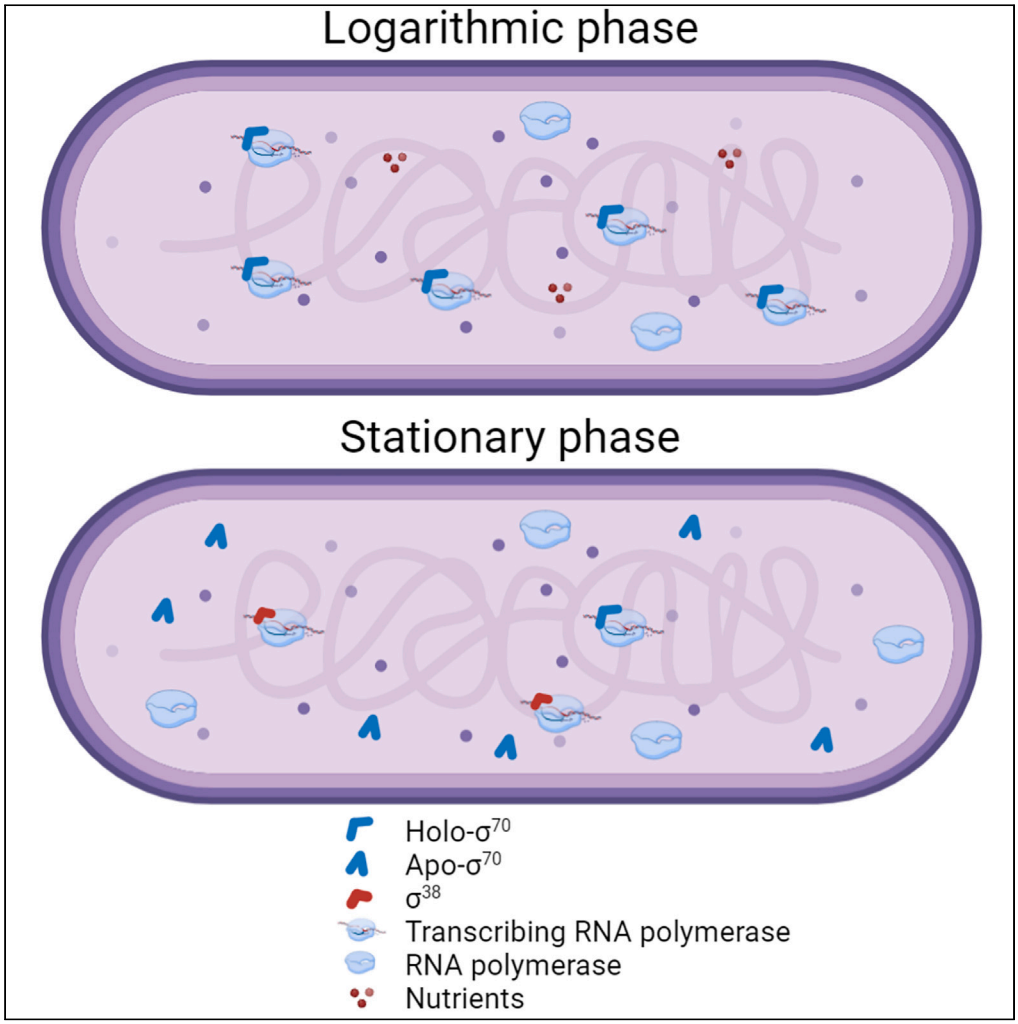
iScience

CellPress
OPEN ACCESS

Article

Evidence for a compact σ^{70} conformation *in vitro* and *in vivo*



Khalil Joron, Joanna Zamel, Nir Kalisman, Eitan Lerner

eitan.lerner@mail.huji.ac.il

Highlights

smFRET and CL-MS point toward a compact conformation of apo-σ^{70}

Structural basis for inhibition of σ^{70}'s DNA binding when not part of transcription

Apo-σ^{70} in its compact conformation exists *in vivo* during stationary phase

Joron et al., iScience 27, 110140
June 21, 2024
https://doi.org/10.1016/j.isci.2024.110140

iScience

CellPress
OPEN ACCESS

Article

Evidence for a compact σ^{70} conformation *in vitro* and *in vivo*

Khalil Joron,[1] Joanna Zamel,[1] Nir Kalisman,[1,2] and Eitan Lerner[1,2,3,*]

SUMMARY

The initiation of transcription in Escherichia coli (*E. coli*) is facilitated by promoter specificity factors, also known as σ factors, which may bind a promoter only as part of a complex with RNA polymerase (RNAP). By performing *in vitro* cross-linking mass spectrometry (CL-MS) of apo-σ^{70}, we reveal structural features suggesting a compact conformation compared to the known RNAP-bound extended conformation. Then, we validate the existence of the compact conformation using *in vivo* CL-MS by identifying cross-links similar to those found *in vitro*, which deviate from the extended conformation only during the stationary phase of bacterial growth. Conclusively, we provide information in support of a compact conformation of apo-σ^{70} that exists in live cells, which might represent a transcriptionally inactive form that can be activated upon binding to RNAP.

INTRODUCTION

σ promoter specificity factors facilitate the specific binding of RNA polymerase (RNAP) to gene promoters in bacteria.[1,2] Among the different types of σ factors in *E. coli*, σ^{70} binds promoters of house-keeping genes under optimal conditions at the logarithmic phase of bacterial growth.[3] The mechanism of initiating DNA transcription involves the binding of σ^{70} to dsDNA promoter only when σ^{70} is bound to the core RNAP complex.[4] In the RNAP complex, σ^{70} is found in a unique structural organization, in which σ^{70} regions 2 and 4 (σR2 and σR4, respectively) are distant from each other,[4] to expose its DNA-binding residues, and by that facilitate transcription initiation.[5] The formation of the RNAP-promoter closed (RP_C) complex is followed by a cascade of DNA isomerization events,[6–10] which end with a stretch of 10–12 melted bases of promoter DNA upstream to the transcription start site, which forms the DNA transcription bubble, that is later stabilized in the RNAP-promoter open (RP_O) complex.[11,12] Importantly, throughout transcription initiation σ^{70} remains part of the holoenzyme complex and retains the extended conformation. Nonetheless, as all parts of the transcription initiation complex, including σ^{70}, are translated and formed separately, σ^{70} may exist in bacteria also in an unbound apo form, at least as long as it is not bound to RNAP. Apo-σ^{70} can also bind anti-σ factors, which repress transcription by competing with RNAP on the interaction with σ^{70}, and even by sequestering σ^{70} out of the cytoplasm.[13–15] However, different factors may lead to the release of σ^{70} from its interaction with anti-σ factors, and back into the cytoplasm, until it rebinds RNAP.[16–19] Therefore, this process, among others, might dictate the lifetime in which σ^{70} exists in an unbound apo form.

Structurally, the protein databank (PDB) includes many entries of high-resolution structures of transcription initiation complexes, where σ^{70} is present as a subunit.[4,8,20–30] Exploring the PDB for apo-σ^{70} identifies only a few *E. coli*-related structures of its separate regions, such as σ^{70} region 4 (i.e., σR4) bound to several transcription factors,[31–33] the unbound σ^{70} region 2 (i.e., σR2),[34,35] as well as structures of regions in house-keeping σ factors from bacteria other than *E. coli*.[35,36] However, a structural description of full-length apo-σ^{70} has not yet been reported. Nevertheless, previous *in vitro* biochemical works showed that apo-σ^{70} adopts a predominant distinct compact conformation,[37–40] which can undergo an overall structural reorganization upon binding core RNAP. Binding induces a conformational change from an unbound compact state to an RNAP-bound extended state.[41–45]

In this work, we provide evidence for the existence of the compact apo-σ^{70} conformation both *in vitro* and *in vivo*, as well as present possible structure models of the conformation solved using PatchDock.[46–48] First, we perform *in vitro* single-molecule Förster resonance energy transfer (smFRET)[49,50] experiments to verify that apo-σ^{70} is found predominantly in a compact conformation. In addition, we also find that the presence of a promoter DNA alone (i.e., in the absence of RNAP) can induce only a minor population shift toward the extended conformation of σ^{70}. Relying on smFRET data, we further study the structural features of apo-σ^{70} by performing *in vitro* cross-linking mass-spectrometry (CL-MS).[51–56] Comparison of the resolved residue proximities against the PDB full-length σ^{70} structures confirms the description of apo-σ^{70} as having a compact conformation. Furthermore, we show a global decrease in the distance between σR2 and σR4, which could support a DNA-binding auto-inhibitory mechanism, for preventing DNA binding not in the context of transcription.

[1]Department of Biological Chemistry, Alexander Silberman Institute of Life Sciences, Faculty of Mathematics & Science, Edmond J. Safra Campus, Hebrew University of Jerusalem, Jerusalem 9190401, Israel
[2]Center for Nanoscience and Nanotechnology, Hebrew University of Jerusalem, Jerusalem 9190401, Israel
[3]Lead contact
*Correspondence: eitan.lerner@mail.huji.ac.il
https://doi.org/10.1016/j.isci.2024.110140

Finally, results from *in vivo* CL-MS experiments on σ^{70} in *E. coli*, report the abundance of inter-residue proximities that cannot be explained by the extended σ^{70} PDB structures. Interestingly, the compact apo-σ^{70} conformation exists not in the logarithmic bacterial growth phase when σ^{70} is recruited for transcription, but rather in the stationary phase when alternative σ factors are recruited for transcription to replace σ^{70}.

In vitro single-molecule FRET of apo-σ^{70} reveals a predominant compact conformation

To study the inter-region conformational dynamics of σ^{70}, we doubly labeled σR2 and σR4 with fluorescent dyes for smFRET measurements, as shown before[37](see STAR Methods). It was previously suggested that the interaction with RNAP causes extension in the conformation of σ^{70} by localizing σR3 and σR4 distant from σR2. Importantly, using an smFRET hybridizing probe-based transcription assay, presented previously by Weiss and co-workers,[57,58] we show that both the *wt* and the Cys mutant used for dye-labeling are transcriptionally active (Figure S1). The Cys residues form a covalent bond with the dye originating maleimide groups through thiol chemistry. The double-labeled σ^{70} has an expected FRET efficiency of ∼0.5 when it is a subunit of the RNAP holoenzyme structure.[5] SmFRET measurements result in a predominant FRET sub-population with higher FRET efficiency (Figures 1A–1C), which point toward a conformation of σ^{70} where regions 2 and 4 are in close proximity, consistent with previously published data of similar smFRET measurements.[5] Further, we report the distribution of raw FRET values (E_{raw}) of single σ^{70} molecules, probed for freely diffusing molecules through a confocal spot within few ms, in which they can undergo conformational changes at the diffusion timescale or faster. To report the underlying conformational states and the potential within-burst conformational dynamics, we analyze the smFRET burst data using multi-parameter photon-by-photon hidden Markov modeling (mpH^{2}MM)[59] (Figures 1D–1F). Apo-σ^{70} transitions between four sub-populations, with one being an acceptor-photoblinked sub-population, and hence does not report FRET-relevant information (Figure 1D, purple dots), and three more sub-populations with both donor and acceptor fluorescently active (Figure 1D, blue, red, and green dots). The rates of the transitions between the three FRET sub-populations are in the ms timescale, where only rates of 20 s^{-1} or higher can be considered as arising from within-burst dynamics recorded in the experiment, due to burst durations of 50 ms at most (Table S1). Two out of the three FRET sub-populations exhibit almost equal mean E_{raw} values (Figure 1D, black dots within the group of blue and red dots) and together represent one predominant FRET sub-population (Figure 1D, blue and red dots). The third minorly populated FRET sub-population (Figure 1D, green dots) exhibits a slightly lower mean E_{raw} (Figure 1D, black dot within the group of green dots). These results suggest a predominant compact σ^{70} conformation and a minorly populated slightly less compact σ^{70} conformation. Incubating the dye-labeled apo-σ^{70} in the presence of 2 μM dsDNA *lacCONS*[57,60] promoter (Figures 1C and F), gives rise to a small fraction of σ^{70} molecules in an FRET sub-population (Figure 1F, green dots) at lower mean E_{raw} (Figure 1F, black dot within the group of green dots) compared to the same minorly populated sub-population in apo-σ^{70}. These results correspond to a further increase in the distance between σR2 and σR4 for a low fraction of σ^{70} molecules, whereas the majority of apo-σ^{70} stay in the high FRET sub-population. In the presence of lower concentration of dsDNA *lacCONS* promoter (100 nM; Figures 1B and E), this minorly populated sub-population decreases to values similar to those in the absence of dsDNA (as seen relative to Figure 1A). Notably, the dsDNA *lacCONS* promoter should bind σ^{70}, when it is a subunit of the RNAP holoenzyme.[60,61] Indeed, in previous smFRET studies involving σ^{70} a full shift toward lower FRET values occurred in the presence of only a few nM of dsDNA, only when in the presence of RNAP.[9,10] Therefore, a small fraction of apo-σ^{70} molecules can undergo a small conformational change in the presence of μM concentrations of dsDNA. The transition rates from the predominant compact conformation represented by the high FRET sub-populations (blue and red) to the slightly less compact sub-population (green), are in the tens of ms, and become as slow as hundreds of ms in the presence of dsDNA (Table S1). It is important to remember, however, that transition rates slower than 20 s^{-1} only point toward the inability of mpH^{2}MM to resolve the slow conformational dynamics in these cases.

Overall, the results of the *in vitro* smFRET measurements suggest that apo-σ^{70} is in an equilibrium between two states, a predominant compact conformation and a minorly populated slightly less compact conformation that is still more compact than the fully-extended-RNAP-bound σ^{70}, which is well-described by PDB structures. Therefore, we can assume that C_α-C_α distances between σR2 and σR4 in the predominant conformation of apo-σ^{70} may yield the short proximities that allow cross-linking to occur efficiently. Thus, we set out to describe the structural features of the apo-σ^{70} compact conformation using *in vitro* CL-MS.

Structural features of the unique compact apo-σ^{70} conformation using *in vitro* CL-MS

We perform *in vitro* cross-linking of purified recombinant *wt* apo-σ^{70} with either a zero-length primary amine to carboxyl coupler, 4-(4,6-dimethoxy-1,3,5-triazin-2-yl)-4-methyl-morpholinium chloride (DMTMM),[53] or a primary amine to primary amine cross-linker, bis(sulfosuccinimidyl)-substrate (BS^3).[62] The cross-linked sample is then digested with trypsin and analyzed using mass-spectrometry (MS; for further information, see method details). We attain both short and intermediate scale spatial information, which report the C_α-C_α distances that are <16 or <30 Å for DMTMM and BS^3, respectively (Table 1). We show that without cross-linking the majority of σ^{70} (∼66%) is in a monomeric form, while the rest exists as higher-order species (Figure S2A). As a result of cross-linking, the fraction of monomeric σ^{70} decreases by 6%, without affecting the overall layout of the eluted peaks (Figure S2B). Therefore, we suggest that many of the cross-links we achieve originate from the monomeric form. After false detection rate (FDR)-based data filtration (Figures S3A and S3B) according to the fragmentation score of each cross-linked peptide (see method details), the highest-ranking cross-links are summarized in Data S1. We evaluate the satisfaction rate of the recovered C_α-C_α distances against the same pairs of residues in existing PDB structures of σ^{70} (Table S2). In the RNAP holoenzyme PDB structure, σR2 and σR4 are far from each other with a distance larger than the cross-linker's maximal length. However, in the σ^{70} interconnectivity map (Figure S4) we identify pairs of cross-linked residues linking these two regions together, suggesting that σR2 and σR4 in apo-form are in closer proximity at distances below the maximal length of BS^3 and DMTMM. Since all existing PDB structures of σ^{70} possess similar

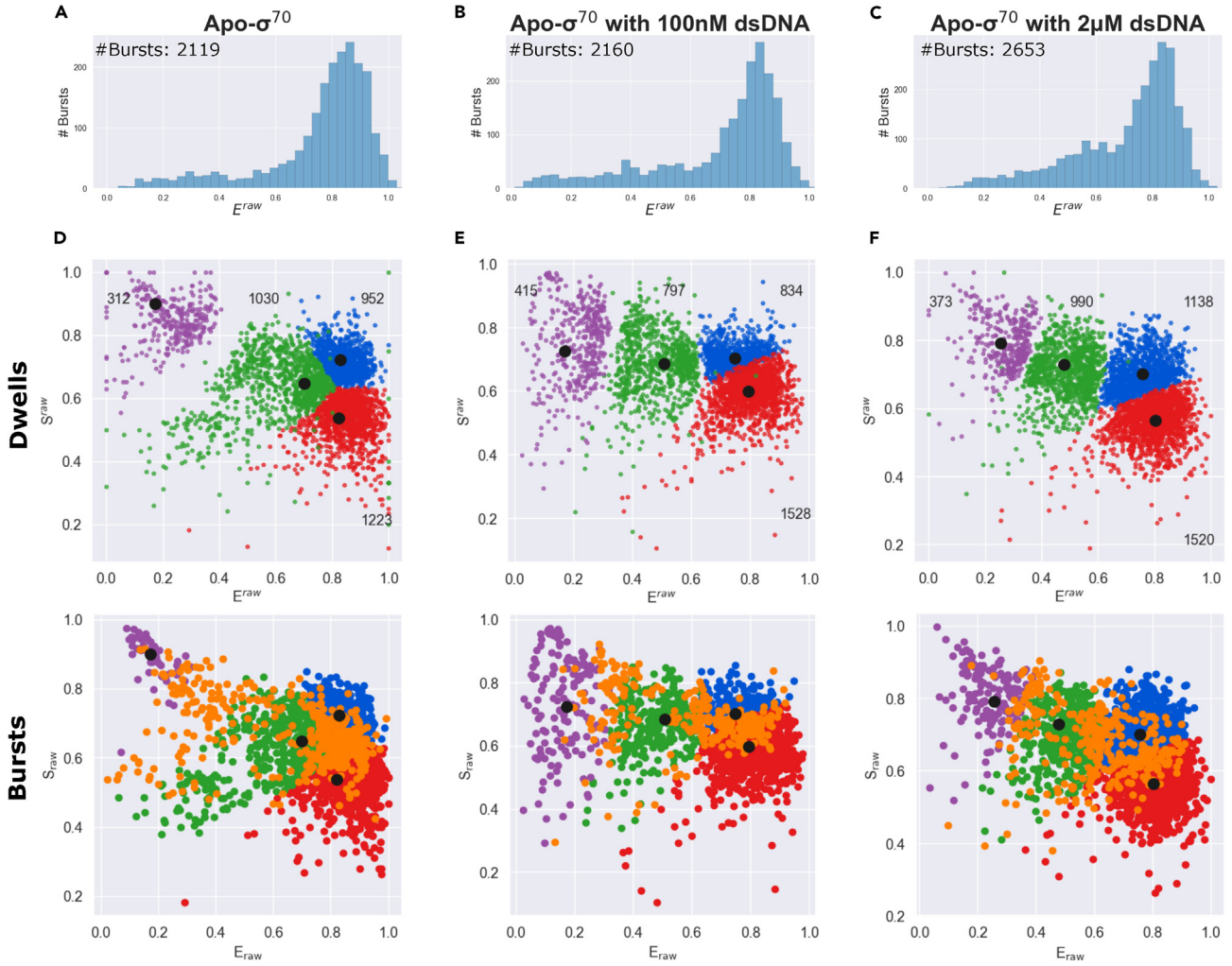


Figure 1. Single molecule FRET bursts and mpH^{2}MM analysis reveals apo-σ^{70} is unable to specifically bind dsDNA, and that within-burst dynamics between predominant compact σ^{70} conformation and less compact σ^{70} conformation as a minor sub-population

(A–C) FRET burst analysis (D–F) mpH^{2}MM analysis of σ^{70} labeled at residues 442 and 579 with ATTO 550 and ATTO 643, respectively. (A) Apo-σ^{70}, a major population exists at high FRET with a tail toward lower FRET efficiencies. (B) Apo-σ^{70} in the presence of 100 nM dsDNA exhibits no significant change in the major high FRET population. (C) Apo-σ^{70} in the presence of 2 μM dsDNA exhibits a slight increase in sub-population with lower FRET. (D–F) Dwells in states within bursts, (dwell-based analysis; top) and burst-based analysis (bottom), where different sub-populations in different states are shown in blue, red, green, and magenta, and the bursts that include transitions between states are shown in orange. (D) Apo-σ^{70}—the predominant sub-populations (blue and red) of the compact conformation exhibit dynamics to a minor sub-population (green) of a less compact conformation. (E) σ^{70} in the presence of 100 nM dsDNA—the green sub-population is shifted, relative to (D), toward lower mean FRET efficiency suggesting binding to dsDNA influences the conformation of a minor sub-population in σ^{70}, while (F) in the presence of 2 μM dsDNA shows a higher an increase in the population of the green sub-population, suggesting the dsDNA-bound σ^{70} occurs slightly at μM concentrations, yet most of σ^{70} still remains in the high FRET sub-population (in the compact conformation; blue and red). Transition times between the blue or red sub-populations and the green sub-population, and vice versa, are in the few to tens of ms, in the absence or presence of dsDNA. The magenta sub-population can be best described by a pure donor-only sub-population shown using uncorrected raw FRET and ALEX data (E_{raw} and S_{raw}). See also Figure S1 and Table S1.

features and are almost identical, we display the recovered cross-links using one representative PDB structure of σ^{70} in the RNAP holoenzyme complex (6P1K)[25] (Figures 2A and 2B), and summarize the results for all PDB structures in a table (Table 1). We find that up to 67% of BS^3 and 75% of DMTMM cross-links are not satisfied and do not fit within the spatial coordinates given by PDB structures. According to PDB structures coordinates most cross-links exceed the C_α-C_α distance range covered by the cross-linkers, suggesting that the PDB structures do not represent all possible monomeric conformations captured in solution during cross-linking. Next, we perform modeling using AlphaFold2[63,64] and RoseTTAFold,[65] to test whether the expected compact conformation may be predicted via training and based on previously acquired structural data alone. AlphaFold2 was able to predict a more compact conformation than the one present in the PDB (Figures 3A–3C). Yet, predicted models retain a high violation percentage of the cross-linking data (Table S2; Figures 3D and 3E), suggesting a more compact

Table 1. Pairs of residues recovered from apo-σ^{70} *in vitro* CL-MS

BS^3 cross-linked Regions	Residue #1	Residue #2	BS^3 cross-linked Regions	Residue #1	Residue #2	DMTMM cross-linked Regions	Residue #1	Residue #2
σR1.2-σNER	**K121**	**K257**	σNER-σR4	K343	K593	σR1.2	**T107**	**E104**
σR.12-σNER	**K121**	**K264**	σNER-σR4	K264	K593	σR1.2-σR2	D96	K376
σR1.2-σNER	K121	K238	σNER-σR4	K236	K578	σR1.2-σR2	E104	K371
σR1.2-σNER	K121	S241	σNER-σR4	K257	K593	σR1.2-σR2	E104	S389
σR1.2-σNER	**K121**	**K220**	σR2	**K371**	**K377**	σR1.2-σR2	**E109**	**K426**
σR1.2-σR2	**K121**	**K371**	σR2	**K371**	**K376**	σR1.2-σR2	**D125**	**K371**
σR1.2-σR2	**K121**	**K426**	σR2-σR3	K371	K462	σR1.2-σR3	**E109**	**K462**
σ1.2-σR3	K121	K462	σR2-σR3	**K393**	**K462**	σNER	K236	E335
σR1.2-σR4	K121	K578	σR2-σR3	K418	K462	σNER	K236	E247
σNER	**K220**	**K251**	σR2-σR3	K377	K462	σNER	E152	K257
σNER	**K220**	**K264**	σR2-σR4	K377	S604	σNER	K296	D332
σNER	**K257**	**K264**	σR2-σR4	K371	K578	σNER-σR4	E343	D613
σNER	K238	T244	σR2-σR4	K377	K557	σNER-σR4	E247	K557
σNER	K236	K264	σR2-σR4	K377	K593	σNER-σR4	E349	K593
σNER	**K220**	**K257**	σR2-σR4	K392	K593	σR2	**E407**	**S442**
σNER	K236	K251	σR2-σR4	K377	K578	σR2-σR3	D445	K462
σNER	**K289**	**K343**	σR3	**K462**	**K496**	σR2-σR4	K376	D613
σNER	K238	S241	σR3	**K462**	**K493**	σR2-σR4	K392	D546
σNER	**K251**	**K264**	σR3-σR4	K462	K557	σR3	**K462**	**E491**
σNER	Y228	K236	σR3-σR4	K462	K593	σR3-σR4	K462	D546
σNER	**K236**	**K257**	σR3-σR4	K462	S604	σR3-σR4	K493	D546
σNER	**Y228**	**K257**	σR3-σR4	K462	K578	σR3-σR4	K493	D613
σNER	**K220**	**S253**	σR3-σR4	K493	K557	σR3-σR4	K462	D613
σNER-σR2	**K257**	**K371**	σR3-σR4	K496	K557	σR4	**D546**	**K557**
σNER-σR2	**K264**	**K371**	σR3-σR4	K493	Y571	σR4	D546	K593
σNER-σR2	**K264**	**K418**	σR3-σR4	K496	K593	σR3-σR4	K462	D613
σNER-σR2	K236	K371	σR3-σR4	K493	K593	σR4	**D546**	**K557**
σNER-σR2	K343	K371	σR4	**K578**	**K593**	σR4	D546	K593
σNER-σR2	K220	K371	σR4	**Y571**	**K593**			
σNER-σR2	**K264**	**K377**	σR4	**K593**	**K597**			
σNER-σR3	K220	K462						
σNER-σR3	K257	K462						
σNER-σR3	K264	K462						
σNER-σR3	K220	K493						
σNER-σR3	K343	K462						
σNER-σR3	K236	K462						
σNER-σR3	S241	K493						
σNER-σR3	K264	K496						

BS^3 and DMTMM pairs of residues recovered from cross-links. Residues are assigned to their originating regions: σ region 1.2 (σR1.2), σ non-essential region (σNER), σ region 2 (σR2), σ region 3 (σR3), and σ region 4 (σR4). Residue numbering is according to the sequence given by the PDB structure 6P1K. Only residues that satisfy the pair-of-residue C_α-C_α distance given by the PDB structure are presented in bold. See also Figures S2–S5, Table S2 and Data S1.

conformation of apo-σ^{70} exists and is not available in the structural database, which the algorithms were trained on. The relatively low predicted local distance difference test score of σR3 and σR4 (Figure 3A), suggest that a more compact conformation which the algorithm was not able to predict may exist. According to the residue-residue alignment confidence plot (Figure 3B), the interaction of the σR2/σR3 with σR4

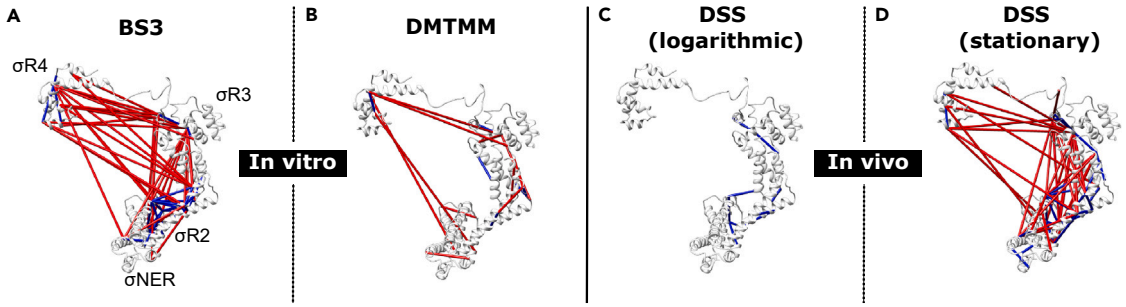


Figure 2. Apo-σ^{70} *in vitro* and *in vivo* distances recovered from CL-MS suggest a compact conformation exists in solution and during stationary phase of bacterial growth

The RNAP-bound σ^{70} conformation from PDB structure (PDB: 6P1K).

(A) *In vitro* BS^3 and (B) DMTMM cross-links.

(C) *In vivo* DSS cross-links at logarithmic growth phase.

(D) *In vivo* DSS cross-links at stationary growth phase. Blue and red—cross-links recovered distances that are within or are not within the C_α-C_α distance covered by the cross-linker, respectively. In (A) σR2, σNER, σR3, and σR4 are indicated. See also Figures S2–S5 and Table S2.

is missing in the AlphaFold2 models, contrary to CL-MS results. Next, we evaluate the satisfaction rate of CL-MS data against the predicted structure models, and find a similar level of satisfaction with up to 36% and up to 50% for the DMTMM and BS^3 data, respectively (Table S2). Overall, across all PDB structures and predicted structure models of σ^{70}, we detect the same cross-linked pairs of residues that cannot be explained by the spatial distances given in these structures. Interestingly, these unique pairs of cross-linked residues include mostly ones connecting σR2/σR3 with σR4 (Table 1), which is the primary DNA-binding regions.

Interactively docked models of apo-form σ^{70} are more compact than holo-form

SmFRET and CL-MS data both report intramolecular proximities in apo-σ^{70}. This warrants the use of recovered pairs of proximal residues as restraints to derive a structure model of σ^{70} in the potentially compact apo-form. Due to the lack of reactive residues in σR1.1, none of the

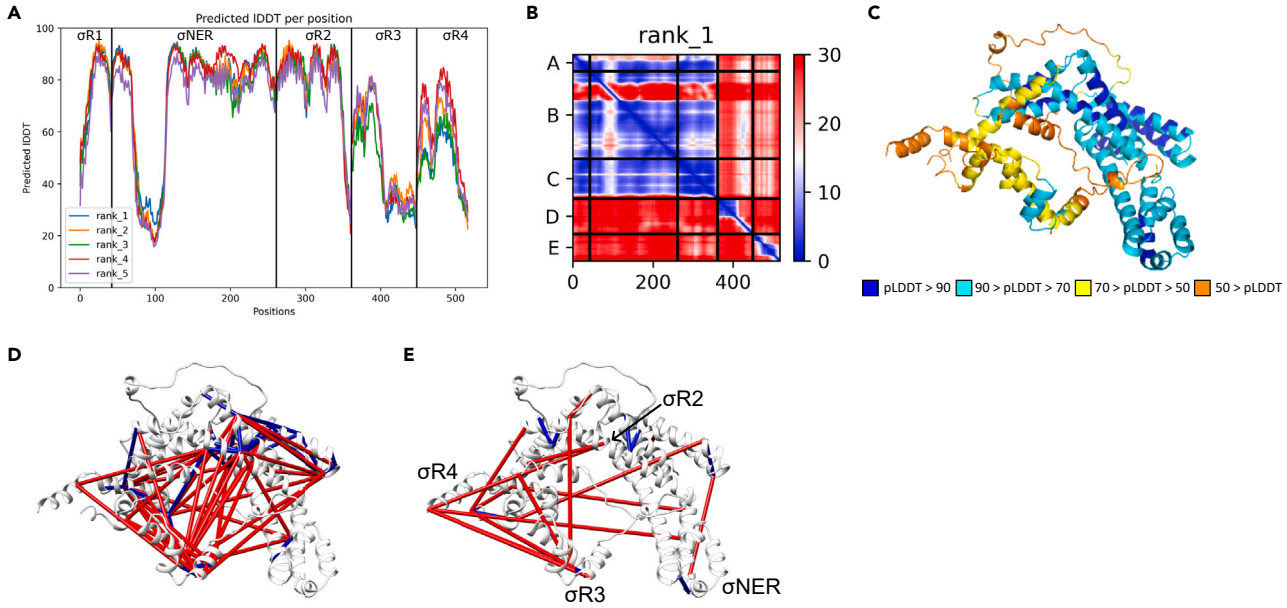


Figure 3. AlphaFold2 Multimer predicts a compact conformation which does not satisfy CL-MS data

(A) Predicted pLDDT (AlphaFold2 output) for each residue in the top 5 predicted model structures. While the domains are predicted with high LDDT, linkers are significantly lower in their values. σ^{70} domains are separated in different boxes.

(B) Predicted inter-residue interactions of the top-ranking model, where A, B, C, D, and E correspond to σR1, σNER, σR2, σR3, and σR4, respectively. Blue color indicates that interaction occurred between the two regions, while red color indicates no interaction.

(C) Rank 1 relaxed model is in a more compact conformation compared to RNAP-bound σ^{70}.

(D) BS^3 and (E) DMTMM cross-links displayed on the rank 1 relaxed model. Blue and red—cross-links recovered distances that are within or are not within the C_α-C_α distance covered by the cross-linker, respectively. In (E) σR2, σNER, σR3, and σR4 are indicated. See also Table S2.

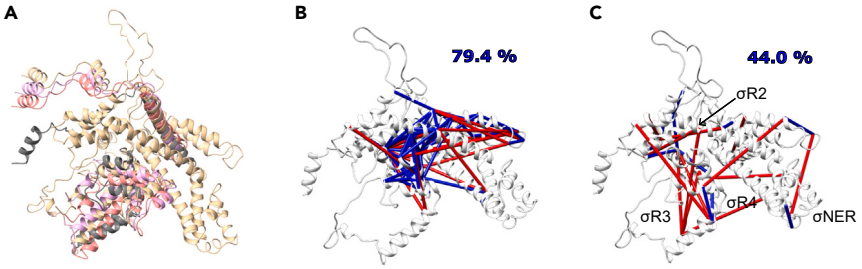


Figure 4. PatchDock CL-MS restraints derived structure models of apo-σ^{70} reveal a compact conformation

PatchDock CL-MS data derived cluster of apo-σ^{70} structure models differing in their C_α chain orientations, yet similar in the percentage of satisfied cross-links. (A) A representative model from the cluster displaying the satisfied (blue) and violated (red) BS^3 (B) and DMTMM (C) apo-σ^{70} cross-links captured *in-vitro*. See % satisfaction above panels. In (C) σR2, σNER, σR3, and σR4 are indicated. See also Data S2.

identified cross-links include residues of σR1.1. Therefore, σR1.1 is not included in the docking process. Comparison of stand-alone σ^{70} regions, solved by nuclear magnetic resonance (NMR) or X-ray diffraction methods, to their structure as part of the whole protein reveals elevated levels of structural similarity.[32] This suggests that σ^{70} regions are rigid domains capable of retaining their structure regardless of the context of the entire protein, which render them suitable for docking-based modeling. Moreover, AlphaFold2 models show the highest local distance difference test (lDDT) in the position of σ non-essential region (σNER) and σR2 in comparison to other regions (Figure 3A), suggesting that σNER and σR2 retain their relative position in apo-σ^{70} with high confidence. These two regions also include the largest number of cross-links, and hence their results are more constrained and include less degrees of freedom. Therefore, our docking process starts with the docking of σNER and σR2. We perform step-by-step integrative docking using PatchDock,[46–48] starting from fragmented regions of the σ^{70} subunit taken from the high-resolution RNAP holoenzyme PDB structure—PDB: 6P1K.[25] The docking process includes three simple stages (see method details). First, we dock σR3 to σNER and σR2, after filtering all models with unacceptable penetrations, we continue the stepwise docking by adding σR4 to the result of the previous step, and then σR1.2. The result is a cluster, where each structure differs in the organization of the main C_α chain and the residues orientation (Figures 4A). The cluster's resolution (RMS) is 1.61 Å, while the root mean square deviation (RMSD) to the extended RNAP-bound structure (PDB: 6P1K) is 3.73 Å. Next, we validate the structure models against the cross-linking data and find a significant increase in the percentage of satisfied cross-links (Figures 4B and 4C). PatchDock structure models are satisfied by ~80% of BS^3 cross-links, and 44% of DMTMM cross-links. It is possible that structural features of apo-σ^{70} lead to steric effects and restrict the DNA binding residues from interacting with the promoter. Therefore, we identify the interaction interface of σ^{70} with DNA in the RP_c complex using the PDB structures summarized in Table S2 and the program PISA.[66] Next, we explore the exposed and buried regions in the extended vs. compact conformations (holo vs. apo), and how these features might affect sidechain solvent accessibility. First, using Discovery Studio Visualizer,[67] we perform solvent accessible surface area (SASA) analysis of the RNAP-bound σ^{70} PDB structure (PDB: 6P1K) and of a representative structure from the apo-σ^{70} cluster (Table 2). We identify a reduction in the SASA of DNA binding residues in the apo conformation, mainly residues located in DNA binding motifs, such as helix-turn-helix. This finding supports the notion that RNAP-σ^{70} interactions occur prior to σ^{70}-DNA interactions, to induce the required conformational change of σ^{70} from a compact to an extended conformation. This, in turn, enables the exposure of σ^{70}'s DNA-binding residues to the solvent to properly form the initial σ^{70} DNA contacts required to initiate RP_c formation.

Overall, we provide structural models and potential roles of structural organizations in regulating the DNA binding capabilities of apo-σ^{70}, all *in vitro*.

In vivo CL-MS identifies structural features of the compact apo-σ^{70} conformation

To explore whether unique features of the predominant compact conformation of apo-σ^{70} exist in living *E. coli* cells, we perform *in vivo* CL-MS during the logarithmic and stationary phases of bacterial growth. Disuccinimidyl suberate (DSS) is a cross-linker, which shares the same end-to-end length and cross-linking chemistry as BS^3. Yet, unlike BS^3, DSS can passively penetrate through cell membranes, including *E. coli*.[68] During the logarithmic phase, σ^{70} is expected to be found mostly in the RNAP-bound form, due to its engagement in the transcription of house-keeping genes.[69] Conversely, during the stationary phase, σ^{38} replaces σ^{70} in the transcription process, which suggest that most of σ^{70} is in an RNAP-unbound form, potentially in the apo form.[70] Following FDR-based filtration of the *in vivo* CL-MS results (Figure S3C), we show that during the logarithmic phase of growth, cross-links can be explained by the RNAP-bound σ^{70} extended conformation (Figure 2C). On the other hand, during the stationary phase, we capture DSS cross-links, which we have previously identified *in vitro* for apo-σ^{70} using BS^3, and are found to describe exclusively its predominant compact conformation (Figures 2A and 2D). Out of the total 68 *in vitro* BS^3 cross-links, we recover only 11 cross-links *in vivo* using DSS during the logarithmic phase (Data S1), but 56 cross-links during the stationary phase. Importantly, while all cross-links captured in the logarithmic phase are satisfied by the extended RNAP-bound σ^{70}, most cross-links captured during the stationary phase are not satisfied and are similar to those captured using BS^3 on apo-σ^{70} *in vitro*. The results of western blot of enriched σ^{70} after *in vivo* DSS cross-linking indicate the cross-linked construct is mostly a monomeric σ^{70} with a small fraction of higher molecular weight complexes (Figure S5). This, in turn, suggests that the majority of the recovered cross-links are a

Table 2. Solvent accessible surface area of DNA interacting residues from σR4 (Number ID: 1–17), σR3 (Number ID: 18–52) and σR2 (Number ID: 53–89) before RP_c formation

Number ID	Residue	Holo-σ^{70} SASA [Å^2]	Apo-σ^{70} SASA [Å^2]
1	E591	29.0	3.4
2	Q589	93.6	82.5
3	R588	128.3	17.3
4	R586	158.2	105.1
5	E585	110.4	13.1
6	R584	148.9	57.7
7	T583	66.8	31.6
8	V582	61.9	34.7
9	D581	97.5	119.0
10	F580	33.7	55.9
11	K578	108.1	116.3
12	E575	83.7	54.5
13	E574	117.6	48.9
14	K573	48.8	23.6
15	T572	55.8	69.4
16	R562	61.9	7.8
17	R554	142.8	130.9
18	S517	28.0	8.6
19	D516	89.9	93.6
20	D514	108.9	53.8
21	D513	84.0	107.2
22	G512	17.6	19.6
23	I511	89.6	67.9
24	P510	58.4	73.5
25	T509	59.1	47.4
26	I505	95.1	96.6
27	E503	92.9	60.0
28	K502	130.7	126.2
29	P478	11.6	4.5
30	E477	110.4	93.5
31	R476	153.3	167.1
32	M474	35.2	65.6
33	E473	86.4	81.0
34	Q472	106.2	119.6
35	K471	67.4	82.5
36	M470	13.6	1.5
37	Q469	106.4	98.5
38	R468	148.7	149.3
39	I466	40.3	25.7
40	R465	143.2	38.4
41	N464	99.0	70.6
42	L463	9.6	10.1
43	K462	88.2	17.9
44	N461	65.5	7.1

(*Continued on next page*)

Table 2. *Continued*

Number ID	Residue	Holo-σ^{70} SASA [Å^2]	Apo-σ^{70} SASA [Å^2]
45	I460	26.2	7.0
46	T459	40.4	33.7
47	E458	111.1	1.5
48	I457	39.2	5.0
49	M456	45.9	25.9
50	H455	116.3	49.0
51	V454	85.0	16.1
52	P453	63.9	4.0
53	R451	149.1	16.4
54	R448	39.8	1.5
55	I443	12.6	18.6
56	R441	147.7	15.8
57	T440	64.8	48.4
58	Q437	61.2	0.0
59	R436	97.8	108.4
60	W434	45.7	19.5
61	W433	137.9	80.1
62	T432	28.8	43.7
63	Y430	47.2	34.2
64	T429	91.9	86.4
65	S428	52.9	5.7
66	F427	6.5	21.1
67	K426	90.4	74.6
68	Y425	118.7	69.6
69	G424	27.2	25.7
70	R423	95.6	60.0
71	E420	79.6	86.8
72	F419	19.6	31.2
73	K418	84.5	90.7
74	K414	117.3	20.1
75	F401	57.9	68.9
76	G398	16.1	27.7
77	R397	104.8	118.6
78	N396	104.8	112.3
79	T395	69.2	63.5
80	Y394	64.1	32.6
81	K393	123.4	121.1
82	K392	105.1	100.7
83	I390	35.7	20.1
84	S389	40.8	6.3
85	I388	11.6	59.9
86	L386	39.8	14.1
87	R385	112.4	129.8
88	N383	14.0	18.1
89	A382	9.6	31.2
Average	–	75.7	53.8

Compared between the extended RNAP-bound holo conformation, and the compact apo conformation. See also Figure S4.

result of conformational changes in the monomeric apo-σ^{70}. The cross-links recovered from *in vivo* CL-MS are summarized in Data S1 and are displayed on top of a representative PDB structure of σ^{70} (PDB: 6P1K; Figures 2C and 2D). Our results suggest that this apo-σ^{70} compact structure in *E. coli* can be anticipated not during the logarithmic phase, where housekeeping genes are needed, but rather during stress when the binding of σ^{70} to RNAP is replaced by an alternative σ factor. Furthermore, we identify the interaction interface of σ^{70} with DNA in the RP_o complex using the PDB structures summarized in Table S2 and the program PISA.[66] We find that part of σR1.2, σR2, σR3, and σR4's DNA interacting residues in the RP_o stage undergo cross-linking in the apo-state using DMTMM, these residues include D96, K392, K493, K593, S389, and K462. As mentioned DMTMM is a carboxy to amine coupler active at a distance of up to 16 Å (C_α-C_α), hence in apo-state σ^{70} DNA binding residues are less accessible to the solvent and are less likely to form interaction with DNA. Based on our knowledge on the interaction of the transcription complex with promoter DNA, together with the fact that all structures of RNAP-bound full-length σ^{70} documented in the PDB are in the extended conformation, we suggest that the unique compact conformation of apo-σ^{70} might be responsible for the inhibition of DNA-interaction, and hence serve as a conformational switch activated upon binding to RNAP.

DISCUSSION

In this work, we provide evidence that apo-σ^{70} is found predominantly in a compact conformation, in which σR4 is in close proximity to σR2 and σR3. We perform CL-MS and further confirm the compact conformation of apo-σ^{70} both *in vitro* and *in vivo*, under stationary phase of the bacterial growth. Yet, an atomic resolution experimental structure of apo-σ^{70} has not yet been elucidated. CL-MS is a high throughput technique able to recover distance restraints between pairs of residues, from just a few micrograms of protein.[51–56]

CL-MS is limited in that it reports solely on pairs of residues that are within a given distance range and not on the pairs of residues that are at longer distances. Yet, we use two different cross-linking reagents to cover a range of inter-residue distances and to increase the accuracy in predicting the suggested compact conformation by attaining both short and intermediate range distances. As can be seen from the western blot and size exclusion chromatography results, our list of cross-links includes mostly ones that originate from the monomeric apo-σ^{70}, while some cross-links may originate from high-order oligomers at lower frequency. Importantly, we could not decipher which of the cross-links found are inter-molecular, hence arise from aggregates, if any. Hence, it is possible that the modeled compact conformation could be slightly influenced by cross-links originating from aggregates. If so, we expect DMTMM, being a short distance linker, to be the main contributor in that matter since shorter distances are more abundant in aggregates. This could be one reason why some cross-links are still violated even by the modeled compact conformation (Figures 4B and 4C, red lines). In detail, no more than 20% of the BS^3 cross-links captured *in vitro* are violated by our structural models. However, more than 50% of the DMTMM cross-links are violated. Nevertheless, the low fraction of these aggregates suggests that the majority of recovered BS^3 cross-linking data arises from the monomeric fraction. Overall, CL-MS results suggest that apo-σ^{70} may attain multiple different compact structural organizations, which are yet to be structurally recorded in the PDB. Our modeling utilizes PatchDock assisted by all the CL-MS data, including DMTMM cross-links. Yet, rather than forcing the cross-links to generate a specific model, the algorithm uses the data to validate the results. Only structural models that satisfy the highest percentage of cross-links are considered. The more cross-links the model satisfies, the more likely it is to resemble the compact conformation that was captured, but it will most probably never represent all conformations captured in solution. This could be another reason why some cross-links are still violated even by the modeled compact conformation (Figures 4B and 4C, red lines). Regardless, the structural models presented here represent the most likely monomeric structure of the apo-form present in solution.

Using smFRET, we show that in the presence of relatively high concentrations of promoter dsDNA, apo-σ^{70} remains predominantly in a compact conformation. Interestingly, a small fraction of σ^{70} in solution exhibits a conformation that is slightly less compact, suggesting that DNA binding may occur to some degree and might induce a slight conformational change between σR2 and σR4 toward a slightly open conformation. Therefore, it raises the question of whether or not the extended conformation of σ^{70} exists intrinsically and dynamically in apo form, interchanging between the two conformations. In fact, previous smFRET studies suggested that such intrinsic conformational dynamics may exist in *E. coli* apo-σ^{70} [5]. As previously shown by *Vishwakarma* et al.,[37] we measure the distance between σ^{70} regions using FRET efficiencies between a donor dye labeling residue 579 of σR4 and an acceptor dye labeling residue 442 of σR2. Overall, the results point toward a preexisting conformational equilibrium between a predominant compact conformation, and a minorly populated less compact conformation of apo-σ^{70}. Furthermore, we show that σ^{70} may partially bind dsDNA in the absence of RNAP, and at low affinity (μM), as opposed to DNA binding in the presence of RNAP, which occurs at high affinity (nM). In that respect, it is important to mention that σ^{70} has been reported to bind DNA structures that deviate from the dsDNA one (e.g., ssDNA in aptamers),[45] however, not at high affinities. The ability of apo-σ^{70} to bind DNA is predictable since σ^{70} undergoes intrinsic dynamic conformational changes, which slightly open the σR2 and σR4 DNA binding interface. One can speculate that if σ^{70} would bind dsDNA not in the context of RNAP, it would probably bind it nonspecifically, hence the interaction would be unstable and will lead to inefficient transcription initiation. Importantly, σ^{70} binds promoter DNA specifically. However, it was previously shown that after its dissociation from RNAP it may re-associate during elongation if a promoter-like sequence is present.[71] After binding to RNAP and forming the holoenzyme complex, σ^{70}'s ability to bind to non-promoter DNA sequences at μM concentrations is reduced.[41] Hence, the interaction at μM concentrations is non-specific and may occur due to the fact that apo-σ^{70} has intrinsic dynamics between closed and extended conformations, where the latter may be prone to this low affinity non-specific binding. Indeed, our results encourage the hypothesis and show that dsDNA-induced σR2-σR4 conformational changes are small, occur only to a small fraction of σ^{70} molecules, and only at relatively high dsDNA promoter concentrations. Therefore, to regulate the transcription process and allow efficient transcription initiation it is necessary to protect the DNA binding residues of σ^{70} from being exposed, until binding to RNAP occurs. Binding will activate the required conformational change in σ^{70}, which will extend it and expose its DNA-binding residues. Moreover, similar to core

RNAP, anti-σ factors bind σ^{70} covering the σR2-σR4 DNA-binding interface.[1] Therefore, the σR2-σR4 interface can serve as a conformational switch that self-inhibits and sterically protects the inner DNA binding residues from being exposed, until needed.

In vivo, we captured cross-links using DSS and compared them to the BS^3 cross-links *in vitro*, since they both share similar thresholds. We found similarity in the DSS and BS^3 cross-links, which were found to be exclusive to the predominant compact conformation of apo-σ^{70}, suggesting that the compact conformation exists also *in vivo*. Importantly, this finding reports on a biologically relevant event, where apo-σ^{70} exists in living *E. coli* cells at stationary phase. In other words, σ^{70} is found in an unbound state and is not degraded, or at least not in full. Nevertheless, these results were achieved using a high copy number plasmid of σ^{70}, and hence might also be influenced by over-expression effects. Yet, if over-expression would pose a problem, we would expect to observe the unique cross-linking signatures of the compact apo-σ^{70} conformation also in the logarithmic phase since at this stage σ^{70} already exist in the cell at high concentration. To confirm that most cross-links arise from conformational changes, we have also shown that the majority of *in vivo* cross-linked purified σ^{70} (i.e., a fraction of 93%) is present in the monomeric form.

In summary, we propose that apo-σ^{70} organizes in a predominant compact conformation, which may lead to self-inhibition of the high-affinity promoter binding capabilities of σ^{70}. Only upon activation by binding to RNAP, a large conformational change is anticipated, which exposes these DNA-binding residues. We provide biological evidence for the existence of the compact apo-σ^{70} conformation in live *E. coli* cells during the stationary phase.

Limitations of the study

It is important to note, that size-exclusion chromatography as well as western blot analysis show slight aggregation of σ^{70} in cross-linked samples compared to non-cross-linked samples. Although most cross-links used as restraints in docking originate from monomeric σ^{70}, the suggested structure models might be influenced by cross-links that originated in aggregates, hence are intermolecular.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - E. coli
  - Statistical analysis
- METHOD DETAILS
  - Recombinant protein expression and purification
  - σ^{70} mutagenesis and double labeling
  - RNAP holoenzyme transcription assay
  - SmFRET measurements and data analysis
  - *In vitro* smFRET-based transcription activity of the purified recombinant RNAP
  - *In vitro* cross-linking and mass-spectrometry preparation
  - *In vivo* cross-linking and mass-spectrometry preparation
  - Mass spectrometry analysis
  - Cross-links identification and filtration
  - Estimation of FDR
  - Structure modeling using AlphaFold and RosettaFold
  - PatchDock CL-MS restraint driven integrative docking
  - Western-blot

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2024.110140.

ACKNOWLEDGMENTS

We would like to thank Dr. Irina Artsimovitch for providing us the plasmid encoding for the RNAP holoenzyme rpoD (σ^{70}) subunit, obtained from Addgene. We would also like to thank Dr. Paul David Harris for consultation regarding mpH^2MM analyses of smFRET data. Finally, we would like to thank Dr. Shimon Weiss for fruitful discussions and for inspecting the text of this work. This work was funded by the Israel Science Foundation (grant numbers 556/22 for E.L. and 1768/15 for N.K.), the National Institutes of Health grant R01 GM130942 (E.L. as a subaward), and by the Yad Hanadiv scholarship (K.J.). Molecular graphics and analyses performed with UCSF ChimeraX, developed by the resource for

biocomputing, visualization, and informatics at the University of California, San Francisco, with support from National Institutes of Health R01-GM129325 and the Office of Cyber Infrastructure and Computational Biology, National Institute of Allergy and Infectious Diseases.

AUTHOR CONTRIBUTIONS

Analytical tools: N.K. protein purification: K.J. experimental data: K.J. and J.Z. analysis of experimental data: K.J. and E.L. writing – original draft: K.J. and E.L. writing – reviewing & editing: K.J., J.Z., N.K., and E.L.

DECLARATION OF INTERESTS

All authors declare no competing interests.

Received: January 5, 2024
Revised: March 28, 2024
Accepted: May 27, 2024
Published: May 28, 2024

REFERENCES

1. Burgess, R.R. (2001). Sigma Factors. In Encyclopedia of Genetics, S. Brenner, J.H. Miller, and B.T.-E.of G., eds. (Academic Press), pp. 1831–1834. https://doi.org/10.1006/rwgn.2001.1192.
2. Feklístov, A., Sharon, B.D., Darst, S.A., and Gross, C.A. (2014). Bacterial Sigma Factors: A Historical, Structural, and Genomic Perspective. Annu. Rev. Microbiol. *68*, 357–376. https://doi.org/10.1146/annurev-micro-092412-155737.
3. Paget, M.S.B., and Helmann, J.D. (2003). The sigma70 family of sigma factors. Genome Biol. *4*, 203. https://doi.org/10.1186/gb-2003-4-1-203.
4. Murakami, K.S. (2013). X-ray Crystal Structure of Escherichia coli RNA Polymerase σ70 Holoenzyme. J. Biol. Chem. *288*, 9126–9134. https://doi.org/10.1074/jbc.M112.430900.
5. Davis, M.C., Kesthely, C.A., Franklin, E.A., and MacLellan, S.R. (2017). The essential activities of the bacterial sigma factor. Can. J. Microbiol. *63*, 89–99. https://doi.org/10.1139/cjm-2016-0576.
6. Saecker, R.M., Record, M.T., and deHaseth, P.L. (2011). Mechanism of Bacterial Transcription Initiation: RNA Polymerase - Promoter Binding, Isomerization to Initiation-Competent Open Complexes, and Initiation of RNA Synthesis. J. Mol. Biol. *412*, 754–771. https://doi.org/10.1016/j.jmb.2011.01.018.
7. Ruff, E.F., Record, M.T., Jr., and Artsimovitch, I. (2015). Initial Events in Bacterial Transcription Initiation. Biomolecules *5*, 1035–1062. https://doi.org/10.3390/biom5021035.
8. Chen, J., Chiu, C., Gopalkrishnan, S., Chen, A.Y., Olinares, P.D.B., Saecker, R.M., Winkelman, J.T., Maloney, M.F., Chait, B.T., Ross, W., et al. (2020). Stepwise Promoter Melting by Bacterial RNA Polymerase. Mol. Cell *78*, 275–288.e6. https://doi.org/10.1016/j.molcel.2020.02.017.
9. Mazumder, A., Ebright, R.H., and Kapanidis, A.N. (2021). Transcription initiation at a consensus bacterial promoter proceeds via a 'bind-unwind-load-and-lock' mechanism. Elife *10*, e70090. https://doi.org/10.7554/eLife.70090.
10. Malinen, A.M., Bakermans, J., Aalto-Setälä, E., Blessing, M., Bauer, D.L.V., Parilova, O., Belogurov, G.A., Dulin, D., and Kapanidis, A.N. (2022). Real-Time Single-Molecule Studies of RNA Polymerase–Promoter Open Complex Formation Reveal Substantial Heterogeneity Along the Promoter-Opening Pathway. J. Mol. Biol. *434*, 167383. https://doi.org/10.1016/j.jmb.2021.167383.
11. Bae, B., Feklistov, A., Lass-Napiorkowska, A., Landick, R., and Darst, S.A. (2015). Structure of a bacterial RNA polymerase holoenzyme open promoter complex. Elife *4*, e08504. https://doi.org/10.7554/eLife.08504.
12. Helmann, J.D. (2019). Where to begin? Sigma factors and the selectivity of transcription initiation in bacteria. Mol. Microbiol. *112*, 335–347. https://doi.org/10.1111/mmi.14309.
13. Missiakas, D., and Raina, S. (1998). The extracytoplasmic function sigma factors: role and regulation. Mol. Microbiol. *28*, 1059–1066. https://doi.org/10.1046/j.1365-2958.1998.00865.x.
14. Helmann, J.D. (1999). Anti-sigma factors. Curr. Opin. Microbiol. *2*, 135–141. https://doi.org/10.1016/S1369-5274(99)80024-1.
15. Paget, M.S. (2015). Bacterial Sigma Factors and Anti-Sigma Factors: Structure, Function and Distribution. Biomolecules *5*, 1245–1265. https://doi.org/10.3390/biom5031245.
16. Kang, J.-G., Paget, M.S., Seok, Y.-J., Hahn, M.-Y., Bae, J.-B., Hahn, J.-S., Kleanthous, C., Buttner, M.J., and Roe, J.-H. (1999). RsrA, an anti-sigma factor regulated by redox change. EMBO J. *18*, 4292–4298. https://doi.org/10.1093/emboj/18.15.4292.
17. Pané-Farré, J., Lewis, R.J., and Stülke, J. (2005). The RsbRST Stress Module in Bacteria: A Signalling System That May Interact with Different Output Modules. Microb. Physiol. *9*, 65–76. https://doi.org/10.1159/000088837.
18. Heinrich, J., and Wiegert, T. (2009). Regulated intramembrane proteolysis in the control of extracytoplasmic function sigma factors. Res. Microbiol. *160*, 696–703. https://doi.org/10.1016/j.resmic.2009.08.019.
19. Anne, F.-C., Julia, F., Christian, R., Zingg, E.J., Benjamin, G., and A, V.J. (2009). Sigma factor mimicry involved in regulation of general stress response. Proc. Natl. Acad. Sci. USA *106*, 3467–3472. https://doi.org/10.1073/pnas.0810291106.
20. Mechold, U., Potrykus, K., Murphy, H., Murakami, K.S., and Cashel, M. (2013). Differential regulation by ppGpp versus pppGpp in Escherichia coli. Nucleic Acids Res. *41*, 6175–6189. https://doi.org/10.1093/nar/gkt302.
21. Bae, B., Davis, E., Brown, D., Campbell, E.A., Wigneshweraraj, S., and Darst, S.A. (2013). Phage T7 Gp2 inhibition of Escherichia coli RNA polymerase involves misappropriation of σ70 domain 1.1. Proc. Natl. Acad. Sci. USA *110*, 19772–19777. https://doi.org/10.1073/pnas.1314576110.
22. Degen, D., Feng, Y., Zhang, Y., Ebright, K.Y., Ebright, Y.W., Gigliotti, M., Vahedian-Movahed, H., Mandal, S., Talaue, M., Connell, N., et al. (2014). Transcription inhibition by the depsipeptide antibiotic salinamide A. Elife *3*, e02451. https://doi.org/10.7554/eLife.02451.
23. Zuo, Y., and Steitz, T.A. (2015). Crystal Structures of the E. coli Transcription Initiation Complexes with a Complete Bubble. Mol. Cell *58*, 534–540. https://doi.org/10.1016/j.molcel.2015.03.010.
24. Narayanan, A., Vago, F.S., Li, K., Qayyum, M.Z., Yernool, D., Jiang, W., and Murakami, K.S. (2018). Cryo-EM structure of Escherichia coli σ70 RNA polymerase and promoter DNA complex revealed a role of σ non-conserved region during the open complex formation. J. Biol. Chem. *293*, 7367–7375. https://doi.org/10.1074/jbc.RA118.002161.
25. Chen, J., Gopalkrishnan, S., Chiu, C., Chen, A.Y., Campbell, E.A., Gourse, R.L., Ross, W., and Darst, S.A. (2019). E. coli TraR allosterically regulates transcription initiation by altering RNA polymerase conformation. Elife *8*, e49375. https://doi.org/10.7554/eLife.49375.
26. Wang, F., Shi, J., He, D., Tong, B., Zhang, C., Wen, A., Zhang, Y., Feng, Y., and Lin, W. (2020). Structural basis for transcription inhibition by E. coli SspA. Nucleic Acids Res. *48*, 9931–9942. https://doi.org/10.1093/nar/gkaa672.
27. Travis, B.A., Ramsey, K.M., Prezioso, S.M., Tallo, T., Wandzilak, J.M., Hsu, A., Borgnia, M., Bartesaghi, A., Dove, S.L., Brennan, R.G., and Schumacher, M.A. (2021). Structural Basis for Virulence Activation of Francisella tularensis. Mol. Cell *81*, 139–152.e10. https://doi.org/10.1016/j.molcel.2020.10.035.
28. Shi, J., Wen, A., Jin, S., Gao, B., Huang, Y., and Feng, Y. (2021). Transcription activation by a sliding clamp. Nat. Commun. *12*, 1131. https://doi.org/10.1038/s41467-021-21392-0.
29. Saecker, R.M., Chen, J., Chiu, C.E., Malone, B., Sotiris, J., Ebrahim, M., Yen, L.Y., Eng, E.T., and Darst, S.A. (2021). Structural origins of Escherichia coli RNA polymerase open

promoter complex stability. Proc. Natl. Acad. Sci. USA *118*. e2112877118. https://doi.org/10.1073/pnas.2112877118.

30. Wang, F., Feng, Y., Shang, Z., and Lin, W. (2021). A unique binding between SspA and RNAP β'NTH across low-GC Gram-negative bacteria facilitates SspA-mediated transcription regulation. Biochem. Biophys. Res. Commun. *583*, 86–92. https://doi.org/10.1016/j.bbrc.2021.10.048.
31. Lambert, L.J., Wei, Y., Schirf, V., Demeler, B., and Werner, M.H. (2004). T4 AsiA blocks DNA recognition by remodeling sigma70 region 4. EMBO J. *23*, 2952–2962. https://doi.org/10.1038/sj.emboj.7600312.
32. Patikoglou, G.A., Westblade, L.F., Campbell, E.A., Lamour, V., Lane, W.J., and Darst, S.A. (2007). Crystal structure of the Escherichia coli regulator of sigma70, Rsd, in complex with sigma70 domain 4. J. Mol. Biol. *372*, 649–659. https://doi.org/10.1016/j.jmb.2007.06.081.
33. Blanco, A.G., Canals, A., Bernués, J., Solà, M., and Coll, M. (2011). The structure of a transcription activation subcomplex reveals how σ70 is recruited to PhoB promoters. EMBO J. *30*, 3776–3785. https://doi.org/10.1038/emboj.2011.271.
34. Malhotra, A., Severinova, E., and Darst, S.A. (1996). Crystal Structure of a σ70 Subunit Fragment from E. coli RNA Polymerase. Cell *87*, 127–136. https://doi.org/10.1016/S0092-8674(00)81329-X.
35. Campbell, E.A., Muzzin, O., Chlenov, M., Sun, J.L., Olson, C.A., Weinman, O., Trester-Zedlitz, M.L., and Darst, S.A. (2002). Structure of the Bacterial RNA Polymerase Promoter Specificity σ Subunit. Mol. Cell *9*, 527–539. https://doi.org/10.1016/S1097-2765(02)00470-7.
36. Li, W., Stevenson, C.E.M., Burton, N., Jakimowicz, P., Paget, M.S.B., Buttner, M.J., Lawson, D.M., and Kleanthous, C. (2002). Identification and structure of the anti-sigma factor-binding domain of the disulphide-stress regulated sigma factor sigma(R) from Streptomyces coelicolor. J. Mol. Biol. *323*, 225–236. https://doi.org/10.1016/S0022-2836(02)00948-8.
37. Vishwakarma, R.K., Cao, A.-M., Morichaud, Z., Perumal, A.S., Margeat, E., and Brodolin, K. (2018). Single-molecule analysis reveals the mechanism of transcription activation in M. tuberculosis. Sci. Adv. *4*, eaao5498. https://doi.org/10.1126/sciadv.aao5498.
38. Dombroski, A.J., Walter, W.A., Record, M.T., Jr., Siegele, D.A., and Gross, C.A. (1992). Polypeptides containing highly conserved regions of transcription initiation factor sigma 70 exhibit specificity of binding to promoter DNA. Cell *70*, 501–512. https://doi.org/10.1016/0092-8674(92)90174-b.
39. Dombroski, A.J., Walter, W.A., and Gross, C.A. (1993). Amino-terminal amino acids modulate sigma-factor DNA-binding activity. Genes Dev. *7*, 2446–2455. https://doi.org/10.1101/gad.7.12a.2446.
40. Schwartz, E.C., Shekhtman, A., Dutta, K., Pratt, M.R., Cowburn, D., Darst, S., and Muir, T.W. (2008). A Full-Length Group 1 Bacterial Sigma Factor Adopts a Compact Structure Incompatible with DNA Binding. Chem. Biol. *15*, 1091–1103. https://doi.org/10.1016/j.chembiol.2008.09.008.
41. Callaci, S., and Heyduk, T. (1998). Conformation and DNA Binding Properties of a Single-Stranded DNA Binding Region of σ70 Subunit from Escherichia coli RNA Polymerase Are Modulated by an Interaction with the Core Enzyme. Biochemistry *37*, 3312–3320. https://doi.org/10.1021/bi972041m.
42. Callaci, S., Heyduk, E., and Heyduk, T. (1999). Core RNA Polymerase from E. coli Induces a Major Change in the Domain Arrangement of the σ70 Subunit. Mol. Cell *3*, 229–238. https://doi.org/10.1016/S1097-2765(00)80313-5.
43. Gruber, T.M., Markov, D., Sharp, M.M., Young, B.A., Lu, C.Z., Zhong, H.J., Artsimovitch, I., Geszvain, K.M., Arthur, T.M., Burgess, R.R., et al. (2001). Binding of the initiation factor sigma(70) to core RNA polymerase is a multistep process. Mol. Cell *8*, 21–31. https://doi.org/10.1016/S1097-2765(01)00292-1.
44. Konstantin, K., Leonid, M., Anita, N.-M.,L.,D.S., Dragana, R.,E.,N.B., Ann, H., Tomasz, H., and Konstantin, S. (2002). A Role for Interaction of the RNA Polymerase Flap Domain with the σ Subunit in Promoter Recognition. Science *295*, 855–857. https://doi.org/10.1126/science.1066303.
45. Feklistov, A., Barinova, N., Sevostyanova, A., Heyduk, E., Bass, I., Vvedenskaya, I., Kuznedelov, K., Merkienė, E., Stavrovskaya, E., Klimašauskas, S., et al. (2006). A Basal Promoter Element Recognized by Free RNA Polymerase σ Subunit Determines Promoter Recognition by RNA Polymerase Holoenzyme. Mol. Cell *23*, 97–107. https://doi.org/10.1016/j.molcel.2006.06.010.
46. Schneidman-Duhovny, D., Inbar, Y., Nussinov, R., and Wolfson, H.J. (2005). PatchDock and SymmDock: servers for rigid and symmetric docking. Nucleic Acids Res. *33*, W363–W367. https://doi.org/10.1093/nar/gki481.
47. Schneidman-Duhovny, D., Rossi, A., Avila-Sakar, A., Kim, S.J., Velázquez-Muriel, J., Strop, P., Liang, H., Krukenberg, K.A., Liao, M., Kim, H.M., et al. (2012). A method for integrative structure determination of protein-protein complexes. Bioinformatics *28*, 3282–3289. https://doi.org/10.1093/bioinformatics/bts628.
48. Braitbard, M., Schneidman-Duhovny, D., and Kalisman, N. (2019). Integrative Structure Modeling: Overview and Assessment. Annu. Rev. Biochem. *88*, 113–135. https://doi.org/10.1146/annurev-biochem-013118-111429.
49. Ha, T., Enderle, T., Ogletree, D.F., Chemla, D.S., Selvin, P.R., and Weiss, S. (1996). Probing the Interaction between Two Single Molecules: Fluorescence Resonance Energy Transfer between a Single Donor and a Single Acceptor. Proc. Natl. Acad. Sci. USA *93*, 6264–6268. https://doi.org/10.1073/pnas.93.13.6264.
50. Lerner, E., Cordes, T., Ingargiola, A., Alhadid, Y., Chung, S., Michalet, X., and Weiss, S. (2018). Toward dynamic structural biology: Two decades of single-molecule Förster resonance energy transfer. Science *359*, eaan1133. https://doi.org/10.1126/science.aan1133.
51. Kalisman, N., Adams, C.M., and Levitt, M. (2012). Subunit order of eukaryotic TRiC/CCT chaperonin by cross-linking, mass spectrometry, and combinatorial homology modeling. Proc. Natl. Acad. Sci. USA *109*, 2884–2889. https://doi.org/10.1073/pnas.1119472109.
52. Murakami, K., Elmlund, H., Kalisman, N., Bushnell, D.A., Adams, C.M., Azubel, M., Elmlund, D., Levi-Kalisman, Y., Liu, X., Gibbons, B.J., et al. (2013). Architecture of an RNA polymerase II transcription pre-initiation complex. Science *342*, 1238724. https://doi.org/10.1126/science.1238724.
53. Leitner, A., Joachimiak, L.A., Unverdorben, P., Walzthoeni, T., Frydman, J., Förster, F., and Aebersold, R. (2014). Chemical cross-linking/mass spectrometry targeting acidic residues in proteins and protein complexes. Proc. Natl. Acad. Sci. USA *111*, 9455–9460. https://doi.org/10.1073/pnas.1320298111.
54. Walker, R.G., Deng, X., Melchior, J.T., Morris, J., Tso, P., Jones, M.K., Segrest, J.P., Thompson, T.B., and Davidson, W.S. (2014). The Structure of Human Apolipoprotein A-IV as Revealed by Stable Isotope-assisted Cross-linking, Molecular Dynamics, and Small Angle X-ray Scattering. J. Biol. Chem. *289*, 5596–5608. https://doi.org/10.1074/jbc.M113.541037.
55. Lima, D.B., Melchior, J.T., Morris, J., Barbosa, V.C., Chamot-Rooke, J., Fioramonte, M., Souza, T.A.C.B., Fischer, J.S.G., Gozzo, F.C., Carvalho, P.C., and Davidson, W.S. (2018). Characterization of homodimer interfaces with cross-linking mass spectrometry and isotopically labeled proteins. Nat. Protoc. *13*, 431–458. https://doi.org/10.1038/nprot.2017.113.
56. Slavin, M., and Kalisman, N. (2018). Structural analysis of protein complexes by cross-linking and mass spectrometry. In Methods in Molecular Biology, J.A. Marsh, ed. (Springer), pp. 173–183. https://doi.org/10.1007/978-1-4939-7759-8_11.
57. Lerner, E., Chung, S., Allen, B.L., Wang, S., Lee, J., Lu, S.W., Grimaud, L.W., Ingargiola, A., Michalet, X., Alhadid, Y., et al. (2016). Backtracked and paused transcription initiation intermediate of Escherichia coli RNA polymerase. Proc. Natl. Acad. Sci. USA *113*, E6562–E6571.
58. Chung, S., Lerner, E., Jin, Y., Kim, S., Alhadid, Y., Grimaud, L.W., Zhang, I.X., Knobler, C.M., Gelbart, W.M., and Weiss, S. (2019). The effect of macromolecular crowding on single-round transcription by Escherichia coli RNA polymerase. Nucleic Acids Res. *47*, 1440–1450. https://doi.org/10.1093/nar/gky1277.
59. Harris, P.D., Narducci, A., Gebhardt, C., Cordes, T., Weiss, S., and Lerner, E. (2022). Multi-parameter photon-by-photon hidden Markov modeling. Nat. Commun. *13*, 1000. https://doi.org/10.1038/s41467-022-28632-x.
60. Kim, S., Streets, A.M., Lin, R.R., Quake, S.R., Weiss, S., and Majumdar, D.S. (2011). High-throughput single-molecule optofluidic analysis. Nat. Methods *8*, 242–245.
61. Mukhopadhyay, J., Kapanidis, A.N., Mekler, V., Kortkhonjia, E., Ebright, Y.W., and Ebright, R.H. (2001). Translocation of sigma(70) with RNA polymerase during transcription: fluorescence resonance energy transfer assay for movement relative to DNA. Cell *106*, 453–463. https://doi.org/10.1016/S0092-8674(01)00464-0.
62. Kahraman, A., Herzog, F., Leitner, A., Rosenberger, G., Aebersold, R., and Malmström, L. (2013). Cross-Link Guided Molecular Modeling with ROSETTA. PLoS One *8*, e73411. https://doi.org/10.1371/journal.pone.0073411.
63. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. Nature *596*, 583–589. https://doi.org/10.1038/s41586-021-03819-2.
64. Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., Yuan, D., Stroe, O., Wood, G., Laydon, A., et al. (2022).

AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. Nucleic Acids Res. *50*, D439–D444. https://doi.org/10.1093/nar/gkab1061.

65. Baek, M., DiMaio, F., Anishchenko, I., Dauparas, J., Ovchinnikov, S., Lee, G.R., Wang, J., Cong, Q., Kinch, L.N., Schaeffer, R.D., et al. (2021). Accurate prediction of protein structures and interactions using a three-track neural network. Science *373*, 871–876. https://doi.org/10.1126/science.abj8754.
66. Krissinel, E., and Henrick, K. (2007). Inference of Macromolecular Assemblies from Crystalline State. J. Mol. Biol. *372*, 774–797. https://doi.org/10.1016/j.jmb.2007.05.022.
67. BIOVIA, Dassault Systèmes (2020). Discovery Studio Visualizer (Dassault Systèmes).
68. Ryl, P.S.J., Bohlke-Schneider, M., Lenz, S., Fischer, L., Budzinski, L., Stuiver, M., Mendes, M.M.L., Sinn, L., O'Reilly, F.J., and Rappsilber, J. (2020). In Situ Structural Restraints from Cross-Linking Mass Spectrometry in Human Mitochondria. J. Proteome Res. *19*, 327–336. https://doi.org/10.1021/acs.jproteome.9b00541.
69. Gruber, T.M., and Gross, C.A. (2003). Multiple Sigma Subunits and the Partitioning of Bacterial Transcription Space. Annu. Rev. Microbiol. *57*, 441–466. https://doi.org/10.1146/annurev.micro.57.030502.090913.
70. Landini, P., Egli, T., Wolf, J., and Lacour, S. (2014). sigmaS, a major player in the response to environmental stresses in Escherichia coli: role, regulation and mechanisms of promoter recognition. Environ. Microbiol. Rep. *6*, 1–13. https://doi.org/10.1111/1758-2229.12112.
71. Bar-Nahum, G., and Nudler, E. (2001). Isolation and characterization of sigma(70)-retaining transcription elongation complexes from Escherichia coli. Cell *106*, 443–451. https://doi.org/10.1016/S0092-8674(01)00461-5.
72. Joron, K., Zamel, J., Dvir, S., Kalisman, N., and Lerner, E. (2022). The structural basis for the self-inhibition of DNA binding by apo-σ70 - smFRET raw data and analyses pipeline. https://doi.org/10.5281/zenodo.7173886.
73. Svetlov, V., and Artsimovitch, I. (2015). Purification of Bacterial RNA Polymerase: Tools and Protocols. In Bacterial Transcriptional Control: Methods and Protocols, I. Artsimovitch and T.J. Santangelo, eds. (Springer New York), pp. 13–29. https://doi.org/10.1007/978-1-4939-2392-2_2.
74. Laurence, T.A., Kong, X., Jäger, M., and Weiss, S. (2005). Probing Structural Heterogeneities and Fluctuations of Nucleic Acids and Denatured Proteins. Proc. Natl. Acad. Sci. USA *102*, 17348–17353. https://doi.org/10.1073/pnas.0508584102.
75. Müller, B.K., Zaychikov, E., Bräuchle, C., and Lamb, D.C. (2005). Pulsed Interleaved Excitation. Biophys. J. *89*, 3508–3522. https://doi.org/10.1529/biophysj.105.064766.
76. Ingargiola, A., Lerner, E., Chung, S., Weiss, S., and Michalet, X. (2016). FRETBursts: An Open Source Toolkit for Analysis of Freely-Diffusing Single-Molecule FRET. PLoS One *11*, e0160716.
77. Nir, E., Michalet, X., Hamadani, K.M., Laurence, T.A., Neuhauser, D., Kovchegov, Y., and Weiss, S. (2006). Shot-Noise Limited Single-Molecule FRET Histograms: Comparison between Theory and Experiments. J. Phys. Chem. B *110*, 22103–22124. https://doi.org/10.1021/jp063483n.
78. Chen, J., Zaer, S., Drori, P., Zamel, J., Joron, K., Kalisman, N., Lerner, E., and Dokholyan, N.V. (2021). The structural heterogeneity of α-synuclein is governed by several distinct subpopulations with interconversion times slower than milliseconds. Structure *29*, 1048–1064.e6. https://doi.org/10.1016/j.str.2021.05.002.
79. Chaudhury, P., Tripp, P., and Albers, S.-V. (2018). Expression, Purification, and Assembly of Archaellum Subcomplexes of Sulfolobus acidocaldarius. In Protein Complex Assembly: Methods and Protocols, J.A. Marsh, ed. (Springer), pp. 307–314. https://doi.org/10.1007/978-1-4939-7759-8_19.
80. Meng, E.C., Goddard, T.D., Pettersen, E.F., Couch, G.S., Pearson, Z.J., Morris, J.H., and Ferrin, T.E. (2023). UCSF ChimeraX: Tools for structure building and analysis. Protein Sci. *32*, e4792. https://doi.org/10.1002/pro.4792.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and virus strains		
E. *coli* (DE3) chemically competent cells	Sigma-Aldrich	CMC0014
Chemicals, peptides, and recombinant proteins		
vector for expression of single E. *coli* RNAP subunit (rpoD)	Addgene	#104399
vector for the expression of E. *coli* core RNAP	Addgene	#104398
Deposited data		
cross-linking mass spectrometry data	PRIDE	PXD037183
single molecule FRET data	Zenodo	https://doi.org/10.5281/zenodo.7173886
model structures cluster	PDB-Dev	PDBDEV_00000377
Software and algorithms		
multi-parameter hidden markov modeling code	Paul David Harris	https://doi.org/10.1038/s41467-022-28632-x
PatchDock	Dina Schneidman-Duhovny	https://doi.org/10.1093/nar/gki481
ChimeraX	Thomas Ferrin	https://doi.org/10.1002/pro.3943

RESOURCE AVAILABILITY

Lead contact

Further information and requests should be directed to and will be fulfilled by the lead contact, Eitan Lerner (eitan.lerner@mail.huji.ac.il).

Materials availability

The plasmid of the Cys-less σ^{70} from which we prepared the doubly-labeled σ^{70} variant for our smFRET measurements, was provided to us as a gift from the laboratory of Dr. Shimon Weiss, UCLA. Plasmid may be provided by the lead contact upon request. WT E. *coli* σ^{70} plasmid is available in Addgene under the accession code PIA586 and catalog number 104399, E. *coli* core RNAP is available in Addgene under the catalog number 104398. BL-21 (DE3) *E.coli* cells used for protein production are available at Sigma-Aldrich under product number CMC0014.

Data and code availability

- SmFRET data have been deposited at https://doi.org/10.5281/zenodo.7173886 [72] and are publicly available as of the date of publication. CL-MS data have been deposited at Proteomics Identification Database (PRIDE); The structural basis for the self-inhibition of DNA binding by apo-σ^{70} – CLMS data, with the following accession code (ebi.ac.uk; PRIDE: PXD037183).
- Model structures cluster – available at PDB-Dev as accession code PDB-Dev: PDBDEV_00000377.
- All data is publicly accessible. Additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

E. coli

The work presented here was done *in vitro* on σ^{70} using the WT σ^{70} plasmid expressed in E. *coli* (DE3) chemically competent cells and purified. For *in vivo* experiments we used E. *coli* (DE3) chemically competent cells expressing the WT σ^{70} and the core RNAP.

Statistical analysis

We perform statistical analysis on the data recovered from mass spectrometry to estimate the false detection rate of cross-links. For more details refer to method details – estimation of FDR.

METHOD DETAILS

Recombinant protein expression and purification

Plasmid encoding for the RNAP holoenzyme rpoD (σ^{70}) subunit was obtained from Addgene as a gift from Dr. Irina Artsimovitch (Addgene plasmid #104399).[73] Plasmid pET28b(+)_reverse_comp_rpoD(-C) expressing the *E. coli* σ^{70} without cysteines was a gift from the laboratory of Shimon Weiss at UCLA. Purification is performed following previously published protocol[27,73] with minor adjustments. Briefly, plasmids are transformed into *E.coli* BL-21 competent cells. A single colony of the transformed BL-21 cells is suspended in 10 mL of LB media containing 50 μg/mL of Kanamycin and incubated overnight at 37°C and 250 rpm. Overnight culture is then diluted (1:100) into a newly autoclaved 2 L erlenmeyer flask, containing 1 L of LB media and supplemented with 50 μg/mL of Kanamycin. Culture growth is monitored by periodically measuring $OD_{\lambda=600nm}$; once $OD_{\lambda=600nm}$ reached 0.6 (after ~4 hours), recombinant protein expression is induced with IPTG addition at a final concentration of 1 mM. After 3 hours of incubation at 37°C and 250 rpm, cells are harvested by centrifuging at 6,000 g and 4°C. The bacterial pellet is resuspended on ice in lysis buffer (50 mM HEPES-KOH (pH 7), 500 mM NaCl, 5% glycerol) containing 1 tablet of cOmplete™ EDTA-free protease inhibitors cocktail (Roche). Cell suspension is then supplemented with lysozyme to a final concentration of 1 mg/mL. Cells are then disrupted by ultrasonication on ice at 60% amplitude for 10-12 cycles (20 second pulses with 50 second intervals). Cell debris was removed by centrifuging at 12,000 rpm and 4°C for 30 minutes. Then for an additional 15 minutes after removing the pellet. Ni-Sepharose 4 mL column is used to separate the His-tagged σ^{70} from the supernatant, with all washings and elution performed at 4°C. Fractions suspected of containing σ^{70} are collected and run in SDS-PAGE for purity assessment. Pure fraction containing high concentration of σ^{70}, are combined, and further purified by running through Resource 15Q 1 mL column for ion exchange chromatography. Pure fractions are assessed by running SDS-PAGE and Coomassie staining. Then, pure fractions are dialyzed against dialysis buffer containing (20 mM HEPES-KOH (pH 7.5), 200 mM NaCl, 0.2 mM DTT and protease inhibitor cocktail). Protein concentration is then determined, and protein is stored in 50% glycerol at 20°C.

σ^{70} mutagenesis and double labeling

Point mutations are introduced to the Cys-less *E. coli* σ^{70} plasmid, pET28b(+)_reverse_comp_rpoD(-C), at positions 442 and 579 following a 2-step PCR protocol[73] and using a set of specific primers. We rely on previous characterizations, suggesting that these regions and specifically these two residues are solvent accessible.[37] Mutagenesis results are confirmed by Sanger sequencing and comparison. Purification of the mutant σ^{70} containing two cysteines at positions 442 and 579, is performed similarly to the *wt* σ^{70} purification described above. For labeling, mutant σ^{70} is dialyzed against (50 mM HEPES-KOH pH=7.2 and 0.5 mM EDTA) overnight at 4°C to reduce the glycerol concentration. After dialysis, a final concentration of 1.25 μM mutant or *wt* σ^{70} are suspended in (50 mM HEPES-KOH pH=7.2, and 0.1 mM EDTA) at final volume of 1 mL with 5 μM of Tris(2-carboxyethyl) phosphine (TCEP), to reduce any disulfide bridges without introducing free reacting thiols, for a more efficient labeling. The reaction is stirred for 45 minutes at room temperature. ATTO 643 and ATTO 550, with linkers to an iodoacetamide group, are then added to a final concentration of 1.25 μM, in a dark room and the reaction is stirred at room temperature for 4-5 hours. The reaction is then transferred to 4°C overnight. To terminate the reaction, β-mercaptoethanol is added to a final concentration of 5 μM at ~1 mL final volume and incubated with stirring for 30 minutes at room temperature. The sample is then dialyzed against (50 mM HEPES-KOH (pH=7.2), and 0.1 mM EDTA) overnight at 4°C, and stored at -20°C with 50% glycerol.

RNAP holoenzyme transcription assay

While in the absence of an RNA transcript the ssDNA FRET probe shows a mean FRET efficiency of ~0.5 (Figure S1B), once a transcript is introduced the ssDNA FRET probe undergoes hybridization, which stretches the DNA probe, leading to a decrease of the FRET efficiency to ~0.0 (Figure S1D). Therefore, the decrease in the FRET efficiency indirectly indicates the presence of transcripts that were successfully synthesized, hence active transcription complex and proper binding of σ^{70} to both RNAP and promoter.

SmFRET measurements and data analysis

SmFRET measurements are performed on dually labeled mutant σ^{70} on a confocal-based microscopy setup (ISS™, Champaign, IL, USA) assembled on top of an Olympus IX73 inverted microscope stand (Olympus, Tokyo, Japan). Samples are measured in a high glass bottom μ-slides (Ibidi), with an acquisition time of 1 hour per technical repeat at room temperature (23°C). We use 532±1 nm (FL-532-PICO, CNI, China) and a 640±1 nm (QuixX ® 642-140 PS, Omicron, GmbH) pulsed picosecond fiber and diode lasers, respectively, operating at 20 MHz repetition rate. Scattering and fluorescence photons return in the excitation path, collected through the same objective. Then, while scattering photons are reflected, fluorescence photons are transmitted through a major dichroic mirror with high reflectivity at 532 nm and 640 nm (ZT532/640rpc, Chroma, Bellows Falls, Vermont, USA) and is focused with an achromatic lens (25 mm Diameter x 100 mm FL, VIS-NIR Coated, Edmund Optics) onto a 100 μm diameter pinhole (variable pinhole, motorized, tunable from 20 μm to 1 mm, custom made by ISS™), and then re-collimated with another achromatic lens (AC254-060-A, Thorlabs). Fluorescence originating from either donor or acceptor dyes is then split into two detection channels using a 605 nm cutoff dichroic mirror, then further cleaned using a 698/70 nm bandpass filter for acceptor emission and a 585/40 nm bandpass filter for donor emission. Single fluorescence photons are then detected using hybrid PMTs (Model R10467U-40, Hamamatsu, Japan), and single photon detection event time tags are collected using a time-correlated single-photon counting card (SPC 150N, Becker & Hickl, GmbH). All measurements are performed in the same buffer (20 mM HEPES-KOH at pH 7.0, 50 mM KCl, 10 mM $MgCl_2$, 2 mM β-mercaptoethanol). SmFRET experiments are performed using nanosecond alternating laser excitation, nsALEX,[74]

also known as pulsed interleaved excitation, PIE[75]), which provides information on the excitation origin of the donor and acceptor photons. Data is then analyzed using FRETbursts[76] and mpH^{2}MM.[59] A burst identified using a dual channel burst search[77] is considered a single molecule event, only if for each consecutive 10 photons a count rate of at least 16 times higher than the background rate, and only if it includes at least 30 photons originating from donor excitation and at least 30 photons originating from acceptor excitation.

In vitro smFRET-based transcription activity of the purified recombinant RNAP

The smFRET based transcription activity assay is performed as previously mentioned.[57,58] Briefly, The RNA polymerase is incubated in KG7 buffer (40 mM HEPES (KOH), pH=7.0, 10 mM $MgCl_2$, 1 mM DTT, 5% glycerol, supplemented with 1 mM TROLOX and 10 mM MEA) with the mutant-σ^{70} (442C, 579C) at 37°C for 30 minutes to form the holoenzyme complex. Next, linear dsDNA *lacCONS* promoter, which promotes the synthesis of a nascent RNA transcript with 20 adenine (20A; Figure S1A) bases at the 3′-end, is introduced to the holoenzyme complex at 37°C to form the RNAP-promoter open complex. Later, all nucleotides are introduced together with RNAse inhibitor (BioLabs), and the sample is again incubated at 37°C. Guanidinium-HCl is then added at room temperature (25°C) to quench the reaction. Then, a ssDNA probe consisting of 20 deoxy-thymine (20dT) and labeled by a donor dye (ATTO 488) at the 5′-end, and an acceptor dye (ATTO 647N) at the 3′-end, is used for detecting nascent RNA transcripts via hybridizing to them, inducing a reduction in the end-to-end FRET values relative to that in the free ssDNA FRET probe. In this manner, the probe is able to produce FRET signals as long as it is not degraded. If the probe is degraded the FRET signal efficiency would be reduced to E=0. A final concentration of 50 pM of the probe is incubated in 100 μL of the prepared sample. If RNA transcripts are produced, we expect to detect a decrease in the FRET efficiency due to hybridization of the 20A bases of the nascent RNA transcript to the 20dT bases of the probe.

In vitro cross-linking and mass-spectrometry preparation

In vitro CL-MS is performed as previously described.[51,78] Briefly, BS^3 powder is used to prepare a 10 mM stock solution using HEPES buffer (20 mM HEPES (KOH) pH=7.0, 50 mM KCl, 10 mM $MgCl_2$ and 2 mM 2-mercaptoethanol). Similarly, DMTMM is prepared at a 70 mM stock solution. For cross-linking, a final concentration of 1 mM BS^3 or 7 mM DMTMM is incubated with 5-10 μg of *wt* σ^{70} at 30°C for 1.5 hours with shaking at 600 rpm. Next, to terminate the reaction, ammonium bicarbonate is added at three times the concentration of the crosslinker from a 1 M stock solution (so the sample is not diluted), and incubated at 25°C for 30 minutes with shaking. Later, preparation of the samples for mass spectrometry and mass spectrometry RAW data files analysis, is performed as previously described[78,79] with minor adjustments, see below the section on *crosslinks identification and estimation of* false detection rate (*FDR*).

In vivo cross-linking and mass-spectrometry preparation

Plasmid encoding for the RNAP holoenzyme rpoD (σ^{70}) subunit was obtained from Addgene as a gift from Dr. Irina Artsimovitch (Addgene plasmid #104399).[73] Plasmid encoding for the core RNAP holoenzyme was obtained from Addgene as a gift from Dr. Irina Artsimovitch (Addgene plasmid #104398).[73] After transformation and isolation of a single colony containing the plasmid as described above (Recombinant protein expression and purification), an overnight culture is diluted (1:100) into a newly autoclaved 250 mL erlenmeyer flask, containing 100 mL of LB media and supplemented with 50 μg/mL of kanamycin. The flask is incubated at 37°C with shaking until $OD_{\lambda=600nm}$ reaches 0.6, then IPTG is added, and the flask is incubated again at 37°C. After 3 hours a bacterial pellet from 5 mL of bacterial culture is collected and stored at -80°C. The flask is incubated again at 28°C with shaking. After overnight incubation, bacterial growth was closely monitored by measuring $OD_{\lambda=600nm}$ every 15 minutes. Once $OD_{\lambda=600nm}$ values stop elevating (stationary phase; monitored periodically using a spectrophotometer) a bacterial pellet from 5 mL of bacterial culture is collected and stored at -80°C. For the cross-linking reaction a 250 mM stock solution of DSS is prepared in DMSO. The collected bacterial pellets are resuspended in lysis buffer (50 mM HEPES (KOH) pH=7.0, 500 mM NaCl and 5% glycerol) containing 10 mM DSS to increase membrane permeability. First, DSS is resuspended from the stock solution in 1 mL of lysis buffer, since DSS is not completely soluble in water it forms a precipitate, and we take only the dissolved fraction which is ~1 mL of lysis buffer supplemented with a final concentration of ~10 mM DSS. The cross-linking reaction is then incubated at 30°C for 20 minutes with shaking. Later, the cross-linking reaction is quenched by adding 40 mM final concentration of ammonium bicarbonate and incubating at room temperature (~25°C) with shaking for another 20 minutes. Centrifuge at 6,000 g and 4°C for 5 minutes to collect the cells, pellets are then stored at -80°C. Pellets are resuspended in 1 mL of lysis buffer and lysed by sonication using the Qsonica 422 Ultrasonic probe at 60% Amp for 12 cycles of 10 seconds ON and 25 seconds intervals (OFF). Centrifuge to discard cell debris, keep the supernatant. Ni-Sepharose high performance beads (GE healthcare) are washed thoroughly with lysis buffer to rinse off the ethanol. The supernatant is then incubated with 10 μL of the Ni beads and incubated at 4°C with minimal shaking (to prevent sinking of the beads) for 4 hours. Following, samples are centrifuged to take out the flow-through, and the beads are resuspended with wash buffer (lysis buffer + 20 mM Imidazole), centrifuge at 950 g and 4°C for 2 minutes and take out the liquid, the washing process is repeated three times. For elution, the beads are then resuspended in 50 μL of elution buffer (lysis buffer with 400 mM Imidazole), and kept at room temperature for 30 minutes while inverting the tubes every 2 minutes to prevent beads from sinking to the bottom. Centrifuge at 950 g and 4°C for 2 minutes and collect the liquid without disrupting the beads. Then samples are prepared for mass spectroscopy as previously described.[78,79] Briefly, the protein is precipitated in 1 mL of acetone (-80°C) for 1 hour, followed by centrifugation at 14,000 g. The pellet is resuspended in 20 mL of 8 M urea. The urea is diluted by adding 200 mL of digestion buffer (25 mM TRIS, pH = 8.0; 10% Acetonitrile). We add 0.5 mg of trypsin (Promega, Madison, Wisconsin) to the diluted urea and digest the protein overnight at 37°C under agitation. Following digestion, the peptides are desalted on C18 stage-tips and eluted by 55% acetonitrile. The eluted peptides are dried in a SpeedVac, reconstituted in 0.1% formic acid, and measured in the mass spectrometer.

Mass spectrometry analysis

The samples are analyzed by a 120-min 0-to-40% acetonitrile gradient on a liquid chromatography system coupled to a Q-Exactive HF mass-spectrometer. The analytical column is an EasySpray 25 cm heated to 40°C. The method parameters of the run are as follows: Data-Dependent Ac-quisition; Full MS resolution 70,000; MS1 AGC target 1e6; MS1 Maximum IT200 ms; Scan range 450 to 1,800; dd-MS/MS resolution 35,000; MS/MS AGCtarget 2e5; MS2 Maximum IT 600 ms; Loop count Top 12; Isolation window1.1; Fixed first mass 130; MS2 Minimum AGC target 800; Peptide match - off; Exclude isotope - on; Dynamic exclusion 45 seconds. Each cross-linked sample is measured twice in two different HCD energies (NCE): 26, and stepped 25, 30, and 35. All cross-linked samples are measured with the following charge exclusion: unassigned,1,2,3,8,>8. Proteomics samples are measured with the following charge exclusion: unassigned,1,8,>8.

Cross-links identification and filtration

The RAW data files are converted to MGF using Proteome Discoverer (Thermo). Then, FindXL[51] is used to exhaustively enumerate all the possible peptide pairs originating from BS^3 or DMTMM crosslinks, with the following search parameters: i) Sequence database - σ^{70} sequence taken from UniProt (P00579); ii) Protease – trypsin, allowing up to three mis-cleavage sites; iii) Variable modifications: methionine oxidation, lysine with hydrolyzed mono-link; iv) Cross-linking must occur between two lysine residues or lysine and glutamic acid/aspartic acid for (BS^3/DSS or DMTMM, respectively); v) Cross-linker is not cleaved; vi) MS/MS fragments to consider: b-ions, y-ions; vii) MS1 tolerance – 6 ppm; viii) MS2 tolerance – 8 ppm; and ix) Cross-linker mass – one of three possible masses: 138.0681, 138.0681 + 1.00335, and 138.0681 + 2.0067 for the BS^3/DSS and -18 Da for DMTMM which represent the release of a water molecule. The three masses address the occasional incorrect assignment of the mono-isotopic mass by the mass spectrometer.

Estimation of FDR

The FDR is estimated by repeating the cross-link identification analysis 20 times with an erroneous cross-linker mass of 138.0681 × N/138 Da (for BS^3 and DSS), where N = 160, 161, 162, ... 179. This leads to bogus identifications with fragmentation scores that are generally much lower than the scores obtained with the correct cross-linker mass. For the identification of true cross-links, we set the threshold on the fragmentation score according to the desired FDR value (Figure S5). For DMTMM, same as mentioned above with the erroneous cross-linker mass of -18.01056 × N/18 Da (Figures S5–S7). With the fragmentation scores 1 and 2 set to 2.0 and 0.7, respectively, we get an FDR of 0.65% for BS3 (Figure S5), 1.6% for DMTMM (Figure S6) and 0.55% for DSS (Figure S7) cross-links. The fragmentation score suggests how many times a certain peptide was detected. We filter cross-links above a threshold fragmentation score of 0.7, where the FDR decreases dramatically, making sure we do not include false cross-links.

Structure modeling using AlphaFold and RosettaFold

Using AlphaFold and RosettaFold we submitted a single sequence of the entire σ^{70}, and got an output of different structure models ranked according to the scoring parameters of each algorithm. However, using ColabFold version 1.5.2, a protein complex structure prediction using AlphaFold2-multimer version 3 (v3), we submitted each region of the σ^{70} as a separate sequence, allowing AlphaFold2-multimer v3 to reorganize the structure of apo-σ^{70} from rigid domains while taking into account the high flexibility of the linkers. Using MMseqs2 the code was able to generate sequence alignments, which created the data library that Colabfold used to predict the structure model. We used the "unpaired_paired" pair mode to pair sequences from the same species, but also from different species as well. This is possible due to the high conservation rate between housekeeping σ factors of different species. We performed three recycles of the run to increase the validity of the resultant structure models.

PatchDock CL-MS restraint driven integrative docking

We use the *E. coli* RNAP holoenzyme PDB structure PDB: *6P1K*,[25] to isolate the σ^{70} subunit from the complex using *ChimeraX*.[80] Then, we separate the different σ^{70} regions into multiple PDB files. According to the sequence given by the PDB structure, σR1.1 ranges between amino acids (aa) 1-96, σR1.2 between aa 97-141, σNER between aa 142-355, σR2 between aa 356-453, σR3 between aa 454-548, σR4 between aa 549-616. Next, we perform step-by-step integrative structural docking in a step-wise fashion to include all six σ^{70} regions (σR1.2, σNER, σR2, σR3, σR4) one by one, starting with σNER and σR2 serving as the base structure. During each step of docking, in addition to the C_α-C_α distance ranges between cross-linked residues, for the sake of the proper reconnection of the backbone of the polypeptide chain, we assign a distance equivalent to a covalent bond between the N- and C-termini of consecutive regions along the primary sequence of the protein. In PatchDock, we input the two PDB files we wish to dock together with a list of restraints derived from the cross-linking data and assign the larger PDB file as the *receptor* and the other PDB file as the *ligand*. In our case, σNER and σR2 as a complex is the first *receptor* and σR3 is the first *ligand*, the result of this docking step will be the *receptor* in the next docking step where σR4 is the *ligand*, and so on. In the first stage PatchDock computes the molecular surface and applies a segmentation algorithm to detect geometric patches (concave, convex and flat surfaces). In the next stage, Patchdock applies a hybrid of geometric hashing and pose-clustering matching techniques to match concave patches with convex and flat patches with any type of patches. In the final stage, candidate structure models are examined, and ones with unacceptable penetrations of the atoms are discarded. Finally, the remaining possible solutions are ranked according to a geometric shape complementarity score and CL-MS data[46] (Data S2). Top scoring cluster representative is used to continue to the next docking step. In docking σNER and σR2 with σR3, a model is filtered if it does not surpass at least 81% agreement across all cross-links. To the top-ranking structure of σNER-σR2-σR3 we add

σR4, models are then filtered if not surpass at least 86% agreement across all cross-links. We continue the process by extracting the top-ranking structure model of the cluster after filtration. A structure model is then filtered if it does not surpass at least 71% agreement across all cross-links.

Western-blot

In vivo cross-linking and lysis were performed as described above. The lysate was incubated with Ni-beads for 4 hours at 4°C. Next the solution was centrifuged at 800 g to precipitate the Ni-beads binding σ^{70}, and the supernatant was separated from the precipitate. Then the beads were washed three times with wash buffer to get rid of all nonspecific Ni binders. Finally, the σ^{70} was eluted using elution buffer. The elution was run in a 10% acrylamide SDS-PAGE and transferred to a PVDF membrane. Detection of σ^{70} was performed by incubating the membrane at 4°C overnight in a 0.5 μg primary antibody solution, mouse IgG2b from Bio-legend, followed by an alkaline phosphatase tagged secondary antibody.