Published online 5 July 2012 *Nucleic Acids Research, 2012, Vol. 40, No. 18* ***9139–9152***
doi:10.1093/nar/gks661

A common feature from different subunits of a homomeric AAA+ protein contacts three spatially distinct transcription elements

Nan Zhang, Nicolas Joly* and Martin Buck*

Division of Biology, Sir Alexander Fleming Building, Imperial College London, Exhibition Road, London, SW7 2AZ, UK

Received March 29, 2012; Revised June 12, 2012; Accepted June 13, 2012

ABSTRACT

Initiation of σ^{54}-dependent transcription requires assistance to melt DNA at the promoter site but is impeded by numerous protein–protein and nucleo–protein interactions. To alleviate these inhibitory interactions, hexameric bacterial enhancer binding proteins (bEBP), a subset of the ATPases associated with various cellular activities (AAA+) protein family, are required to remodel the transcription complex using energy derived from ATP hydrolysis. However, neither the process of energy conversion nor the internal architecture of the closed promoter complex is well understood. *Escherichia coli* Phage shock protein F (PspF), a well-studied bEBP, contains a surface-exposed loop 1 (L1). L1 is key to the energy coupling process by interacting with Region I of σ^{54} ($\sigma^{54}{}_{RI}$) in a nucleotide dependent manner. Our analyses uncover new levels of complexity in the engagement of a multimeric bEBP with a basal transcription complex via several L1s. The mechanistic implications for these multivalent L1 interactions are elaborated in the light of available structures for the bEBP and its target complexes.

INTRODUCTION

In bacteria, the multi-subunit core RNA polymerase (RNAP or E) catalyses transcription and is directed to the promoter DNA by association with a sigma factor (σ). Bacterial σ factors fall into two classes: σ^{70} binds to the consensus sequences at –35 (TTGACA) and –10 (TATAAT), whereas σ^{54} binds to the consensus sequences at –24 (GG) and –12 (GC). In σ^{70}-dependent transcription, RNAP forms a closed complex (RP_C) on the promoter that can spontaneously isomerize to an open complex (RP_O) (1). In σ^{54}-dependent transcription, the isomerization from RP_C to RP_O is energetically unfavourable due to the presence of a stably engaged upstream fork junction DNA around the –12 site. Within the stable RP_C, the –12 fork junction is evident (2), although the downstream DNA melting has not occurred and the +1 transcription start site is outside of the RNAP at this stage (3). The stable RP_C is thought to be preceded by an unstable RP_C in which the –12 fork junction has yet to form (4). The isomerization from RP_C to RP_O requires ATP hydrolysis by bacterial enhancer binding proteins (bEBPs), resembling in part the eukaryotic Pol II system that utilizes TFIIH and ATP for DNA melting (5). σ^{54}-dependent transcription not only regulates various adaptive responses (6,7), but is also responsible for regulating pathogenesis determinants in disease-causing agents such as *Borrelia burgdorferi* (the agent of Lyme disease) and *Vibrio cholera* (the agent of epidemic diarrheal disease) (8,9). Therefore, an understanding of the σ^{54}-transcription pathway is valuable for identification of new antibacterial drug targets (10).

Current information on the organization of σ^{54}-transcription complexes have been drawn from: the low-resolution Cryo-electron microscopy (Cryo-EM) studies with purified (E)σ^{54}–bEBP complexes (11,12), the NMR and SAXS structures of some regions of σ^{54} (13–16) and crystal studies of bEBPs (12,17–20). Three regions have been identified in σ^{54} (Figure 1B). Region I of σ^{54} ($\sigma^{54}{}_{RI}$) interacts with bEBPs, core RNAP and the –12 promoter region (21–23), participating in promoter melting and isomerization processes (24,25). Region II of σ^{54} ($\sigma^{54}{}_{RII}$) is dispensable for interactions with RNAP and DNA. Region III of σ^{54} ($\sigma^{54}{}_{RIII}$) contains several functionally important modules, including the RpoN box required for the recognition of the –24 promoter element (22).

bEBPs belong to Clade 6 of the ATPases associated with various cellular activities (AAA+) protein family

*To whom correspondence should be addressed. Tel: +33 157 278 023; Email: joly@ijm.univ-paris-diderot.fr
Correspondence may also be addressed to Martin Buck. Tel: +44 2075 945 443; Fax: +44 2075 045 419; Email: m.buck@imperial.ac.uk
Present address:
Nicolas Joly, Institut Jacques Monod, CNRS UMR 7592, Universite Paris Diderot, Batiment Buffon, 15 rue Helene Brion, 75205 Paris cedex 13, France.

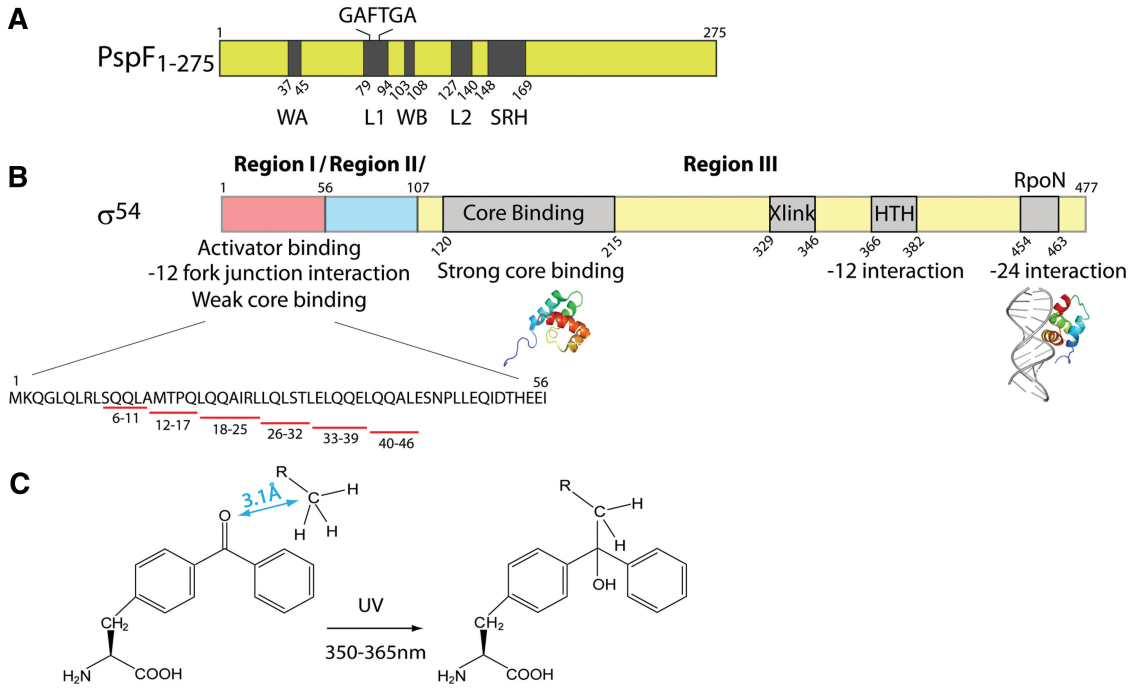

Figure 1. Sequence organization and functional properties of $PspF_{1–275}$ and σ^{54}. (**A**) Domain organization of $PspF_{1–275}$. WA stands for Walker A motif, L1 for Loop 1 (containing the 'GAFTGA' motif), WB for Walker B motif and SRH for second region of homology. (**B**) Domain organization of σ^{54}. HTH stands helix-turn-helix motif. Xlink stands for DNA cross-linking region. σ^{54} Regions I–III are separated by slashes. Available structures of two σ^{54} fragments are depicted under the corresponding sequences. Six σ^{54} Region I fragments (underlined in red) are generated for the following binding assays. (**C**) The rationale of the *p*Bpa-based UV cross-linking assay. The ketone group in the *p*Bpa artificial amino acid cross-links to any C–H bond within 3.1 Å under UV irradiation.

(26), and their functionality is dependent on an ability to self-associate, typically to form hexamers. The biochemical functions of bEBPs have been widely studied using as examples the *Escherichia coli* Phage shock protein F (PspF), which is involved in membrane stress responses (27), and the nitrogen control proteins NtrC and NtrC1. PspF contains an AAA+ domain (residues 1–275, $PspF_{1–275}$, Figure 1A) essential for oligomerization and ATP hydrolysis and a C-terminal DNA-binding domain. Two surface-exposed loops of PspF (Figure 1A), namely Loop 1 (L1) and Loop 2 (L2), are particularly important for coupling ATP hydrolysis to the RP_C remodelling event. L1 contains a highly conserved 'GAFTGA' motif amongst bEBPs. Residues Phe (F85) and Thr (T86) in the PspF 'GAFTGA' motif interact with $\sigma^{54}{}_{RI}$ during RP_O formation (21,28). Structural modelling indicates that L1 may have additional roles (11). These roles may be accommodated by the potential availability of up to six L1s across a PspF hexamer (11,12,29).

To seek evidence for functional specialization amongst L1s, we incorporated a photoreactive artificial amino acid, *p*-benzoyl-L-phenylalanine (*p*Bpa), using an orthogonal tRNA/tRNA synthetase pair to each PspF L1 'GAFTGA' position for identification of potential interacting partners (30); *p*Bpa can cross-link to any C–H bond within 3.1 Å upon UV irradiation [Figure 1C, (31,32)]. Here we provide direct evidence that: (i) L1 contacts the DNA non-template strand immediately upstream of the –24 promoter element, and (ii) the DNA immediately upstream of the –24 element is important for the isomerization from RP_C to RP_O as well as formation of one transcription intermediate. Using a fragmentation approach, we were able to identify two previously unknown PspF L1-binding patches within $\sigma^{54}{}_{RI}$ (residues 18–25 and 33–39). The above observations provide evidence that L1 is multifunctional, and makes at least three distinct nucleotide-dependent interactions within its target complex in driving RP_O formation.

MATERIALS AND METHODS

Plasmids

Plasmid pPB1 [encoding the *E. coli* $pspF_{1–275}$ sequence, (21)] was used as a template for the subsequent site-directed mutagenesis studies. Each 'GAFTGA' position was mutagenized in the context of pPB1 to an amber stop codon (TAG) to yield pET28b-$pspF_{1–275\ variant}$ plasmids (Supplementary Table S1).

DNA probes and peptide fragments

The linear DNA probes used in this study are summarized in Supplementary Table S2. The $\sigma^{54}{}_{RI}$ fragments were purchased with the highest purity level from Insight Biotechnology.

Protein expression and purification

The expression of $PspF_{1–275}$ *p*Bpa variants depends on two plasmids: (i) the pET28b-*pspF*$_{1–275\ variant}$ (Supplementary Table S1) and (ii) the pDULE-*p*Bpa [encoding the *Methanococcus jannaschii* tRNA/tRNA synthetase pair to specifically charge the intrinsic stop codon with *p*Bpa, (30)]. Typically, 0.26 g of *p*Bpa (Bachem) were dissolved under alkaline conditions and added to a L1 culture. The $PspF_{1–275}$ *p*Bpa variants were expressed and purified as previously described (28), treated with thrombin to remove the $(His)_{x6}$ tag, and stored in TGED buffer 1 (20 mM Tris-HCl pH 8.0, 50 mM NaCl, 1 mM DTT, 0.1 mM EDTA and 5% glycerol) at −80°C.

Klebsiella pneumoniae σ^{54} was purified as previously described and stored in TGED buffer 2 (20 mM Tris-HCl pH 8.0, 200 mM NaCl, 1 mM DTT, 0.1 mM EDTA and 50% glycerol) at −80°C (33). *K. pneumoniae* heart muscle kinase (HMK) tagged full-length σ^{54} (HMK-σ^{54}_{FL}) and HMK-tagged σ^{54} fragments (HMK-σ^{54}_{RI} and HMK-$\sigma^{54}_{\Delta RI}$) were purified and radio-labelled as previously described (34). *E. coli* core RNAP was purchased from Cambio.

ATPase activity assay

Typically in a 10 µl volume, 4 µM $PspF_{1–275}$ was pre-incubated with the ATPase buffer (20 mM Tris-HCl pH 8.0, 50 mM NaCl, 15 mM $MgCl_2$, 0.1 mM EDTA, 10 µM DTT) at 37°C for 5 min. ATP hydrolysis was initiated by addition of 1 mM unlabelled ATP and 0.6 µCi/µl [α-^{32}P] ATP (3000 Ci/mmol) and incubated for various time spans at 37°C. Reactions were quenched by addition of 5 volumes of 2 M formic acid. The [α-^{32}P] ADP was separated from the [α-^{32}P] ATP by thin layer chromatography (Macherey–Nagel) in 0.4 M K_2HPO_4/0.7 M boric acid. Radioactivity was scanned by PhosphoImager (Fuji Bas-1500) and analysed by Aida software. The ATP turnover rate (k_{cat}) of each $PspF_{1–275}$ *p*Bpa variant was expressed as a percentage of $PspF_{1–275}$ wild type (WT) activity. All experiments were minimally performed in triplicate.

Native gel mobility shift assay

Reactions were performed in 10 µl volumes and supplemented with 1 µM ^{32}P-HMK-σ^{54}_{FL} (or its fragments), ± 0.3 µM core RNAP, ± 50 nM radio-labelled DNA, 5 mM NaF and 4 mM nucleotides (ATP, ADP or AMP) in STA buffer [2.5 mM Tris-acetate pH 8.0, 8 mM Mg-acetate, 10 mM KCl, 1 mM DTT, 3.5% (w/v) PEG 8000] at 37°C for 5 min. Ten µM $PspF_{1–275}$ and 0.4 mM $AlCl_3$ were added for a further 15 min incubation to allow 'trapped' complex formation at 37°C. Complexes were either analysed on native gels or subject to UV cross-linking and then analysed.

Gel filtration assay

$PspF_{1-275}$ WT or the ADP–AlF_x 'trapped' complex was pre-incubated at 4°C with gel filtration buffer (20 mM Tri-HCl pH 8.0, 50 mM NaCl, 15 mM $MgCl_2$) for 5 min. A Superdex 200 column (10/30, 24 ml, GE Healthcare) assembled on the AKTA system (GE Healthcare) was equilibrated with buffer. Chromatography was carried out at a flow rate of 0.5 ml/min at 4°C.

*p*Bpa-based UV cross-linking assay

'Trapped' complexes were formed with either ^{32}P-HMK-σ^{54} (or its fragments) or ^{32}P-DNA. Reaction mixtures were UV irradiated at 365 nm on ice for 5 min, 15 min and 30 min then analysed on both native and SDS PAGE gels. The cross-linked protein–protein or nucleo–protein species were scanned by a Fuji PhosphoImager and analysed by Aida software.

Proteinase K-ExoIII footprinting assay

The UV cross-linked nucleo–protein species were generated with ^{32}P-DNA and subject to Proteinase K-ExoIII footprinting assays as previously described (35). The UV-irradiated samples (20 µl) were digested with 1 µl of 20 mg/ml Proteinase K (Sigma) at 37°C for 1 h to remove the protein components. ^{32}P-DNA containing the *p*Bpa peptide was phenol-extracted and isopropanol-precipitated. Twenty units of exonuclease III (ExoIII, USB) were added to the DNA sample to a 10 µl final volume. The ExoIII digestion proceeded for various time spans before being quenched by 4 µl of 3X formamide stop dye (3 mg xylene cyanol, 3 mg bromophenol blue, 0.8 ml 250 mM EDTA, 10 ml deionised formamide in 10 ml). The reaction mixtures were heated at 97°C for 5 min before separated on a sequencing gel.

In vitro RP_O formation assay

Open complex formation was measured in 10 µl final volumes containing: 4 µM $PspF_{1–275}$, 100 nM holoenzyme (1:4 ratio of E: σ^{54}), 20 U RNase inhibitor, 5% (v/v) glycerol, 4 mM dATP and 20 nM *Sinorhizobium meliloti nifH* promoter (Supplementary Table S2) in STA buffer at 37°C. Transcription was activated for various lengths of time before 0.5 mM dinucleotide primer UpG, 0.2 µCi/µl [α-^{32}P GTP] (3000 Ci/mmol) and 0.2 mg/ml heparin were added. After extension at 37°C for 10 min, the reaction mixtures were quenched by addition of 4 µl of 3X formamide stop dye and run on a sequencing gel. The activator-bypass activities of the σ^{54} variants were examined in a similar experimental procedure without the addition of $PspF_{1–275}$ activators and dATP.

In vitro spRNA assay

The ADP–AlF_x 'trapped' complexes were initially formed on the late-melted –10–1/WT DNA probe. Without the addition of dATP, the ADP–AlF_x complexes were allowed to synthesize a UpG-primed RNA product UpGpGpG (the spRNA) in the presence of α-^{32}P GTP in a manner similar to the RP_O formation assay as described above.

RESULTS

$PspF_{1-275}$ G83*p*Bpa can cross-link to σ^{54} but not to core RNAP

The photo-reactive artificial amino acid *p*Bpa was incorporated at each L1 'GAFTGA' position, generating

six $PspF_{1–275}$ *p*Bpa variants (G83*p*Bpa, A84*p*Bpa, F85*p*Bpa, T86*p*Bpa, G87*p*Bpa and A88*p*Bpa). Characterization of each *p*Bpa variant for bEBP functions is summarized in Table 1. Position 83 in the PspF L1 'GAFTGA' motif (the position of interest, as the following experiments demonstrate) was substituted with Ala (to remove the side chain) and Phe (to mimic the *p*Bpa cross-linker). The resulting G83A and G83F variants were also characterized (Table 1).

To investigate whether introducing the bulky hydrophobic *p*Bpa cross-linker could affect the overall L1 exposure and the local σ^{54} interaction, we performed the well-established 'trapping' assay. The transient L1 'GAFTGA'– σ^{54} interaction is stabilized in the presence of an ATP transition state analogue, ADP–AlF_x. The resulting 'trapped' complex, $PspF_{1–275}$-(E)σ^{54}–ADP–AlF_x, reflects one of the intermediate states (RP_Is) *en route* to RP_O formation (36,37). G83*p*Bpa was the only variant tested capable of maintaining the σ^{54} interaction in an ADP–AlF_x dependent manner (Table 1 and Figure 2A). By irradiating the 'trapped' $PspF_{1–275}$ G83*p*Bpa-σ^{54}–ADP–AlF_x complex and analysed under denaturing conditions, multiple cross-linked $PspF_{1–275}$ G83*p*Bpa × σ^{54} species were observed (Figure 2B). Prolonged irradiation shifted the cross-linked species towards higher molecular forms (Figure 2B); it is likely that higher oligomeric states of G83*p*Bpa (G83*p*Bpa can self-associate and self-crosslink) associated with multiple σ^{54} regions (as demonstrated below). The cross-linked $PspF_{1–275}$G83*p*Bpa × σ^{54} species was only be observed in the 'trapped' complexes (in the presence of ADP–AlF_x but not in the presence of ATP, ADP or AMP, Supplementary Figure S1), consistent with the proposal that at the point of ATP hydrolysis, the PspF L1s assume a raised conformation to contact and thereby cross-link to σ^{54} (20). The G83*p*Bpa variant was unable to drive RP_O formation (Table 1) possibly due to sub-optimal L1 exposure (50% of WT activity) and low ATPase hydrolysis rate (10% of WT activity). However, this derivative of PspF did support partial functionalities required for forming RP_O, and was able to function for engagement of RP_C and support RP_O formation in a mixed oligomer with WT subunits (see below).

Based on the Cryo-EM structure, Bose *et al.* (11) proposed that up to three subunits in a PspF hexamer could potentially contact the RNAP holoenzyme via the PspF surface-exposed L1s. In addition, it has been shown that the AAA+ domain of the *S. meliloti* DctD (another well-studied hexameric bEBP) can cross-link to the core RNAP β subunit (38). To assess whether PspF L1s can directly contribute to core RNAP binding, we added core RNAP to the cross-linking reactions. As shown in Figure 2E, the cross-linking pattern between $PspF_{1–275}$ G83*p*Bpa and σ^{54} was not altered by the presence of core RNAP. The outcomes did not provide evidence to support a direct contact between the L1 'GAFTGA' motif and core RNAP, rather the protein contacts appear to be primarily with σ^{54}.

Taken together, formation of the ADP–AlF_x 'trapped' complexes with the G83*p*Bpa variant suggests it is a potentially useful reagent, as demonstrated in the following experiments, to elucidate the L1-interacting partners.

Table 1. Functional characterization of $PspF_{1–275}$ *p*Bpa variants

$PspF_{1–275}$	σ54 interaction (% of WT)		ATPase activity (% of WT)	RP_o formation (% of WT)
	ADP–AlF_X	AMP–AlF_X		
G83*p*Bpa	50	0	10	0
A84*p*Bpa	0	0	116	0
F85*p*Bpa	0	0	96	0
T86*p*Bpa	0	0	146	0
G87*p*Bpa	0	0	138	0
A88*p*Bpa	0	0	100	0
G83A	56	0	23	0
G83F	90	29	10	0.1

The $PspF_{1–275}$ *p*Bpa variant and the two G83 variants (substitution with Ala to remove the side chain and substitution with Phe to closely mimic the *p*Bpa cross-linker) were characterized in terms of σ^{54} interaction (in the presence of the ADP–AlF_x 'trapping' reagent), ATPase activity and RP_O formation (on a super-coiled *S. meliloti nifH* promoter).

Two novel PspF L1-binding patches within σ^{54}_{RI} were identified

The Cryo-EM structure of the ADP–AlF_x 'trapped' complex indicated that the PspF hexamer contacted two opposing sites in σ^{54} (12). If one contact site is the PspF L1 target site–σ^{54}_{RI} (36), the other contact site might be outside σ^{54}_{RI}. To explore this proposal, different radio-labelled σ^{54} fragments (σ^{54}_{RI}, $\sigma^{54}_{\Delta RI}$ and mixed $\sigma^{54}_{RI/\Delta RI}$) were used to form the cross-linked 'trapped' complexes with G83*p*Bpa. Both $PspF_{1–275}$ WT and G83*p*Bpa can form stable complexes with σ^{54}_{FL} and σ^{54}_{RI} but not with $\sigma^{54}_{\Delta RI}$ (Figure 2C). With the G83*p*Bpa variant, a major cross-linked species of ~39 kDa was observed (Figure 2D), corresponding to one $PspF_{1–275}$ G83*p*Bpa (33 kDa) cross-linked to one σ^{54}_{RI} (6 kDa). Other faint cross-linked G83*p*Bpa × σ^{54}_{RI} species with higher molecular weights were also observed (Figure 2D), suggesting that more than one L1 could contact σ^{54}_{RI}.

From the above observations, we set to explore the PspF L1-binding patches within σ^{54}_{RI} by generating six σ^{54}_{RI} peptide fragments (Figure 1B). As we screened the σ^{54}_{RI} fragments for their ability to stably bind $PspF_{1–275}$ hexamers under ADP–AlF_x 'trapping' conditions, a slower migrating complex was detected with either σ^{54}_{Frag} 18–25 or 33–39 but not the four other peptides tested (Figure 3A). Formation of the 'trapped' complexes in the presence of σ^{54}_{Frag} 18–25 and 33–39 was confirmed by gel filtration, providing evidence that the two peptides can associate with PspF in its ADP–AlF_x-bound state (Figure 3D). Given their relatively small size, these two σ^{54}_{RI} fragments are unlikely to assume 'complete' secondary structures in solution; thus their interaction with the $PspF_{1–275}$ hexamer may be largely dependent on their primary sequence and independent of their forming a well-ordered structure prior to binding to PspF. From here on, the σ^{54}_{RI} residues 18–25 will be named Patch 1 and residues 33–39 Patch 2.

To demonstrate that binding of the patch fragments to $PspF_{1–275}$ is dependent on established determinants for

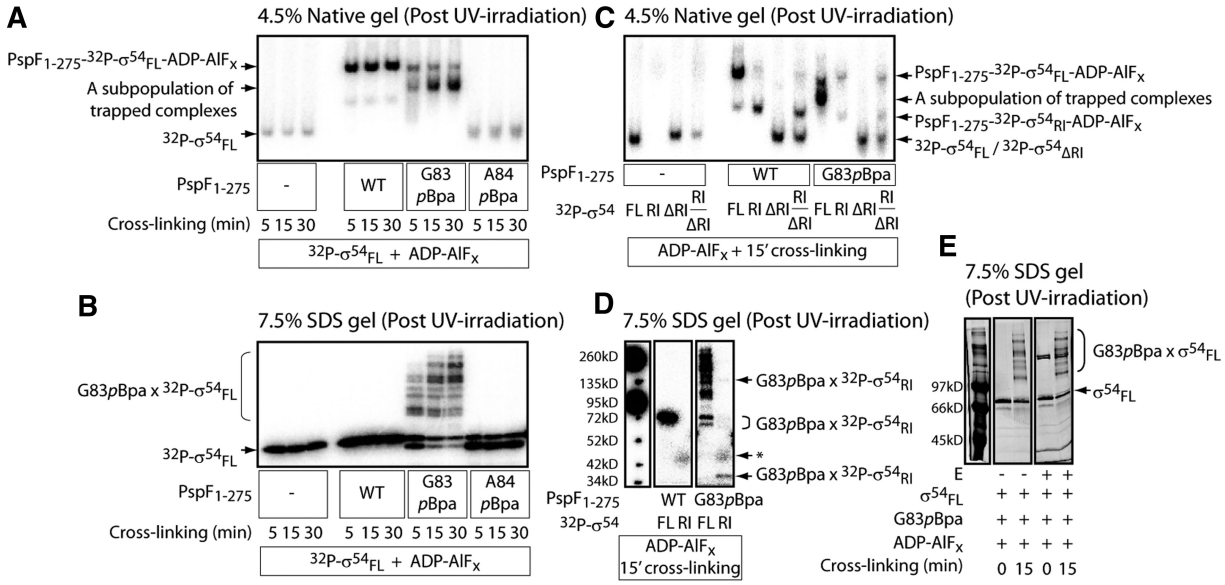

Figure 2. PspF L1 cross-links to multiple locations within σ^{54} Region I (σ^{54}_{RI}) but not to core RNAP. The HMK-tagged full-length σ^{54} was radio-labelled (^{32}P-σ^{54}_{FL}) and used in (A–D) but not in (E). The native gel (**A**) and the SDS PAGE gel (**B**) depict the binding and cross-linking, respectively, between $PspF_{1–275}$ *p*Bpa variants and ^{32}P-σ^{54}_{FL} in the presence of the ADP–AlF_x 'trapping' reagent. A protein band (G83*p*Bpa, Panel A) increases with UV-irradiation time, possibly corresponding to a subpopulation of the 'trapped' complex due to cross-linking. The A84*p*Bpa variant was used as a negative control for cross-linking as it cannot bind to σ^{54}. The cross-linked species between the G83*p*Bpa variant and ^{32}P-σ^{54}_{FL} were depicted as G83*p*Bpa × ^{32}P-σ^{54}_{FL}. From this point, all the following UV cross-linking reactions were performed with 15 min irradiation (the ADP–AlF_x-dependent complexes are most stable within 30 min—in this case, 15 min for complex formation and 15 min for UV irradiation). To assess the interaction and cross-linking between PspF L1 and different regions of σ^{54}, the ADP–AlF_x-dependent 'trapping' reactions were performed with radio-labelled Region I (^{32}P-σ^{54}_{RI}), Region I delete (^{32}P-$\sigma^{54}_{\Delta RI}$) and mixed RI/ΔRI (a 3:1 ratio). The presence of 'trapped' complexes was analysed on a native gel (**C**); the cross-linking event was analysed on an SDS PAGE gel (**D**). The band in the G83*p*Bpa × ^{32}P-σ^{54}_{RI} cross-linking reaction (D) indicated by an asterisk corresponds to an artefact also present in ^{32}P-σ^{54}_{RI} alone. To assess whether PspF L1 cross-links to core RNAP, core RNAP was added to the ADP–AlF_x-dependent 'trapping' reaction (**E**). The SDS PAGE gel was stained with Invitrogen Sypro Ruby stain and scanned by PhosphoImager. The presence of core RNAP does not seem to alter the cross-linking profile between G83*p*Bpa and σ^{54}. Similar results were obtained when ^{32}P-σ^{54}_{FL} was used.

the binding of PspF to σ^{54}, the L1 'GAFTGA' variants (F85Y and T86S) were used in the ADP–AlF_x reactions (Figure 3B). The F85Y variant appears to interact with the two patch fragments (the PspF band intensity increases dramatically in the presence of these fragments). However, their binding is possibly different from that of $PspF_{1–275}$ WT, as the complex band does not shift much nor into a compact band. The T86S variant can shift the complexes to the same level as WT with Patch 2 fragment but not with Patch 1 fragment. The above observations suggest that the L1 'GAFTGA' motif is indeed the direct target of the σ^{54}_{RI} patches. Based on the formation of stable 'trapped' complexes in a titration experiment (Figure 3C), the Patch 1 fragment exhibited a much higher affinity (at least 10-fold) towards $PspF_{1–275}$ than did the Patch 2 fragment.

Sequence alignment of the two σ^{54}_{RI} patches from different organisms (Figure 4A) revealed three highly conserved residues in Patch 1 and two highly conserved residues in Patch 2. Substitutions of these residues with Ala (QAQAAARL and AQQEAQQ, as underlined) produced variant fragments unable to detectably bind the hexameric $PspF_{1–275}$ (Figure 4B), suggesting that these residues are key to PspF L1 interactions with σ^{54}.

Taken together, we have identified three highly conserved residues in σ^{54}_{RI} Patch 1 contributing to the high affinity PspF L1 binding, and two highly conserved residues in Patch 2 to a lower affinity PspF L1 binding.

L1–σ^{54}_{RI} sequence-specific interactions play different roles along the activation pathway

After establishing that two σ^{54}_{RI} patches are sequence-specific for PspF L1 interactions, we next examined their impact in the context of σ^{54} and RNAP holoenzyme binding interactions and different steps in the transcription activation pathway. Thus, we generated three full-length σ^{54} variants: σ^{54}_{scm} Patch 1 (harbouring the scrambled Patch 1—'QAQAAARL'), σ^{54}_{scm} Patch 2 (harbouring the scrambled Patch 2—'AQQEAQQ') and σ^{54}_{scm} Patches 1 and 2 (harbouring both the scrambled patches). The initial assessment of the σ^{54}_{scm} patch variants revealed no large defect in forming the 'trapped' complexes (Figure 4C), suggesting additional sequences of the σ^{54}_{RI} along with promoter DNA contacts (see below) could compensate for the loss of L1 sequence-specific interactions in forming the 'trapped' transcription intermediate (RP_I).

Burrows *et al.* (39) devised an assay in which the ADP–AlF_x-dependent RP_I could carry out dinucleotide-primed short RNA (spRNA) transcription when the '–10 to –1' transcription bubble was pre-formed. Using this assay, we assessed the impact of the σ^{54}_{scm} patch variants on the

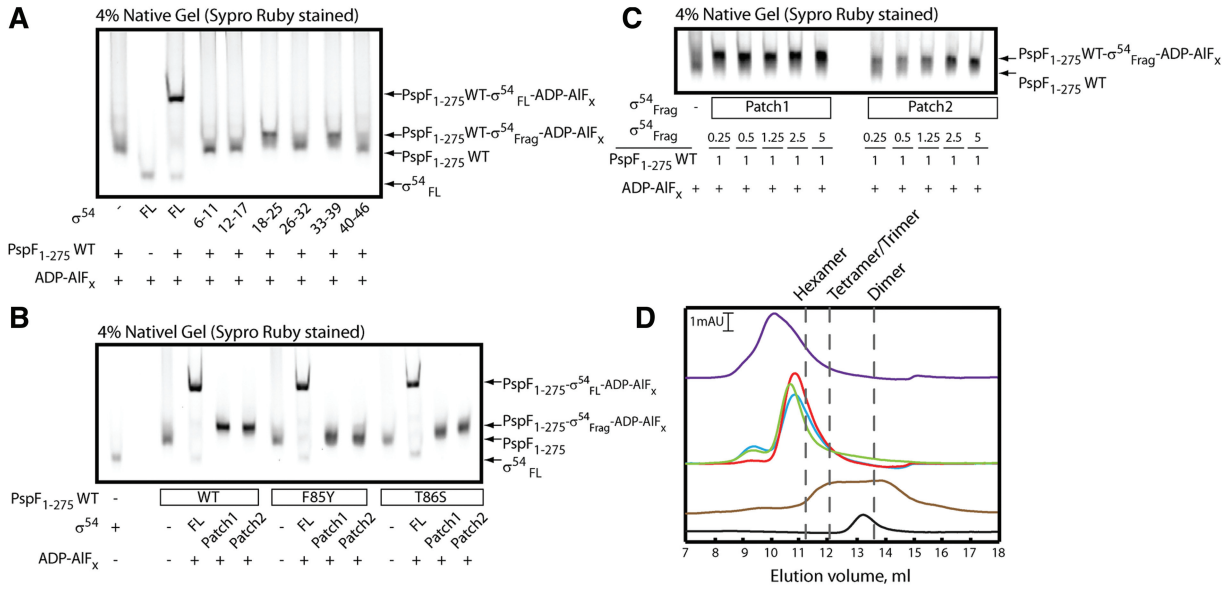

Figure 3. Two σ^{54}_{RI} peptide fragments bind to $PspF_{1-275}$ with different affinities. (**A**) $PspF_{1–275}$ WT hexamers bind to two σ^{54}_{RI} fragments (residues 18–25 and 33–39) in the presence of the ADP–AlF_x 'trapping' reagent. From here on, fragments 18–25 is depicted as Patch1 and fragments 33–39 as Patch2. (**B**) Mutations in the PspF 'GAFTGA' motif result in sensitivity to σ^{54}_{RI} patch binding. (**C**) σ^{54}_{RI} Patch1 fragments bind to $PspF_{1–275}$ WT hexamers with a markedly higher affinity than Patch2 fragments. A titration experiment was performed with a constant $PspF_{1–275}$ WT concentration while the concentration of each Patch fragment was gradually increased. (**D**) Gel filtration profiles of the 'trapped' complexes at 4°C using a Superdex 200 column. The black trace corresponds to 2 μM σ^{54}_{FL}. The brown trace corresponds to 20 μM $PspF_{1-275}$ WT in the presence of ADP. The green trace corresponds to 20 μM $PspF_{1–275}$ WT in the presence of ADP–AlF_x. The red trace corresponds to 20 μM $PspF_{1-275}$ WT binding to σ^{54}_{RI} Patch1 fragments in the presence of ADP–AlF_x. The blue trace corresponds to 20 μM $PspF_{1-275}$ WT binding to σ^{54}_{RI} Patch2 fragments in the presence of ADP–AlF_x. The purple trace corresponds to 20 μM $PspF_{1–275}$ WT binding to σ^{54}_{FL} in the presence of ADP–AlF_x. The gel elution volumes corresponding to apparent hexamers, tetramers/trimers and dimers are marked by dotted lines.

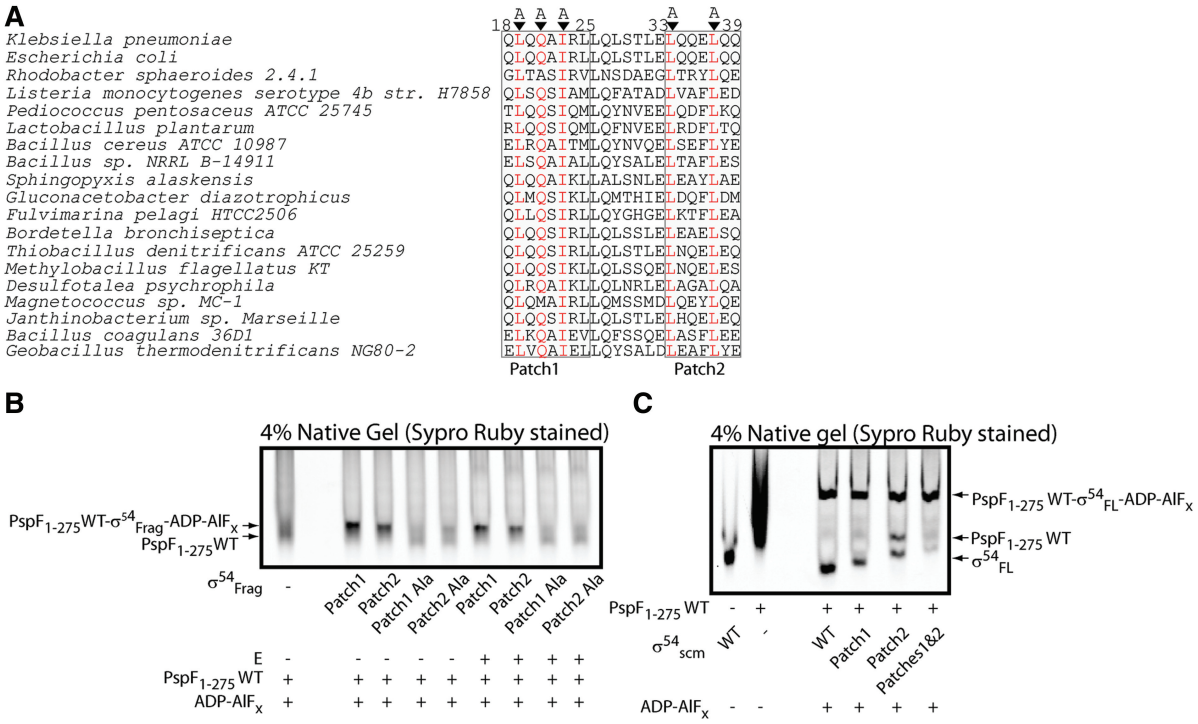

Figure 4. Scrambled σ^{54}_{RI} Patches fail to bind $PspF_{1–275}$ WT in the context of fragments but can bind $PspF_{1–275}$ WT in the context of full-length proteins. (**A**) Sequence alignment of σ^{54}_{RI} Patches 1 and 2 from different bacteria using NCBI BLAST. Highly conserved residues are highlighted in red and subsequently replaced with Ala ('Scrambled' or 'scm'). (**B**) The scrambled patches failed to bind $PspF_{1–275}$ WT hexamers in the context of fragments in the presence of ADP–AlF_x 'trapping' reagents. (**C**) The scrambled patches were able to bind $PspF_{1–275}$ WT hexamers in the context of full-length σ^{54} in the presence of ADP–AlF_x 'trapping' reagents. σ^{54}_{scm} Patches 1 and 2 corresponds to the full-length σ^{54} harbouring both the scrambled patches.

amount of transcriptionally 'active' RP_Is generated. As shown in Figure 5B, despite starting with a similar amount of ADP–AlF_x-dependent RP_Is (a slightly more pronounced reduction in RP_I was observed with $\sigma^{54}{}_{scm}$ Patch 2 and Patches 1 and 2, Figure 5A), all three $\sigma^{54}{}_{scm}$ patch variants were able to produce significantly more spRNAs (4–8-fold) than was the σ^{54} WT, indicating the RP_Is were more active than with σ^{54} WT. In contrast, the ATPase-dependent RP_O formation assays replacing ADP–AlF_x with dATP on the same DNA probe revealed significant defects in all three $\sigma^{54}{}_{scm}$ patch variants (Figure 5C). Interestingly, all three $\sigma^{54}{}_{scm}$ patch variants can generate an RP_O in the absence of the $PspF_{1–275}$ activator and hydrolysable nucleotides on the pre-melted DNA probe (Figure 5D), so revealing an activator-bypass phenotype (40). Considering the RP_O generated from the activator-bypass activity may have contributed to the amount of RP_O as observed in Figure 5C, the defect of the $\sigma^{54}{}_{scm}$ patch variants in the ATPase-driven isomerization may be more pronounced.

Taken together, the above data demonstrate that the PspF L1–$\sigma^{54}{}_{RI}$ sequence-specific interactions may play an inhibitory role in the activity of RP_I, possibly to keep the complex in check before moving to RP_O. Once the inhibitory PspF L1–$\sigma^{54}{}_{RI}$ interactions were disrupted (by scrambling the Region I patches), the spontaneous transition from RP_I to RP_O is clearly increased as seen in the activator-bypass assays. In contrast, the PspF L1–$\sigma^{54}{}_{RI}$ sequence-specific interactions are needed for ATPase-driven RP_O formation, suggesting important roles for the patches in making RP_I from RP_C and in limiting the activity of RP_I in the activator-dependent pathway. Transient interactions between RP_C and PspF in the ATPase-driven reaction may therefore be more dependent upon the integrity of the $\sigma^{54}{}_{RI}$ than is the stably engaged RP_I created with ADP–AlF_x.

'Doped' WT/G83*p*Bpa heterohexamers can directly cross-link to promoter DNA

After demonstrating the G83*p*Bpa variant can cross-link to σ^{54} in the ADP–AlF_x 'trapped' complex, the DNA probe harbouring the early-melted *nifH* promoter (–12–11/WT, mimicking the –12 fork junction DNA in the RP_C, Figure 6) was added to the reaction mixture. If this fork junction DNA conformation was successfully accommodated in the 'trapped' complex, the spatial proximity between G83*p*Bpa and the corresponding promoter region could in principle be determined.

As shown in Figure 6, the radio-labelled –12–11/WT DNA probe was not efficiently covalently bound into the ADP–AlF_x 'trapped' complexes with G83*p*Bpa homohexamers (ratio of WT/G83*p*Bpa was 0/6), resulting in a cross-linked species with abundance only slightly above the background. Since the $PspF_{1–275}$ WT homohexamers

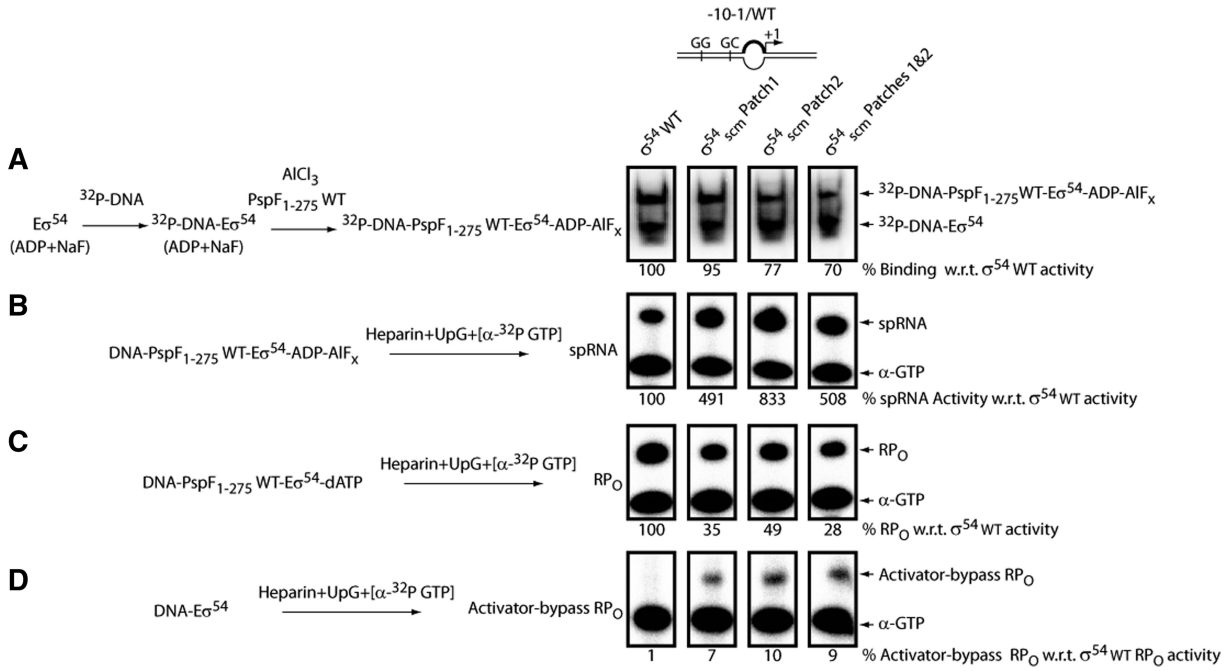

Figure 5. Functional importance of the PspF L1–$\sigma^{54}{}_{RI}$ patch interactions along the activation pathway. A simplified reaction scheme is depicted for each assay. Full-length σ^{54} harbouring the scrambled patches and radio-labelled –10–1/WT DNA probes were used in the following reactions. The linear –10–1/WT DNA probe harbours a mismatch from –10 to –1 on the non-template strand to mimic the DNA conformation in RP_O. (**A**) The ability of scrambled σ^{54} patch variants ($\sigma^{54}{}_{scm}$ patch) to form the ADP–AlF_x-dependent 'trapped' complexes, each expressed as a percentage of that of σ^{54} WT. (**B**) Each $\sigma^{54}{}_{scm}$ patch variant was allowed to form the ADP–AlF_x-dependent RP_I complex (DNA–$PspF_{1–275}$ WT–Eσ^{54}–ADP–AlF_x). The resulting RP_I complexes were tested for their ability to support transcription in the presence of the spRNA mixture (heparin, dinucleotide primer UpG and [α-^{32}P GTP]). The extent of transcription activity (correlates with the amount of spRNA synthesis) was expressed as a percentage of that of σ^{54} WT. (**C**) The amount of RP_O generated by $\sigma^{54}{}_{scm}$ patch variants in the presence of hydrolysable nucleotide dATP, each expressed as a percentage of that of σ^{54} WT. (**D**) The amount of activator-bypass RP_O generated by $\sigma^{54}{}_{scm}$ patch variants in the absence of $PspF_{1–275}$ WT, each expressed as a percentage of that of σ^{54} WT in the activator-dependent assay (C).

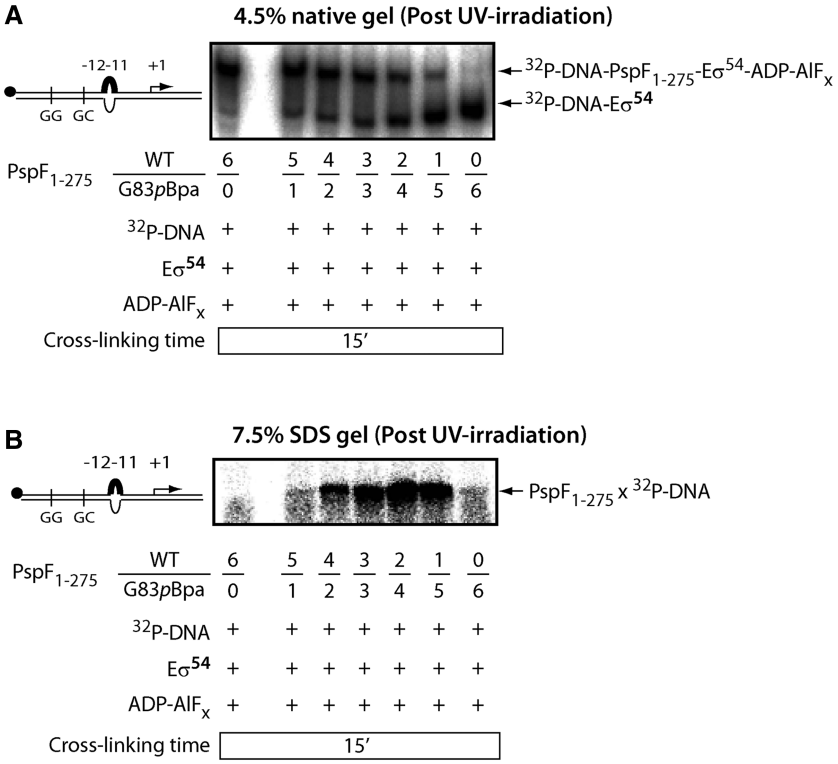

Figure 6. The 'doped' $PspF_{1–275}$ WT/G83*p*Bpa heterohexamers cross-link to promoter DNA. The ADP–AlF_x-dependent 'trapping' reaction was performed in the presence of radio-labelled –12–11/WT DNA probe. The –12–11/WT DNA probe harbours a mismatch from –12 to –11 on the non-template strand to mimic the DNA conformation in RP_C. $PspF_{1–275}$ WT and G83*p*Bpa subunits were mixed with different ratios (6/0, 5/1, 4/2, 3/3, 2/4, 1/5 and 0/6) in the 'trapping' reaction. Core RNAP was also added to the reaction mixture to ensure all the transcriptional components were in place. Samples were loaded on a native gel (**A**) and on an SDS PAGE gel (**B**). A single cross-linked $PspF_{1–275} \times$ DNA species was observed.

(ratio of WT/G83*p*Bpa was 6/0) were able to efficiently form the 'trapped' complexes in the presence of DNA (Figure 6A), we reasoned that by 'doping' G83*p*Bpa with different ratios of WT subunits, accommodation of the –12–11/WT DNA may be achieved. Not only did the 'doping' experiment successfully restore the 'trapped' complex formation, but also generated a single $PspF_{1–275} \times$ DNA cross-linked species (Figure 6B). The above observation provides clear evidence that subunit mixing indeed occurred (see also Figure 7) and that the reconstituted WT/G83*p*Bpa heterohexamers likely contacted the promoter DNA via position 83 or adjacent residues in the L1 'GAFTGA' motif. Taken together, the above data strongly support a chemical bonding interaction between L1 'GAFTGA' and the promoter DNA in the 'doped' WT/G83*p*Bpa heterohexamers, as the *p*Bpa cross-linking chemistry requires a distance of 3.1 Å (recall the H-bonding distance is 2 Å). Spatial organizations reported for RNAP-σ^{54} and the L1-DNA cross-linking event together suggest that L1s from different subunits of the hexamer must be involved in DNA and $\sigma^{54}{}_{RI}$ contacts.

Next we 'doped' the rest of the *p*Bpa variants (recall all failed to form the 'trapped' complexes, Table 1 and Supplementary Figure S2) with the $PspF_{1–275}$ WT subunits for DNA cross-linking complementation. Strikingly, the 'doped' WT/T86*p*Bpa heterohexamers showed a comparatively strong DNA cross-linking signal with the –12–11/WT DNA probe (Supplementary Figure S2B). This outcome implies a role for the conserved residue T86 or residues adjacent to it in DNA contact in RP_C. Previously, T86 has only been characterized as being a σ^{54}-contacting residue of L1.

To assess whether the 'doped' WT/G83*p*Bpa heterohexamers were biologically relevant to the *bona fide* WT homohexamers, we examined the activity in the context of their self-association and RP_O formation by mixing an equimolar amount of $PspF_{1–275}$ WT and G83*p*Bpa subunits (Figure 7). By gel filtration of apo forms, we establish that $PspF_{1–275}$ WT exists as a mixture of apparent tetramers/dimers (12 ml/13.3 ml elution volumes, Figure 7A purple trace) at a 30 μM injection concentration. The G83*p*Bpa variant exists predominantly as apparent octamers/tetramers (10.7 ml/12 ml elution volumes, Figure 7A blue trace) at a 30 μM injection concentration. The 'doped' WT/G83*p*Bpa mixture generated an apparent hexameric peak (11.2 ml elution volume, Figure 7 red trace), eluting at the same volume as the WT homohexamers [11.18 ml elution volume, (28,41)]. This apparent hexameric peak was absent in the theoretical sum of each individual subunit profile (Figure 7 green trace). We thus conclude that the 'doped' WT/G83*p*Bpa heterohexamer is very similar in overall geometry to the WT homohexamer and subunit mixing indeed occurred. We also tested the ability of the WT/G83*p*Bpa heterohexamer to form RP_O under three different total concentrations (Figure 7B). The

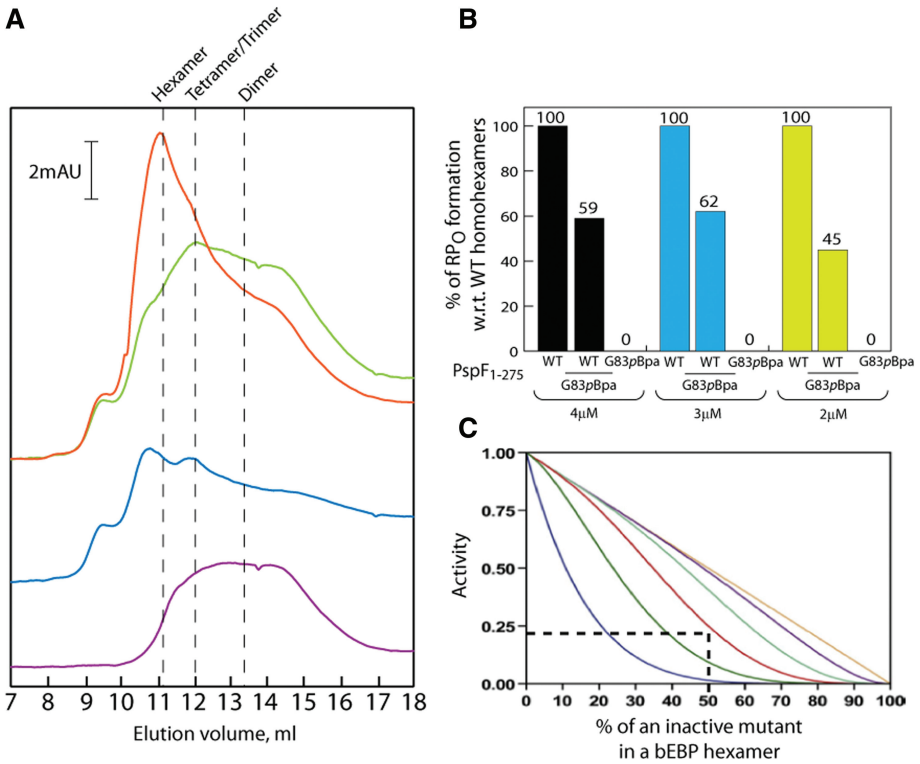

Figure 7. The 'doped' $PspF_{1–275}$ WT/G83*p*Bpa heterohexamers are transcriptionally active. (**A**) Gel filtration profiles of $PspF_{1–275}$ WT, G83*p*Bpa and WT/G83*p*Bpa mixture at 4°C in the absence of nucleotides. The purple trace corresponds to 30 μM $PspF_{1–275}$ WT. The blue trace corresponds to 30 μM $PspF_{1–275}$ G83*p*Bpa. The green trace corresponds to the theoretical profile of mixing 30 μM WT subunits with 30 μM G83*p*Bpa subunits. The red trace corresponds to the experimental profile of mixing 30 μM WT subunits with 30 μM G83*p*Bpa subunits. A prominent hexameric peak was observed in the experimental profile of the WT/G83*p*Bpa mixture, corresponding to the presence of true heterohexamers. (**B**) The amount of RP_O formation in the presence of WT homohexamers, WT/G83*p*Bpa heterohexamers (an equimolar ratio of WT and G83*p*Bpa subunits) and G83*p*Bpa heterohexamers under three different total $PspF_{1-275}$ concentrations. The –10–1/WT DNA probe was used in this assay. The average amount of RP_O generated by WT/G83*p*Bpa heterohexamer is 55% of that of WT homohexamers. (**C**) Theoretical activities of the 'doped' heterohexamers [taken from (Werbeck *et al*. (42)]. One inactive subunit 'doped' with five WT subunits (blue), two inactive subunits 'doped' with four WT subunits (green), three inactive subunits 'doped' with three WT subunits (red), four inactive subunits 'doped' with two WT subunits (mint), five inactive subunits 'doped' with one WT subunit (purple) and six inactive subunits (brown) are necessary to abolish enzyme activity. The WT/G83*p*Bpa heterohexamers (WT/G83*p*Bpa subunit ratio 3/3) should theoretically generate 25% the RP_O of that of WT homohexamers (dotted lines).

'doped' WT/G83*p*Bpa heterohexamers generated on average 55% the RP_O of the WT homohexamers (Figure 7B). Based on the statistical model for mixing experiments (42), an equimolar amount of WT and an inactive variant (in this case G83*p*Bpa) should theoretically generate ~25% the RP_O of the WT homohexamers (Figure 7C dotted lines). The above considerations imply that the reconstituted WT/G83*p*Bpa heterohexamers are indeed active in RP_O formation, and that the G83*p*Bpa subunits contribute to this activity.

L1 cross-links to the non-template '–29' region in RP_C/RP_I

To determine the precise DNA region cross-linked by L1, we employed a Proteinase K-ExoIII footprinting method (35). The rationale of this approach is to remove all the protein components by Proteinase K after UV irradiation; a stably cross-linked *p*Bpa peptide will remain attached to the DNA cross-linking site and physically block the read-through of ExoIII (a 3′–5′ exonuclease).

The G83*p*Bpa homohexamer is able to weakly bind and cross-link to DNA (Figure 6B). However, owing to its inability to activate transcription (Table 1 and Figure 7B), we chose the transcriptionally active 'doped' WT/G83*p*Bpa heterohexamer for footprinting. The L1 cross-linking site was mapped initially on the –12–11/WT DNA probe as this gave a very clear cross-linking signal (Supplementary Figure S2B). An ExoIII-resistant site was observed from approximately –30 to –27 on the non-template strand but not observed on the template strand (compare Figure 8A with 8B). The above data clearly demonstrate that in RP_C/RP_I, a L1 contacts the non-template strand of the promoter region between –30 and –27 (abbreviated as the '–29 region'), immediately upstream of the consensus '–24' GG element (located at –26–25 in the *nifH* promoter). Removal of the entire upstream sequence of the consensus GG yielded a near 60% reduction in L1-DNA cross-linking, further confirming the '–29' region is the major target site of L1 'GAFTGA' interaction (Figure 8C).

The '–29' region is important for activator-dependent RP_O formation

We next addressed whether the '–29' region was important for isomerization from RP_C to RP_O and forming 'trapped'

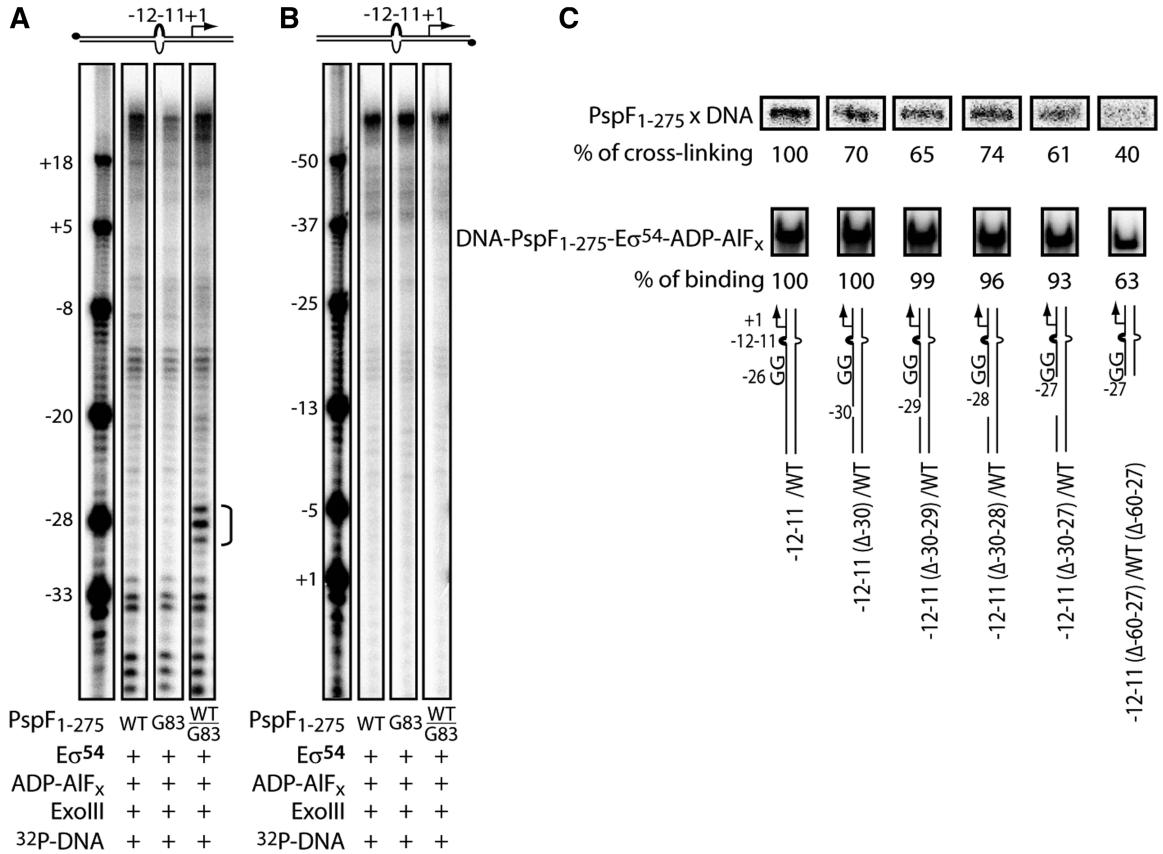

Figure 8. L1 cross-links to the non-template '–29' region of the –12–11/WT DNA probe. (**A** and **B**) $PspF_{1–275}$ WT homohexamers, WT/G83*p*Bpa heterohexamers and G83*p*Bpa homohexamers were allowed to form the 'trapped' complexes with $E\sigma^{54}$ in the presence of ADP–AlF_x. The 'trapped' complexes were then subject to Proteinase K-ExoIII footprinting assays. Proteinase K removes all the protein components. The *p*Bpa cross-linker in theory should remain attached to the promoter site and subsequently blocks the read-through of ExoIII nuclease. The cross-linking site is depicted by a half bracket from –27 to –30 (collectively called the '–29' site). Both the template strand, (**B**) and the non-template strand (**A**) were radio-labelled on the –12–11/WT DNA probe (^{32}P as black dots). (**C**) Nicking and truncation of the '–29' region reduced the L1 cross-linking efficiency. DNA strand nicking/truncation around the non-template '–29' region was made on the –12–11/WT DNA probe to assess the impact on UV cross-linking and DNA binding. Sequential nicking of the non-template '–29' region reduced the UV cross-linking efficiency by ~30–40% (top panel). The reduction in UV cross-linking was not due to unfavourable DNA binding (DNA binding was largely unaffected, middle panel), except for the –12–11(Δ–60–27)/WT(Δ–60–27) DNA probe where the entire upstream region of the consensus GG was removed. Nicking/truncation may cause the linear DNA probes to adopt slightly different confirmations to the unmodified probe (e.g. secondary sites of cross-linking), which could explain the absence of complete abolishment in UV cross-linking.

complexes to make RP_I. The RP_O formation assays were performed and the amount of RP_O formed from the long DNA probe was compared with that from the short DNA. Both DNA probes harboured a mismatch from –10 to –1 on the non-template strand since this DNA conformation gave the strongest activation signal amongst the three linear probes used in this assay. As shown in Figure 9, although the RP_C formation with the probe lacking DNA upstream of '–29' was reduced by 2-fold (Figure 9 and Supplementary Figure S3), 'trapped' complexes (RP_I) with activators were reduced by 18-fold (Figure 9 and Supplementary Figure S4A) and activator-dependent RP_O formation was reduced by 35-fold (Figure 9 and Supplementary Figure S4B). Since truncation of the '–29' region did not reduce the stability of RP_O by more than 2-fold (Supplementary Figure S6), the large RP_O formation defect cannot be attributed to an unstable RP_O generated from the shortened DNA probe. We conclude that the cross-linking of '–29' region DNA to L1 of the activator is important for forming RP_I and RP_O. Parallel experiments with fully duplexed probes confirmed the importance of the '–29' sequence for trapping and activator-dependent formation of RP_O (Supplementary Figure S5).

After showing that the activator-dependent RP_O formation (with $PspF_{1–275}$ WT-σ^{54} WT) involves the interaction between L1 'GAFTGA' motif and the '–29' region, we addressed whether the L1–DNA interaction was important for activator-independent RP_O formation. Residue R336 of σ^{54} is located in the DNA cross-linking sequence of Region III (Figure 1B). The σ^{54} R336A induces an activator-bypass phenotype (40) where the RP_O formation is independent of PspF hexamers and activating nucleotides on a super-coiled or on a linear –10 to –1 pre-opened *nifH* promoter.

In contrast to the activator-dependent RP_O formation (with $PspF_{1–275}$ WT-σ^{54} WT) which was greatly reduced when the '–29' region was truncated (a 35-fold reduction

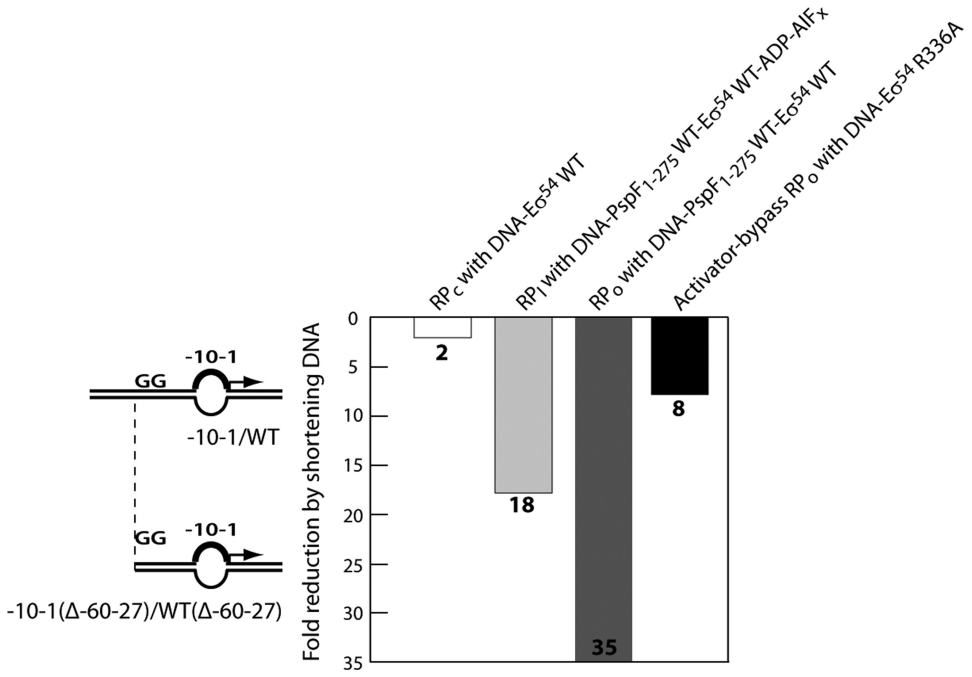

Figure 9. Truncation of the '–29' region affects both activator-dependent and -independent transcription activation. The –10–1/WT and –10–1(Δ–60–27)/WT(Δ–60–27) DNA probes were used in the assays to assess the impact of truncation of the '–29' region on activator-dependent $RP_C/RP_I/RP_O$ formation and on activator-independent RP_O formation. The activator-dependent RP_C formation was assessed by the stability of DNA–Eσ^{54} WT complexes (Supplementary Figure S3). The activator-dependent RP_I formation was assessed by the stability of DNA–Eσ^{54} WT–$PspF_{1-275}$ WT–ADP–AlF_x 'trapped' complexes (Supplementary Figure S4A). The activator-dependent RP_O formation was assess by the amount of RP_O generated by DNA-Eσ^{54} WT- $PspF_{1-275}$ WT (Supplementary Figure S4B and Supplementary Figure S7 left panel). The activator-independent RP_O formation was assessed by the amount of RP_O generated by DNA-Eσ^{54} R336A (Supplementary Figure S7 right panel).

on average, Figure 9 and Supplementary Figure S7), the activator-independent RP_O formation with σ^{54} R336A remained relatively constant on both DNA probes and was only reduced by 8-fold when the '–29' region was removed (Figure 9 and Supplementary Figure S7). The differences in RP_O formation were not due to different affinities of Eσ^{54} WT and Eσ^{54} R336A holoenzymes towards DNA (Supplementary Figure S3). The above observations demonstrate that the activator-independent pathway for RP_O formation is 4-fold less sensitive to loss of the '–29' region than is the activator-dependent pathway. Clearly, the '–29' region is important for σ^{54} WT-containing RNAP to form RP_O in an activator-dependent manner.

DISCUSSION

By using a fragmentation approach, we were able to identify precisely two amino acid patches within $\sigma^{54}{}_{RI}$ responsible for PspF L1 contact (residues 18–25 and 33–39). Both $\sigma^{54}{}_{RI}$ patches are located within the Leu heptad/hexad repeats (residues 19–44) where the activator-bypass mutations can be found (43,44). It is not known whether these two $\sigma^{54}{}_{RI}$ patches are contacted by two L1s simultaneously or in sequence when RP_C passes to RP_I and then to RP_O. Based on the shared phenotypes of the three full-length $\sigma^{54}{}_{scm}$ patch variants (Figure 5), it is possible the two $\sigma^{54}{}_{RI}$ patches are contacted by two L1s from adjacent PspF subunits in a synchronized manner. This interaction with $\sigma^{54}{}_{RI}$ may be responsible for initially holding the holoenzyme and the PspF hexamer together through interactions at the –12 part of the promoter DNA in RP_C and then organizing the RP_I to accept the melted DNA (Figure 10C). The latter reorganization may involve the LI to upstream DNA contact as described below.

Two structural features of the RP_C are thought to impede spontaneous RP_O formation: (i) the $\sigma^{54}{}_{RI}$ which interacts with core RNAP and blocks DNA entry, functionally reminiscent to $\sigma^{70}{}_{1.1}$ in RP_C (45–48). (ii) The –12 DNA melting site which is modelled in an upstream position with respect to its place in RP_O and is, therefore, misaligned with the active channel of the holoenzyme (11). We speculate that the contact between L1 and the upstream promoter DNA facilitates RP_I and RP_O formation. Although both L1 and $\sigma^{54}{}_{RIII}$ contact the non-template '–29' region (Figure 8 and Figure 10A), they do not appear to contact one another (Figure 2C). L1 and $\sigma^{54}{}_{RIII}$ may access the '–29' region from different DNA grooves to hold the promoter DNA firmly in place (Figure 10A). Doucleff *et al.* (14) proposed that $\sigma^{54}{}_{RIII}$ may interact with the β G flap. Thus, nucleotide-dependent conformational changes directed by L1 may facilitate the re-alignment of the holoenzyme with the –12 DNA melting site via the proposed $\sigma^{54}{}_{RIII}$–β G flap interaction. In principle, reorganization of $\sigma^{54}{}_{RIII}$ triggered by L1 movement could be transmitted to $\sigma^{54}{}_{RI}$ and result in disruption of the previously established inhibitory interactions at the –12 fork junction DNA maintained by $\sigma^{54}{}_{RI}$. An LI-directed torsion generated on $\sigma^{54}{}_{RI}$ may facilitate the propagation of DNA melting from –12

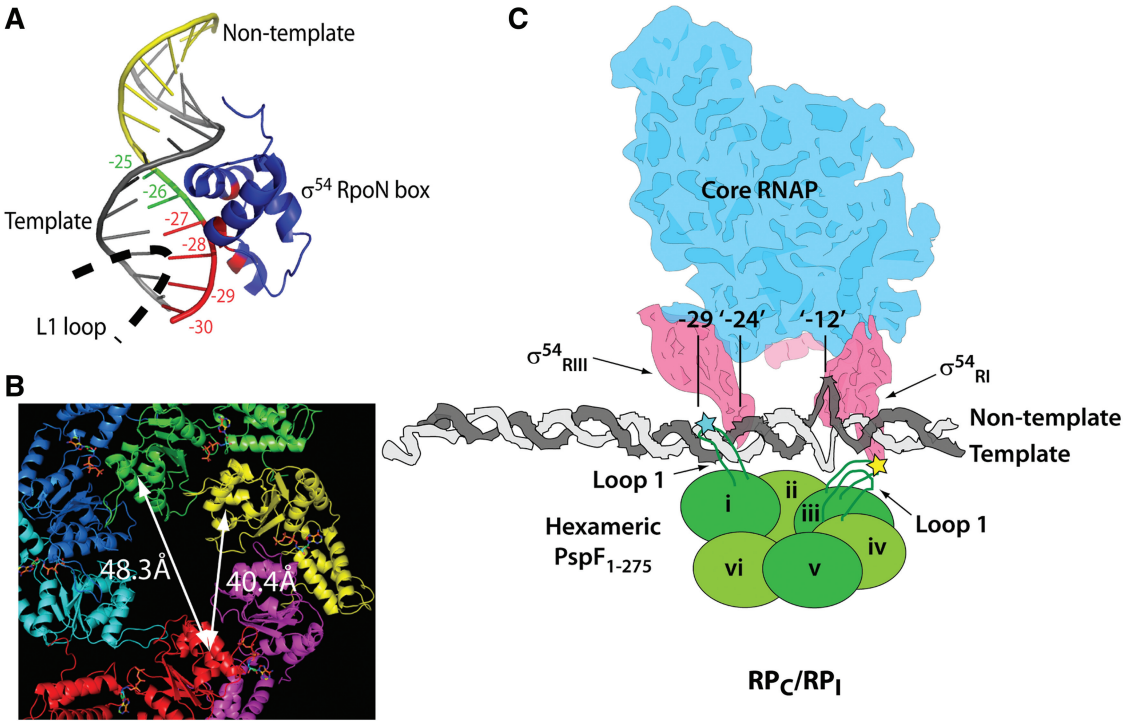

Figure 10. The proposed organizsation of L1s in engaged RP_C/RP_I. (**A**) L1 and σ^{54}_{RIII} might contact the promoter '–29' region from different grooves. The σ^{54}_{RIII} RpoN box (blue) binds to the non-template strand of the *nifH* promoter (PDB 2O8K). The consensus GG element at –25 and –26 is highlighted in green. The '–29' region is highlighted in red. (**B**) The $PspF_{1–275}$ WT hexameric structure with an open spiral (based on energy minimization of monomeric ATP-bound crystal structures, courtesy of M. Rappas) shows that the distances between the centres of two L1s across the hexameric plane are 40.4 Å [measured from subunit (i) to subunit (iii)] and 48.3 Å [measured from subunit (i) to subunit (iv)], consistent with the distance between the boundaries of the '–29' element and the –12 fork junction (~40–47 Å). (**C**) The proposed model based on the cross-linking data. The L1–DNA and L1–σ^{54}_{RI} cross-linking sites are depicted by the cyan and yellow stars, respectively. Alternate subunits of the $PspF_{1–275}$ open spiral are highlighted by different depths of green.

to the start site. Thus, disruption of the L1–DNA interaction around the '–29' region could partially contribute to the activation defect as observed when this upstream interaction cannot be established (Figure 9). As PspF progresses through cycles of ATP hydrolysis, L1s contacting both the '–29' region and σ^{54}_{RI} are likely to change from an extended state to a folded down state (19,20,49). These changes may move the DNA downstream to facilitate the re-alignment of the –12 fork junction with the active site of the holoenyzme (11). This is analogous to the DNA threading observed in AAA+ helicases by the 'staircasing' ssDNA-binding hairpins (50).

An open spiral hexameric configuration has been observed in many AAA+ proteins (51–56). Joly *et al.* (29) proposed that the $PspF_{1-275}$ hexamer also assumed an open spiral configuration employing at least two functional L1s for σ^{54} binding. Thus we incorporated the $PspF_{1–275}$ spiral into a proposed transcription complex organization within RP_C/RP_I (Figure 10C). Structural analyses indicate that the distances between the centres of the base of L1 across the $PspF_{1–275}$ spiral are 40.4 Å [measured from subunit (i) to subunit (iii)] and 48.3 Å [measured from subunit (i) to subunit (iv), Figure 10B]. This fits well with a 40–47 Å span between the boundary of the '–29' region (position –27, contacted by a L1) and the boundary of the '–12' element [position –14, contacted by σ^{54}_{RI} (57)].

We believe the underlying mechanism proposed can be extended from PspF to other bEBPs. In this context, mutation of the second Gly in the 'GAFTGA' motif of *Salmonella typhimurium* NtrC results in a 'super' DNA binding activity (58). The authors suggested that the 'GAFTGA' motif may be close to DNA in such inactive bEBP dimmers, and the binding activity may be non-specific due to the additional charge introduced (Lys in place of Gly). In this study, we provide evidence to show that the L1 'GAFTGA' motif is presented for a specific and direct promoter DNA engagement within active bEBP hexamers. Another bEBP that warrants discussion is the *Aquifex aeolicus* NtrC1 protein. NtrC1 forms closed oligomers in solution [90% heptamers and 10% hexamers (59)]. The negative-stain EM data suggest that although heptameric NtrC1 can engage σ^{54}, a significant portion of the density for σ^{54} is missing in the co-complex (17). Quite how far the NtrC1 heptamer–σ^{54} complex might functionally deviate from the transcriptionally active complexes forming with more usual hexameric assemblies of bEBPs, such as PspF, NtrC, ZraR, DpmR, NorR and HrpR/S (18,41,60–63), is unknown. A heptameric arrangement of NtrC1 as compared to, for example, the hexameric ZraR (62) is anticipated to be distinct in terms of the details of interfacial subunit–subunit contacts, some of which are known to control the nucleotide-dependent remodelling output of

PspF (64) and to precisely define sites where the ATPase can become uncoupled from remodelling by simple mutation. Although the triple L1 contacts observed in PspF is dictated by the stoichiometry and arrangement of the ring, it can still apply to NtrC1 if the exchange between heptamers and hexamers occurs frequently to allow a faithful and productive engagement of σ^{54} and promoter DNA.

To conclude, our work provides clear evidence that discrete L1s make interactions with three distinct and well separated elements within RP_C/RP_I, these are the two $\sigma^{54}{}_{RI}$ patches and the '–29' promoter region. The triple contacts by a single feature of a bEBP contrast directly with many AAA+ proteins (e.g. unfoldases and helicases) that contact only either protein or DNA and the classic bacterial activators (e.g. CRP-cAMP receptor protein) that contact protein and DNA via two distinct and spatially well separated domains.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Tables 1 and 2 and Supplementary Figures 1–7.

ACKNOWLEDGEMENTS

We gratefully acknowledge Dr J.W. Chin and Dr D.P. Nguyen for providing the pDULE-*p*Bpa plasmid. We thank E. Lawton and Dr T. Simpson for their stimulating comments on the manuscript. We thank all the past and present MB group members for their friendly support. N.Z., N.J. and M.B. conceived and designed the experiments included in this manuscript. N.Z. performed the experiments.

FUNDING

Biotechnology and Biological Sciences Research Council project (BBSRC) [BB/J002828/1 and BB/G001278/1 to M.B.]. Funding for open access charge: BBSRC [BB/J002828/1 and BB/G001278/1].

Conflict of interest statement. None declared.

REFERENCES

1. Browning,D.F. and Busby,S.J. (2004) The regulation of bacterial transcription initiation. *Nat. Rev. Microbiol.*, **2**, 57–65.
2. Morris,L., Cannon,W., Claverie-Martin,F., Austin,S. and Buck,M. (1994) DNA distortion and nucleation of local DNA unwinding within sigma-54 (sigma N) holoenzyme closed promoter complexes. *J. Biol. Chem.*, **269**, 11563–11571.
3. Burrows,P.C., Wigneshweraraj,S.R. and Buck,M. (2008) Protein-DNA interactions that govern AAA+ activator-dependent bacterial transcription initiation. *J. Mol. Biol.*, **375**, 43–58.
4. Friedman,L.J. and Gelles,J. (2012) Mechanism of transcription initiation at an activator-dependent promoter defined by single-molecule observation. *Cell*, **148**, 679–689.
5. Lin,Y.C., Choi,W.S. and Gralla,J.D. (2005) TFIIH XPB mutants suggest a unified bacterial-like mechanism for promoter opening but not escape. *Nat. Struct. Mol. Biol.*, **12**, 603–607.
6. Studholme,D.J. and Buck,M. (2000) The biology of enhancer-dependent transcriptional regulation in bacteria: insights from genome sequences. *FEMS Microbiol. Lett.*, **186**, 1–9.
7. Wolfe,A.J., Millikan,D.S., Campbell,J.M. and Visick,K.L. (2004) Vibrio fischeri sigma54 controls motility, biofilm formation, luminescence, and colonization. *Appl. Environ. Microbiol.*, **70**, 2520–2524.
8. Correa,N.E., Lauriano,C.M., McGee,R. and Klose,K.E. (2000) Phosphorylation of the flagellar regulatory protein FlrC is necessary for Vibrio cholerae motility and enhanced colonization. *Mol. Microbiol.*, **35**, 743–755.
9. Fisher,M.A., Grimm,D., Henion,A.K., Elias,A.F., Stewart,P.E., Rosa,P.A. and Gherardini,F.C. (2005) Borrelia burgdorferi sigma54 is required for mammalian infection and vector transmission but not for tick colonization. *Proc. Natl Acad. Sci. USA*, **102**, 5162–5167.
10. Villain-Guillot,P., Bastide,L., Gualtieri,M. and Leonetti,J.P. (2007) Progress in targeting bacterial transcription. *Drug Discov. Today*, **12**, 200–208.
11. Bose,D., Pape,T., Burrows,P.C., Rappas,M., Wigneshweraraj,S.R., Buck,M. and Zhang,X. (2008) Organization of an activator-bound RNA polymerase holoenzyme. *Mol. Cell*, **32**, 337–346.
12. Rappas,M., Schumacher,J., Beuron,F., Niwa,H., Bordes,P., Wigneshweraraj,S., Keetch,C.A., Robinson,C.V., Buck,M. and Zhang,X. (2005) Structural insights into the activity of enhancer-binding proteins. *Science*, **307**, 1972–1975.
13. Doucleff,M., Malak,L.T., Pelton,J.G. and Wemmer,D.E. (2005) The C-terminal RpoN domain of sigma54 forms an unpredicted helix-turn-helix motif similar to domains of sigma70. *J. Biol. Chem.*, **280**, 41530–41536.
14. Doucleff,M., Pelton,J.G., Lee,P.S., Nixon,B.T. and Wemmer,D.E. (2007) Structural basis of DNA recognition by the alternative sigma-factor, sigma54. *J. Mol. Biol.*, **369**, 1070–1078.
15. Hong,E., Doucleff,M. and Wemmer,D.E. (2009) Structure of the RNA polymerase core-binding domain of sigma(54) reveals a likely conformational fracture point. *J. Mol. Biol.*, **390**, 70–82.
16. Svergun,D.I., Malfois,M., Koch,M.H., Wigneshweraraj,S.R. and Buck,M. (2000) Low resolution structure of the sigma54 transcription factor revealed by X-ray solution scattering. *J. Biol. Chem.*, **275**, 4210–4214.
17. Chen,B., Sysoeva,T.A., Chowdhury,S., Guo,L., De Carlo,S., Hanson,J.A., Yang,H. and Nixon,B.T. (2010) Engagement of arginine finger to ATP triggers large conformational changes in NtrC1 AAA+ ATPase for remodeling bacterial RNA polymerase. *Structure*, **18**, 1420–1430.
18. De Carlo,S., Chen,B., Hoover,T.R., Kondrashkina,E., Nogales,E. and Nixon,B.T. (2006) The structural basis for regulated assembly and function of the transcriptional activator NtrC. *Genes Dev.*, **20**, 1485–1495.
19. Lee,S.Y., De La Torre,A., Yan,D., Kustu,S., Nixon,B.T. and Wemmer,D.E. (2003) Regulation of the transcriptional activator NtrC1: structural studies of the regulatory and AAA+ ATPase domains. *Genes Dev.*, **17**, 2552–2563.
20. Rappas,M., Schumacher,J., Niwa,H., Buck,M. and Zhang,X. (2006) Structural basis of the nucleotide driven conformational changes in the AAA+ domain of transcription activator PspF. *J. Mol. Biol.*, **357**, 481–492.
21. Bordes,P., Wigneshweraraj,S.R., Schumacher,J., Zhang,X., Chaney,M. and Buck,M. (2003) The ATP hydrolyzing transcription activator phage shock protein F of Escherichia coli: identifying a surface that binds sigma 54. *Proc. Natl Acad. Sci. USA*, **100**, 2278–2283.
22. Buck,M., Gallegos,M.T., Studholme,D.J., Guo,Y. and Gralla,J.D. (2000) The bacterial enhancer-dependent sigma(54) (sigma(N)) transcription factor. *J. Bacteriol.*, **182**, 4129–4136.
23. Sasse-Dwight,S. and Gralla,J.D. (1988) Probing the Escherichia coli glnALG upstream activation mechanism in vivo. *Proc. Natl Acad. Sci. USA*, **85**, 8934–8938.
24. Cannon,W., Gallegos,M.T. and Buck,M. (2001) DNA melting within a binary sigma(54)-promoter DNA complex. *J. Biol. Chem.*, **276**, 386–394.
25. Gallegos,M.T. and Buck,M. (1999) Sequences in sigmaN determining holoenzyme formation and properties. *J. Mol. Biol.*, **288**, 539–553.

26. Erzberger,J.P. and Berger,J.M. (2006) Evolutionary relationships and structural mechanisms of AAA+ proteins. *Annu. Rev. Biophys. Biomol. Struct.*, **35**, 93–114.
27. Joly,N., Engl,C., Jovanovic,G., Huvet,M., Toni,T., Sheng,X., Stumpf,M.P. and Buck,M. (2010) Managing membrane stress: the phage shock protein (Psp) response, from molecular mechanisms to physiology. *FEMS Microbiol. Rev.*, **34**, 797–827.
28. Zhang,N., Joly,N., Burrows,P.C., Jovanovic,M., Wigneshweraraj,S.R. and Buck,M. (2009) The role of the conserved phenylalanine in the sigma54-interacting GAFTGA motif of bacterial enhancer binding proteins. *Nucleic Acids Res.*, **37**, 5981–5992.
29. Joly,N. and Buck,M. (2011) Single chain forms of the enhancer binding protein PspF provide insights into geometric requirements for gene activation. *J. Biol. Chem.*, **286**, 12734–12742.
30. Farrell,I.S., Toroney,R., Hazen,J.L., Mehl,R.A. and Chin,J.W. (2005) Photo-cross-linking interacting proteins with a genetically encoded benzophenone. *Nat. Methods*, **2**, 377–384.
31. Dorman,G. and Prestwich,G.D. (1994) Benzophenone photophores in biochemistry. *Biochemistry*, **33**, 5661–5673.
32. Lin,A.A., Sastri,V.R., Tesoro,G. and Reiser,A. (1988) On the cross-linking mechanism of benzophenone-containing polyimides. *Macromolecules*, **21**, 1165–1169.
33. Cannon,W.V., Gallegos,M.T. and Buck,M. (2000) Isomerization of a binary sigma-promoter DNA complex by transcription activators. *Nat. Struct. Biol.*, **7**, 594–601.
34. Casaz,P. and Buck,M. (1997) Probing the assembly of transcription initiation complexes through changes in sigmaN protease sensitivity. *Proc. Natl Acad. Sci. USA*, **94**, 12145–12150.
35. Buck,M. and Cannon,W. (1994) A simple procedure for visualising protein–nucleic acid complexes by photochemical crosslinking. *Nucleic Acids Res.*, **22**, 1119–1120.
36. Chaney,M., Grande,R., Wigneshweraraj,S.R., Cannon,W., Casaz,P., Gallegos,M.T., Schumacher,J., Jones,S., Elderkin,S., Dago,A.E. *et al.* (2001) Binding of transcriptional activators to sigma 54 in the presence of the transition state analog ADP-aluminum fluoride: insights into activator mechanochemical action. *Genes Dev.*, **15**, 2282–2294.
37. Leach,R.N., Gell,C., Wigneshweraraj,S., Buck,M., Smith,A. and Stockley,P.G. (2006) Mapping ATP-dependent activation at a sigma54 promoter. *J. Biol. Chem.*, **281**, 33717–33726.
38. Lee,J.H. and Hoover,T.R. (1995) Protein crosslinking studies suggest that Rhizobium meliloti C4-dicarboxylic acid transport protein D, a sigma 54-dependent transcriptional activator, interacts with sigma 54 and the beta subunit of RNA polymerase. *Proc. Natl Acad. Sci. USA*, **92**, 9702–9706.
39. Burrows,P.C., Joly,N. and Buck,M. (2010) A prehydrolysis state of an AAA+ ATPase supports transcription activation of an enhancer-dependent RNA polymerase. *Proc. Natl Acad. Sci. USA*, **107**, 9376–9381.
40. Chaney,M. and Buck,M. (1999) The sigma 54 DNA-binding domain includes a determinant of enhancer responsiveness. *Mol. Microbiol.*, **33**, 1200–1209.
41. Joly,N., Schumacher,J. and Buck,M. (2006) Heterogeneous nucleotide occupancy stimulates functionality of phage shock protein F, an AAA+ transcriptional activator. *J. Biol. Chem.*, **281**, 34997–35007.
42. Werbeck,N.D., Schlee,S. and Reinstein,J. (2008) Coupling and dynamics of subunits in the hexameric AAA+ chaperone ClpB. *J. Mol. Biol.*, **378**, 178–190.
43. Bordes,P., Wigneshweraraj,S.R., Chaney,M., Dago,A.E., Morett,E. and Buck,M. (2004) Communication between Esigma(54), promoter DNA and the conserved threonine residue in the GAFTGA motif of the PspF sigma-dependent activator during transcription activation. *Mol. Microbiol.*, **54**, 489–506.
44. Wigneshweraraj,S.R., Casaz,P. and Buck,M. (2002) Correlating protein footprinting with mutational analysis in the bacterial transcription factor sigma54 (sigmaN). *Nucleic Acids Res.*, **30**, 1016–1028.
45. Darst,S.A., Opalka,N., Chacon,P., Polyakov,A., Richter,C., Zhang,G. and Wriggers,W. (2002) Conformational flexibility of bacterial RNA polymerase. *Proc. Natl Acad. Sci. USA*, **99**, 4296–4301.
46. Mekler,V., Kortkhonjia,E., Mukhopadhyay,J., Knight,J., Revyakin,A., Kapanidis,A.N., Niu,W., Ebright,Y.W., Levy,R. and Ebright,R.H. (2002) Structural organization of bacterial RNA polymerase holoenzyme and the RNA polymerase-promoter open complex. *Cell*, **108**, 599–614.
47. Mooney,R.A. and Landick,R. (2003) Tethering sigma70 to RNA polymerase reveals high in vivo activity of sigma factors and sigma70-dependent pausing at promoter-distal locations. *Genes Dev.*, **17**, 2839–2851.
48. Nagai,H. and Shimamoto,N. (1997) Regions of the Escherichia coli primary sigma factor sigma70 that are involved in interaction with RNA polymerase core enzyme. *Genes Cells*, **2**, 725–734.
49. Chen,J., Darst,S.A. and Thirumalai,D. (2010) Promoter melting triggered by bacterial RNA polymerase occurs in three steps. *Proc. Natl Acad. Sci. USA*, **107**, 12523–12528.
50. Enemark,E.J. and Joshua-Tor,L. (2006) Mechanism of DNA translocation in a replicative hexameric helicase. *Nature*, **442**, 270–275.
51. Bowman,G.D., O'Donnell,M. and Kuriyan,J. (2004) Structural analysis of a eukaryotic sliding DNA clamp–clamp loader complex. *Nature*, **429**, 724–730.
52. Costa,A., Ilves,I., Tamberg,N., Petojevic,T., Nogales,E., Botchan,M.R. and Berger,J.M. (2011) The structural basis for MCM2-7 helicase activation by GINS and Cdc45. *Nat. Struct. Mol. Biol.*, **18**, 471–477.
53. Lander,G.C., Estrin,E., Matyskiela,M.E., Bashore,C., Nogales,E. and Martin,A. (2012) Complete subunit architecture of the proteasome regulatory particle. *Nature*, **482**, 186–191.
54. Meinke,G., Bullock,P.A. and Bohm,A. (2006) Crystal structure of the simian virus 40 large T-antigen origin-binding domain. *J. Virol.*, **80**, 4304–4312.
55. Skordalakes,E. and Berger,J.M. (2003) Structure of the Rho transcription terminator: mechanism of mRNA recognition and helicase loading. *Cell*, **114**, 135–146.
56. Story,R.M., Weber,I.T. and Steitz,T.A. (1992) The structure of the E. coli recA protein monomer and polymer. *Nature*, **355**, 318–325.
57. Guo,Y., Lew,C.M. and Gralla,J.D. (2000) Promoter opening by sigma(54) and sigma(70) RNA polymerases: sigma factor-directed alterations in the mechanism and tightness of control. *Genes Dev.*, **14**, 2242–2255.
58. North,A.K., Weiss,D.S., Suzuki,H., Flashner,Y. and Kustu,S. (1996) Repressor forms of the enhancer-binding protein NrtC: some fail in coupling ATP hydrolysis to open complex formation by sigma 54-holoenzyme. *J. Mol. Biol.*, **260**, 317–331.
59. Chen,B., Doucleff,M., Wemmer,D.E., De Carlo,S., Huang,H.H., Nogales,E., Hoover,T.R., Kondrashkina,E., Guo,L. and Nixon,B.T. (2007) ATP ground- and transition states of bacterial enhancer binding AAA+ ATPases support complex formation with their target protein, sigma54. *Structure*, **15**, 429–440.
60. Bush,M., Ghosh,T., Tucker,N., Zhang,X. and Dixon,R. (2011) Transcriptional regulation by the dedicated nitric oxide sensor, NorR: a route towards NO detoxification. *Biochem. Soc. Trans.*, **39**, 289–293.
61. Jovanovic,M., James,E.H., Burrows,P.C., Rego,F.G., Buck,M. and Schumacher,J. (2011) Regulation of the co-evolved HrpR and HrpS AAA+ proteins required for Pseudomonas syringae pathogenicity. *Nat. Commun.*, **2**, 177.
62. Sallai,L. and Tucker,P.A. (2005) Crystal structure of the central and C-terminal domain of the sigma(54)-activator ZraR. *J. Struct. Biol.*, **151**, 160–170.
63. Wikstrom,P., O'Neill,E., Ng,L.C. and Shingler,V. (2001) The regulatory N-terminal region of the aromatic-responsive transcriptional activator DmpR constrains nucleotide-triggered multimerisation. *J. Mol. Biol.*, **314**, 971–984.
64. Joly,N. and Buck,M. (2010) Engineered interfaces of an AAA+ ATPase reveal a new nucleotide-dependent coordination mechanism. *J. Biol. Chem.*, **285**, 15178–15186.