The Role of Bacterial Enhancer Binding Proteins as Specialized Activators of $\sigma^{54}$-Dependent Transcription

Matthew Bush and Ray Dixon

Department of Molecular Microbiology, John Innes Centre, Norwich Research Park, Norwich, United Kingdom

INTRODUCTION

Transcription initiation in bacteria requires sequence-specific promoter recognition by sigma factors and a mechanism for the localized melting of the promoter DNA to provide a single-stranded template to initiate RNA synthesis. Not only do sigma factors direct the binding of the polymerase to specific promoters, they also mediate the DNA duplex-melting event (30). Unlike the $\sigma^{70}$ family, which is represented by a diverse group of housekeeping and alternative sigma factors, the $\sigma^{54}$ family contains just a single member, $\sigma^{54}$, that shows little sequence similarity to the $\sigma^{70}$ class (30,139). Although members of both families associate with the same core RNA polymerase (RNAP) enzyme, the resulting holoenzymes activate transcription by entirely different mechanisms (Fig. 1). In contrast to transcription initiation mediated by $\sigma^{70}$-like sigma factors, $\sigma^{54}$-dependent transcription absolutely requires the presence of an activator that couples the energy generated from ATP hydrolysis to the isomerization of the RNAP-$\sigma^{54}$ closed complex (CC) (187). Such activators typically bind at sites 80 to 150 bp upstream of the promoter, known as upstream activator sequences (UASs) or enhancer sites. This is similar to the
binding of eukaryotic enhancer binding proteins (EBPs), and so activators of σ54-dependent transcription are referred to as bacterial enhancer binding proteins (bEBPs) (reviewed in references 118, 171, 186, 217, and 235). σ54-dependent activators are sufficiently similar in structure and function to be classified as members of the AAA+ (ATPases associated with various cellular activities) family of proteins (151). AAA+ proteins are universal in living organisms, functioning as molecular machines to convert the chemical energy stored in ATP into a mechanical energy that can be used in various cellular processes (92, 133, 151, 155, 208). Examples include the HslU (177) and Lon (24) proteases, the protein-disaggregating chaperone ClpB (144), minichromosome maintenance (MCM) proteins (77), the DNA helicase RuvB (96, 108), and p97, implicated in numerous cellular processes, including the processing of ubiquitylated proteins en route to the proteasome (16, 236). Single-molecule experiments with optical tweezers have directly demonstrated that the AAA+ protease ClpX uses the energy derived from ATP hydrolysis to generate the force necessary to unfold and translocate its protein substrates (7, 134).

Since bEBPs bind relatively far upstream of the transcriptional start site, DNA must bend between the enhancer and the promoter site in order for the activator to interact with the RNAP-σ54 holoenzyme (Fig. 2) (186, 218, 235). Such DNA looping has been visualized by electron microscopy (EM) (201) and scanning force microscopy (SFM) (176). DNA looping is often aided by the integration host factor (IHF) or HU, small heterodimeric proteins which bind between promoter and enhancer sites to bend the DNA up to 180° (6, 47, 99, 230). Since correct interfacing between the bEBP and the holoenzyme is crucial for the activation process, the phasing of the IHF binding site relative to the promoter is important (53, 104). In addition to ensuring efficient activator-holoenzyme contact, IHF-induced changes in DNA topology contribute to the specificity and efficiency of activation (71, 163). For example, at the psaA promoter, IHF has been shown to mediate architectural changes that aid the binding of the bEBP PspF and increase promoter output (120). Once DNA bending has been induced, the bEBP utilizes nucleotide triphosphate (NTP) hydrolysis to drive conformational rearrangements in the holoenzyme that promote the transition of the closed complex to an open complex (OC) (Fig. 2C) (173, 187). Although ATP binding to a bEBP dimer has been demonstrated (180), oligomerization is required to stimulate full ATPase activity (172, 222). Well-characterized examples of bEBPs include nitrogen regulatory protein C (NtrC), C4-dicarboxylic acid transport protein D (DctD), nitrogen fixation regulatory protein (NifA), phage shock protein F (PspF), xylene catabolism regulatory protein (XylR), and 3,4-dimethylphenol catabolism regulatory protein (DmpR) (Table 1) (200). In this review, we first discuss the role of the sigma factor in the activation of bacterial transcription, with particular focus on the alternative sigma factor σ54. The features of σ54-dependent transcription that lead to the requirement of the activator will be introduced before an in-depth look at the structure and function of bEBPs.

FEATURES THAT DISTINGUISH σ54 FROM THE σ70-LIKE FAMILY OF SIGMA FACTORS

Alternative sigma factors such as σ24, σ32, σ28, σ19, and σ34 are all members of the σ54 class. Each member binds to conserved −10 and −35 promoter elements, although the consensus sequences and spacing differ for each sigma factor. The Eσ70 holoenzyme recognizes and binds to the consensus sequences TTGACA at the −35 element and TATAAT at the −10 element, and the spacing between these sequences is crucial for transcription initiation (93). In contrast, σ54 binds to different consensus sequences that are more strongly conserved than those for σ70. Binding occurs at the GG −24 and TGC −12 elements (146) that are part of the wider consensus sequence YTGCCACGrNNTTGGC (where uppercase type indicates highly conserved residues, lowercase type indicates weakly conserved residues, N is nonconserved, Y is pyrimidines, R is purines, and W is A or T) (10). This forms an energetically favorable CC that rarely isomerizes into the OC. In order to form the transcription “bubble,” a specialized activator (a bacterial enhancer binding protein [bEBP]) must bind and use the energy from ATP hydrolysis to remodel the holoenzyme.

![Diagram showing initiation of transcription by the RNAP-σ70 (A) and RNAP-σ54 (B) holoenzymes. The σ70 factor directs the binding of polymerase to the consensus −10 (TATAAT) and −35 (TTGACA) sequences to form an energetically unfavorable closed complex (CC) that is readily converted into an open complex (OC) to initiate transcription. In contrast, the σ54 factor directs the binding of RNAP to conserved −12 (TGC) and −24 (GG) promoter elements that are part of the wider consensus sequence YTGCCACGrNNTTGGC (where uppercase type indicates highly conserved residues, lowercase type indicates weakly conserved residues, N is nonconserved, Y is pyrimidines, R is purines, and W is A or T) (10). This forms an energetically favorable CC that rarely isomerizes into the OC. In order to form the transcription “bubble,” a specialized activator (a bacterial enhancer binding protein [bEBP]) must bind and use the energy from ATP hydrolysis to remodel the holoenzyme.](http://mmbr.asm.org/)

FIG 1 Initiation of transcription by the RNAP-σ70 (A) and RNAP-σ54 (B) holoenzymes. The σ70 factor directs the binding of polymerase to the consensus −10 (TATAAT) and −35 (TTGACA) sequences to form an energetically unfavorable closed complex (CC) that is readily converted into an open complex (OC) to initiate transcription. In contrast, the σ54 factor directs the binding of RNAP to conserved −12 (TGC) and −24 (GG) promoter elements that are part of the wider consensus sequence YTGCCACGrNNTTGGC (where uppercase type indicates highly conserved residues, lowercase type indicates weakly conserved residues, N is nonconserved, Y is pyrimidines, R is purines, and W is A or T) (10). This forms an energetically favorable CC that rarely isomerizes into the OC. In order to form the transcription “bubble,” a specialized activator (a bacterial enhancer binding protein [bEBP]) must bind and use the energy from ATP hydrolysis to remodel the holoenzyme.
conserved elements is critical for 70-dependent transcription. Therefore, the holoenzyme absolutely requires these motifs to be positioned on the same side of the DNA helix in order to function (28, 124, 152).

Both the 70-type and the alternate 54 factors form holoenzyme-promoter complexes with a default closed and nonproductive form (90). However, the requirements for the formation of an open promoter complex differ (Fig. 1). The 70 holoenzyme binds to the consensus −10 and −35 sequences to form an energetically unfavorable closed complex (CC) that is readily converted into an open complex (OC) without a requirement for activators (90). Indeed, recent crystallographic studies indicated that the recognition of the promoter by 70 is closely coupled to the thermally driven nucleation of DNA melting, since 70 recognizes only the melted state of the −10 element, in which bases on the non-template strand are flipped out of the base stack (74). A similar situation may apply in the case of 54, whereby the unpairing of the duplex at the −12 and −11 elements, to form an early-melted (fork junction-like) structure, favors recognition by the sigma factor. However, it appears that, in contrast to 70, the interaction of 54 with the −12/−11 fork junction prevents the binding of the holoenzyme to the non-template strand, a key step in the DNA-melting process (30, 90, 149). As a result, the 54 holoenzyme binds promoter sequences tightly in such a way that isomerization is not spontaneous; the holoenzyme is transcriptionally silent (39, 46). Therefore, the initiation of 54-dependent transcription is unique in that it requires an activator of the AAA+ class that couples the energy produced from ATP hydrolysis to change the structure of the nucleoprotein complex located close to the start site of DNA melting. This energy-dependent step ensures the removal of the inhibitory interaction at the −12/−11 position so that the thermodynamic and kinetic barriers restricting open complex formation in the 54 system are overcome (42, 45, 90, 235).

A sequence comparison of members of the 70 family reveals four conserved helical domains (σ1, σ2, σ3, and σ4) connected by flexible linkers. Each domain can be divided into functionally distinct subregions that have roles in promoter recognition, core binding, and isomerization (95, 149). In contrast, 54 is composed of three regions, based primarily on function rather than structure (Fig. 3) (21, 30, 44, 84). Region I (residues 1 to 56 in Escherichia coli) is commonly a glutamine- and leucine-rich sequence that represents the regulatory domain of 54. Region II (residues 57 to 107 in E. coli) is variable in amino acid composition and length, ranging from 26 residues in Rhodobacter capsulatus to 110 residues in Bradyrhizobium japonicum, but can be characterized by the predominance of acidic residues (196). This region is not essential for 54-dependent transcription; region II is virtually absent in some bacterial species, e.g., Bacillus subtilis (30). However, deletions in Klebsiella pneumoniae 54 region II significantly impair the activity of the holoenzyme in open complex formation (196). Recent evidence suggests that region II has roles in DNA binding (40) and DNA melting (224). The C-terminal region III (residues 108 to 477 in E. coli) is well conserved, containing the major determinants for binding to promoter DNA (residues 329 to 463 in E. coli) (30, 34, 35, 219). These determinants include a DNA cross-linking motif (residues 329 to 346 in E. coli) (41) and an RpoN box (residues 454 to 463 in E. coli) (205). Recently, nuclear magnetic resonance (NMR)-based structural studies of the C-terminal region of Aquifex aeolicus 54 bound to promoter DNA have revealed a helix-turn-helix (HTH) motif in which the RpoN box forms the recognition helix that binds to the −24 promoter element (67, 68). An additional HTH motif, previously suggested to bind to the
<table>
<thead>
<tr>
<th>bEBP Description</th>
<th>Organism</th>
<th>Function (target)</th>
<th>Regulatory domain(s)</th>
<th>Stimulus/mode of regulation</th>
<th>Regulation(s) of AAA⁺ activity</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phage shock protein F</td>
<td><em>Escherichia coli</em></td>
<td>Phage shock response (Er²⁴)</td>
<td>Absent</td>
<td>Interaction in trans with PspA (i) Small-ligand binding (2-oxoglutarate), (ii) in trans in response to NifL (i) Positive (by 2-OG), (ii) negative (by NifL)</td>
<td>Negative (by PspA)</td>
<td>72, 73, 113, 114</td>
</tr>
<tr>
<td>Nitrogen fixation regulatory protein</td>
<td><em>Azotobacter vinelandii</em></td>
<td>Regulation of nitrogenase (Er²⁴)</td>
<td>GAF domain</td>
<td>Phosphorylation by sensor kinase NtrB</td>
<td>Positive</td>
<td>136, 137</td>
</tr>
<tr>
<td>Nitrogen regulatory protein C</td>
<td><em>Salmonella enterica</em></td>
<td>Regulation of glutamine synthetase (and other functions of nitrogen metabolism) (Er²⁴)</td>
<td>RR</td>
<td>Phosphorylation</td>
<td>Negative</td>
<td>8, 62, 66, 216</td>
</tr>
<tr>
<td>Nitrogen regulatory protein C1</td>
<td><em>Aquifex aeolicus</em></td>
<td>Homologue of NtrC (Er²⁴)</td>
<td>RR</td>
<td>Phosphorylation</td>
<td>Negative (R domain releases inhibition of C domain)</td>
<td>11, 12</td>
</tr>
<tr>
<td>Nitrogen regulatory protein C4</td>
<td><em>Aquifex aeolicus</em></td>
<td>Homologue of NtrC (Er²⁴)</td>
<td>RR</td>
<td>Phosphorylation</td>
<td>Negative (R domain releases inhibition of C domain)</td>
<td>11, 12</td>
</tr>
<tr>
<td>Xylene catabolism regulatory protein</td>
<td><em>Pseudomonas putida</em></td>
<td>Regulation of oxidative transformation of toluene, m-xylene, and p-xylene (Er²⁴)</td>
<td>V4R domain</td>
<td>Small-ligand binding (aromatic effectors)</td>
<td>Negative (R domain releases inhibition of C domain)</td>
<td>76, 161, 162</td>
</tr>
<tr>
<td>C₄-Dicarboxylic acid transport protein D</td>
<td><em>Sinorhizobium meliloti</em></td>
<td>Regulation of C₄-dicarboxylic acid transport (Er²⁴)</td>
<td>RR</td>
<td>Phosphorylation</td>
<td>Negative (R domain releases inhibition of C domain)</td>
<td>66, 226, 228</td>
</tr>
<tr>
<td>3,4-Demethylphenol catabolism regulatory protein</td>
<td><em>Pseudomonas sp. strain CF600</em></td>
<td>Regulation of (methyl)phenol catabolism (Er²⁴)</td>
<td>V4R domain</td>
<td>Small-ligand binding (aromatic effectors)</td>
<td>Negative (R domain releases inhibition of C domain)</td>
<td>193, 194, 222</td>
</tr>
<tr>
<td>Nitric oxide reductase regulator</td>
<td><em>Escherichia coli</em></td>
<td>Regulation of NO detoxification system (Er²⁴)</td>
<td>GAF domain</td>
<td>Small-ligand binding (nitric oxide)</td>
<td>Negative (R domain releases inhibition of C domain)</td>
<td>36, 37, 60, 106, 206</td>
</tr>
<tr>
<td>Zinc resistance-associated protein regulator</td>
<td><em>Salmonella Typhimurium</em></td>
<td>Regulation of heavy metal tolerance system (Er²⁴)</td>
<td>RR</td>
<td>Phosphorylation by sensor kinase ZraS</td>
<td>Unknown</td>
<td>130, 181</td>
</tr>
<tr>
<td>Formate-hydrogen lyase activator protein</td>
<td><em>Escherichia coli</em></td>
<td>Regulates formate-hydrogen lyase (which oxidizes formic acid) (Er²⁴)</td>
<td>Two GAF domains</td>
<td>Small-ligand binding (formate)</td>
<td>Negative</td>
<td>100, 183</td>
</tr>
<tr>
<td>Tyrosine regulator</td>
<td><em>Escherichia coli</em></td>
<td>Regulation of aromatic amino acid biosynthesis and transport (Er²⁴)</td>
<td>PAS and ACT domains</td>
<td>Small-ligand binding (phenylalanine, tyrosine, or tryptophan) in the presence of ATP</td>
<td>ATPase inactive (activates or represses Er²⁴ promoters)</td>
<td>164, 165</td>
</tr>
<tr>
<td>Hypersensitive reaction pathogenicity island regulator</td>
<td><em>Pseudomonas syringae</em></td>
<td>Regulation of <em>hrp-hrc</em> pathogenicity island (Er²⁴)</td>
<td>Absent</td>
<td>Interaction with apparent sensor kinase (HrpS)</td>
<td>Negative (HrpS regulated by HrpV; HrpR regulated by HrpG)</td>
<td>105, 122, 170, 188</td>
</tr>
<tr>
<td>Flagellar gene regulator</td>
<td><em>Helicobacter pylori</em></td>
<td>Regulates class III flagellar genes; only bEBP present in organism, activates from solution (no D domain) (Er²⁴)</td>
<td>RR</td>
<td>Phosphorylation by sensor kinase FlgS</td>
<td>Negative (release of repression)</td>
<td>26, 197</td>
</tr>
</tbody>
</table>
FIG 3 Domain organization of $\sigma^{54}$. E. coli $\sigma^{54}$ (residues 1 to 477) consists of 3 regions (regions I to III). DNA binding motifs include the DNA cross-linking (X-link) region, the helix-turn-helix (HTH) motif, and the RpoN box, all present at the C terminus. Region I interacts with the activator of transcription. Region II is often acidic and occasionally absent. The location of the main core RNAP binding determinants (residues 120 to 215) is shown (30).

Structural Basis for the Activator Dependence of $\sigma^{54}$

Although high-resolution crystal structures have been determined for $\sigma^{70}$ family members (38,109, 150, 195), $\sigma^{54}$ has yielded only low-resolution small-angle X-ray scattering (SAXS) and cryo-electron microscopy (cryo-EM) structures (23, 172, 202). Recently, cryo-EM has revealed structural features of the core binding region of $\sigma^{54}$ holoenzyme that explain the stability of the closed complex and, therefore, the need for an activator (23,84). Reconstructions of RNAP-$\sigma^{54}$ in the presence and absence of an activator protein have identified three distinct structural regions of $\sigma^{54}$ (named D1, D2, and D3 by Bose et al. [23]), each positioned on the $\beta'$ side and on the upstream face of the core RNA polymerase (23). The D1 region likely represents the core RNAP binding domain (residues 120 to 215 in E. coli) of region III and is located at the tip of the $\beta'$ subunit, well positioned to contact the $\beta'$-coiled-coil motif, which is the binding site of $\sigma$ factors in the core enzyme (231). This is consistent with NMR studies of the core binding region of Aquifex aeolicus $\sigma^{54}$, in which one surface is negatively charged and predicted to interact with the coiled-coil motif of $\beta'$ (98). Bose et al. attributed the density of the D3 region to the DNA binding domain of region III, which includes the RpoN box (23). Of particular importance is the presence of a strong bridging density (Db) connecting the two “pinchers” of the polymerase enzyme that is more pronounced in RNAP-$\sigma^{54}$ than in an RNAP-$\sigma^{70}$ structure. Since the Db region correlates with the −12 position of the promoter DNA, it was proposed that this connecting density (attributed to region I) obstructs the loading of DNA into the active-site channel of the core enzyme. Therefore, the presence of the bridging density (Db) could explain why the $\sigma^{54}$ holoenzyme forms an energetically favorable closed complex, unlike the $\sigma^{70}$ holoenzyme, which spontaneously isomerizes into an open form. A comparison of the RNAP-$\sigma^{54}$ reconstruction with one also in the presence of an activator and nucleotide transition-state analogue revealed a significant conformational change in region I upon activator binding that is coupled to a movement of the DNA toward the active site of the polymerase (Fig. 4) (23). Indeed, hydroxyl radical footprinting and photo-cross-linking have demonstrated that the ATPase domain of the activator is within 12 Å of the −12 promoter element during open complex formation (34). This is consistent with a role for the activator in the remodeling of the nucleoprotein regulatory center (21, 49, 220). Based on studies of $\sigma^{54}$ (residues 69 to 198) from Aquifex aeolicus, it was also suggested that the hydrophobic interface between the N-terminal and C-terminal subdomains of the core binding region (D1 in the reconstruction by Bose et al.) might be disrupted upon activator binding to region I of $\sigma^{54}$, paving the way for conformational changes in the holoenzyme and the isomerization of the closed complex (98). Overall, structural and biochemical studies have revealed three main roles for the activator in $\sigma^{54}$-dependent transcription (23, 84). First, the activator must stimulate DNA melting at the −12 promoter element, since activator bypass mutants function only in the presence of premelted DNA (49,89). Second, the activator must remodel region I of the $\sigma$ factor, which physically blocks the loading of the DNA into the active site (23). Third, the activator must cause the repositioning of the DNA binding domains of $\sigma^{54}$ downstream, since the −12 promoter element at which DNA melting originates is located too far upstream from the active site of the core enzyme for elongation to proceed (23).

bEBPs ARE SPECIALIZED ACTIVATORS OF $\sigma^{54}$-DEPENDENT TRANSCRIPTION

bEBPs are modular proteins and in general consist of three domains (186, 200). The N-terminal regulatory (R) domain has a role in signal perception and modulates the activity of the bEBP. The central AAA$^+$ domain (C) is responsible for ATP hydrolysis; it is indispensable and often sufficient to activate $\sigma^{54}$-dependent transcription (15,121, 222, 226). Lastly, the C-terminal DNA binding domain (D) contains a helix-turn-helix (HTH) motif (160, 181) that enables specific UAS/enhancer site recognition (227). However, not all activators of $\sigma^{54}$-dependent transcription consist of each of the three domains (Fig. 5). While the presence of the central domain is conserved, some bEBPs lack either the regulatory domain (i.e., consisting of C plus D) or the DNA binding domain (i.e., consisting of R plus C) (26, 73, 105). In addition, the regulatory domains do not share a common homology and contain a variety of sensory motifs...
depending on the signal that is detected (200). Consequently, the bEBP family has been divided into five groups (I to V) based on the organizations of the three domains that are present (Fig. 5) (218). The structures and functions of these three domains will now be discussed in detail.

**ROLE OF THE CENTRAL AAA⁺ DOMAIN**

The central (AAA⁺) domain of bEBPs is responsible for nucleotide binding and hydrolysis, oligomerization, and σ⁵₄ contact. The AAA⁺ domain is the most conserved of the three domains and has been divided into seven conserved regions, C1 to C7 (Fig. 6) (148, 157). While many of the structural features of this domain are common to the AAA⁺ superfamily, bEBPs contain two unique insertions that form surface-exposed loops. Loop 1 (L1) is formed from an insertion that projects out of an α-helix in region C3 and contains a highly conserved GAFTGA motif, whereas loop 2 (L2) is inserted between C5 and C6 (Fig. 6). Both of these loops serve an important function in engagements with σ⁵₄ (see below).

**Conserved Elements in the AAA⁺ Superfamily**

Crystal structures of the PspF, NtrC, and ZraR bEBP AAA⁺ domains (128, 172, 181) reveal monomers arranged in a ring (Fig. 7B), each consisting of an α/β subdomain followed by a smaller α-helical subdomain (Fig. 7C), characteristic of all AAA⁺ proteins (171, 235). The nucleotide binding site is located in the cleft between these subdomains and between two adjacent protomers (Fig. 7A) (171). AAA⁺ domains are characterized by the Walker A and Walker B (WB) motifs, which have roles in nucleotide binding and hydrolysis, respectively (92, 210). The Walker A motif is in the C1 region (Fig. 6) and forms a P loop with the consensus sequence GxxxGK(T/S), which interacts with the phosphates of ATP (182). The requirement for the Walker A motif has been shown for a number of bEBPs, including PspF (187) and NtrC (180). In Pseudomonas putida XylR, the G268N substitution abolishes Pseudomonas putida XylR, the G268N substitution abolishes ATP binding and hydrolysis (162).

Likewise, the Walker B motif of the C4 region (Fig. 6) has a consensus sequence of hhhhDE (where h is any hydrophobic amino acid) and has been shown to be required for nucleotide hydrolysis. The mutagenesis of the key aspartate residue suggests a role in the coordination of Mg²⁺, which is required for ATP hydrolysis (180, 187). Another common feature of AAA⁺ proteins is the presence of the sensor I and sensor II motifs, which are present in the conserved regions C6 and C7, respectively (186). Sensor I residues are located within a loop containing a conserved threonine residue that interacts with the second acidic residue of the WB motif via a water molecule. This threonine residue has been implicated in the coupling of nucleotide hydrolysis to conformational changes in surface-exposed L1 and L2 (185). Sensor II residues within the C7 region are located in the third helix of the α-helical subdomain and have been implicated in nucleotide binding. I226 in PspF has been suggested to be involved in this function, while the adjacent arginine residue points toward the γ-phosphate and may be involved in hydrolysis (186). Indeed, R227 in PspF was suggested to have a role in Mg²⁺ coordination (173, 235). Members of the AAA⁺ superfamily also contain one or two arginine residues (R fingers) that have been implicated in intersubunit catalysis and nucleotide sensing (92, 133, 155). In accordance with this, bEBPs contain two potential R fingers that together with the catalytically important sensor II residues are located at the protomer interface. In PspF, the predicted R finger R168 has been shown to be required for ATP hydrolysis but not for ATP binding (187), while the same has been observed for NtrC for the R-finger residue R294 (R162 in PspF) (179, 180). These phenotypes reflect the observation that the ATPase active site is formed at the interface between adjacent protomers to

![FIG 4](http://mmbr.asm.org/Downloaded from http://mmbr.asm.org on May 27, 2021 by guest)
FIG 5 Domain architecture of the five classical groups (groups I to V) of bEBPs (218). The central AAA+ domain (C) (red) is highly conserved and absolutely essential for σ54-dependent transcription. The C-terminal DNA binding domain (D) (green) consists of an HTH motif that directs the bEBP to specific UAS/enhancer binding sites and is absent in some bEBPs (group V). The N-terminal regulatory domain (R) is not well conserved between members of the bEBP family. Different sensory domains are present depending on the environmental signal to be detected, but in some bEBPs, they are absent (group IV). Group I bEBPs contain a response regulator (RR) domain (blue). Group II bEBPs contain Per, ARNT, and Sim (PAS) domains (orange) or XylR-N and V4R (vinyl 4 reductase) domains (pink). Group III bEBPs contain a cGMP-specific and stimulated phosphodiesterase, Anabaena adenylate cyclase, and E. coli FhlA (GAF) domain (purple). HrpR and HrpS are coactivators of transcription and therefore are grouped together.
The structure of the AAA⁺ protein p97 shows the catalytic arginine protruding from one protomer into the catalytic site of the adjacent protomer, contacting the γ-phosphate of ATP (236). Indeed, the recent publication of the crystal structure of ATP-bound NtrC₁ indicates that the second R finger (R₂₉₉ in NtrC₁) engages the γ-phosphate (51). Therefore, oligomerization is essential for the ability of bEBPs to hydrolyze ATP.

The GAFTGA Motif, a bEBP-Specific Structural Element

The bEBP subfamily of AAA⁺ domains contains specific structural features that enable nucleotide-dependent interactions with σ₅₄ (Fig. 7) (186, 235). Most conserved among these is the GAFTGA motif (in region C₃), which forms a loop on the surface of the AAA⁺ domain that contacts σ₅₄ during the ATP hydrolysis cycle (21). Crystal structures of the NtrC₁ and ZraR AAA⁺ domains (128, 181) show that the GAFTGA motif is located at the α/β subdomain surface at the tip of loop 1 (L₁), which is inserted into helix 3 (H₃) (Fig. 7C). Although the surface-exposed loops appear to point toward the central pore of the oligomeric rings, the GAFTGA motifs are not in an extended conformation. A stable interaction between the GAFTGA loop and σ₅₄ has been observed by studies that used the ATP transition-state analogue ADP·AlFx. Cryo-electron microscopy (cryo-EM) studies of the PspF central domain in complex with σ₅₄ and ADP·AlFx revealed a hexameric bEBP ring in contact with monomeric σ₅₄ (172). Significantly, the reconstruction reveals connecting electron densities between the bEBP and σ₅₄. The fitting of the PspF AAA⁺ crystal structure into the three-dimensional (3D) reconstruction confirms that the GAFTGA-containing L₁, assisted by loop 2 (L₂), mediates this interaction. Therefore, it is likely that these conserved motifs enable a nucleotide-dependent σ₅₄ interaction to initiate the transition of the closed complex. In accordance with this, the GAFTGA motif

FIG 6 Domain map and sequence alignment of the conserved regions of bEBP AAA⁺ domains (C1 to C7) (148). The conserved regions are based on a structure-based sequence alignment (186). Key residues (Walker A, “switch” Asn, GAFTGA, Walker B, and R fingers) are highlighted in yellow, and nonconsensus sequences in the alignments are highlighted in red. Alignments were conducted by using ClustalW (www.ebi.ac.uk/clustalw/), using the following sequences from UniProtKB/Swiss-Prot (http://www.expasy.ch/): PspF (E. coli), NifA (A. vinelandii), XyI (P. putida), DmpR (Pseudomonas sp.), NtrC (E. coli), ZraR (E. coli), NtcC₁ (A. aeolicus), NtcC₄ (A. aeolicus), FlgR (H. pylori), DctD (S. meliloti), FhlA (E. coli), HrpB (P. syringae), NorR (E. coli), and TyrR (E. coli). The R (regulatory) and D (DNA binding) domains are also illustrated, although they are not to scale.
has been shown to be critical for open complex formation (235). The effect of the substitution of residues of the GAFTGA motif has been studied with the bEBPs NtrC (131, 154, 229), DctD (214), NifA (86), DmpR (222), and PspF (20, 21, 49, 58, 234). There seems to be an absolute requirement for an intact GAFTGA motif; a mutation of any one of the six amino acids has severe effects on the ability to hydrolyze ATP, contact $\sigma^{54}$, or activate transcription (Table 2). Unusual variants of the bEBP family that lack this motif, such as E. coli TyrR and Rhodobacter capsulatus NtrC, are unable to activate $\sigma^{54}$-dependent transcription and instead have been shown to regulate transcription at $\sigma^{70}$-dependent promoters (25, 165, 166).

**Residues of the GAFTGA motif.** Sequence alignments of bEBP AAA$^+$ domains indicate a very high level of conservation for the GAFTGA motif (Fig. 6) (234), reflecting its importance in $\sigma^{54}$-dependent transcription. The first glycine residue of the motif appears to be absolutely conserved, although it has not been widely studied in bEBPs. In NtrC, random mutagenesis identified the G215V mutation, which abolished transcriptional activation both in vivo and in vitro (131). In DctD, a substitution of the equivalent residue (G220D) also produced an inactive protein in vivo (214).

Likewise, a substitution of the second amino acid of the GAFTGA motif in the bEBPs NtrC, NifA, and DctD gave rise to variants that were unable to activate transcription (86, 154, 214). However, these variants showed little or no reduction in ATPase activity, suggesting that while this residue may be required for contacting $\sigma^{54}$, it does not communicate with the ATP hydrolysis machinery. Interestingly, in the case of the NtrC A216C variant, the defect in transcriptional activation is relieved by additional substitutions in the helix-turn-helix motif that prevent specific binding to enhancer DNA (229). Therefore, it appears that binding to the enhancers prevents the A216C variant of NtrC from contacting $\sigma^{54}$, suggesting a relationship between enhancer binding and AAA$^+$ function.

The role of the phenylalanine in the GAFTGA motif has been extensively studied (Table 2). The mutagenesis of this residue in NtrC (F217), NifA (F307), DctD (F222), DmpR (F312), and PspF (F85) produced bEBPs that failed to activate transcription (21, 86, 131, 214, 222). The exception is the F307Y variant of NifA, which retained 20% of its activity in vivo (86). Indeed, 16 of 248 bEBPs identified in an alignment have a naturally occurring tyrosine at this position (234). This indicates that an aromatic ring at this position is essential for transcriptional activation. To further investigate the role of the phenylalanine of the GAFTGA motif in $\sigma^{54}$-dependent transcription, the F85 residue of PspF was systematically replaced with 10 other amino acid residues, and the functionality of the resulting variants was assessed in vitro (234). Each of the substitutions rendered the bEBP unable to activate transcription from the nifH promoter. The F85H, F85I, F85W, F85L,
<table>
<thead>
<tr>
<th>bEBP</th>
<th>Residue</th>
<th>Change</th>
<th>Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>NtrC</td>
<td>G215</td>
<td>V</td>
<td>Fails to activate transcription in vivo or in vitro</td>
<td>131</td>
</tr>
<tr>
<td>NtrC</td>
<td>G220</td>
<td>D</td>
<td>Fails to activate transcription in vivo</td>
<td>214</td>
</tr>
<tr>
<td>NtrC</td>
<td>A216</td>
<td>V</td>
<td>Fails to activate transcription despite little reduction in ATPase activity</td>
<td>154</td>
</tr>
<tr>
<td>NtrC</td>
<td>A216</td>
<td>C</td>
<td>Sufficient ATPase; increased oligomerization state; binds enhancer, but activity is perturbed by DNA</td>
<td>229</td>
</tr>
<tr>
<td>NifA</td>
<td>A306</td>
<td>D/N</td>
<td>Fails to activate transcription in vivo</td>
<td>86</td>
</tr>
<tr>
<td>DctD</td>
<td>A221</td>
<td>V/D</td>
<td>Fails to activate transcription in vivo</td>
<td>214</td>
</tr>
<tr>
<td>NtrC</td>
<td>F217</td>
<td>L</td>
<td>Fails to activate transcription in vivo or in vitro</td>
<td>131</td>
</tr>
<tr>
<td>NifA</td>
<td>F307</td>
<td>A/L/I/P/R/H/N</td>
<td>Fails to activate transcription in vivo</td>
<td>86</td>
</tr>
<tr>
<td>NifA</td>
<td>F307</td>
<td>Y</td>
<td>&lt;20% activity in vivo</td>
<td>86</td>
</tr>
<tr>
<td>DctD</td>
<td>F222</td>
<td>L</td>
<td>Fails to activate transcription in vivo or in vitro</td>
<td>214</td>
</tr>
<tr>
<td>DmpR</td>
<td>F312</td>
<td>L</td>
<td>Fails to activate transcription in vivo; ATPase 75-85% of WTΔNTD (ATP) and 23-25% WTΔNTD (dATP)</td>
<td>222</td>
</tr>
<tr>
<td>PspF</td>
<td>F85</td>
<td>A/E/R</td>
<td>&lt;1% activity of WT in in vitro transcription assays; does not form ADP,AlFx-dependent trapped complex; decreased ATPase activity; defective for oligomerization</td>
<td>21</td>
</tr>
<tr>
<td>PspF</td>
<td>F85</td>
<td>C</td>
<td>&lt;1% activity of WT in in vitro transcription assays; does not form ADP,AlFx-dependent trapped complex; decreased ATPase activity; oligomerizes in the presence/absence of nucleotide</td>
<td>234</td>
</tr>
<tr>
<td>PspF</td>
<td>F85</td>
<td>H/I/W</td>
<td>&lt;1% activity of WT in in vitro transcription assays; does not form ADP,AlFx-dependent trapped complex; WT ATPase activity; nucleotide-dependent oligomerization (~WT)</td>
<td>234</td>
</tr>
<tr>
<td>PspF</td>
<td>F85</td>
<td>L/Q</td>
<td>&lt;1% activity of WT in in vitro transcription assays; does not form ADP,AlFx-dependent trapped complex; decreased ATPase activity; nucleotide-dependent oligomerization (~WT)</td>
<td>234</td>
</tr>
<tr>
<td>PspF</td>
<td>F85</td>
<td>Y</td>
<td>&lt;1% activity of WT in in vitro transcription assays; forms ADP,AlFx-dependent trapped complex; decreased ATPase activity; oligomerizes in the presence/absence of nucleotide; cannot form activator-DNA-σ(^5) complex (phenotype rescued by G4L substitution in σ(^5))</td>
<td>234</td>
</tr>
<tr>
<td>NorR</td>
<td>F264</td>
<td>Y</td>
<td>Partial “escape” from repression in vivo; shows WT-like response</td>
<td>36</td>
</tr>
<tr>
<td>NtrC</td>
<td>T218</td>
<td>A/N</td>
<td>Fails to activate transcription in vivo/in vitro</td>
<td>131</td>
</tr>
<tr>
<td>NifA</td>
<td>T308</td>
<td>A/L/M/P/R/V/G/C/S</td>
<td>Fails to activate transcription in vivo</td>
<td>86</td>
</tr>
<tr>
<td>DctD</td>
<td>T223</td>
<td>I</td>
<td>Fails to activate transcription in vivo/in vitro; ATPase 123% of WTΔNTD</td>
<td>214</td>
</tr>
<tr>
<td>DctD</td>
<td>T223</td>
<td>A</td>
<td>Fails to activate transcription; significant ATPase retained</td>
<td>214</td>
</tr>
<tr>
<td>PspF</td>
<td>T86</td>
<td>A</td>
<td>&lt;1% activity of WT in in vitro transcription assays; wild-type ATPase activity</td>
<td>49</td>
</tr>
<tr>
<td>PspF</td>
<td>T86</td>
<td>S</td>
<td>Does not form ADP,AlFx-dependent trapped complex</td>
<td>21</td>
</tr>
<tr>
<td>PspF</td>
<td>T86</td>
<td>V</td>
<td>52% activity of WT in in vitro transcription assays; in vivo, ~25% of WT Region IG4L σ(^5) substitution restores transcription activation activity WT ATPase activity</td>
<td>21</td>
</tr>
<tr>
<td>PspF</td>
<td>T86</td>
<td>V</td>
<td>Forms ADP,AlFx-dependent trapped complex</td>
<td>58</td>
</tr>
<tr>
<td>NtrC</td>
<td>G219</td>
<td>K</td>
<td>50% ATPase but fails to activate transcription; improved DNA binding</td>
<td>154</td>
</tr>
<tr>
<td>NtrC</td>
<td>G219</td>
<td>C</td>
<td>Increased ATPase due to increased oligomerization; binding of enhancer DNA prevents transcription</td>
<td>229</td>
</tr>
<tr>
<td>NorR</td>
<td>G266</td>
<td>D/N</td>
<td>Complete “escape” from repression in vivo; shows increased activity relative to that of WT</td>
<td>36</td>
</tr>
<tr>
<td>NorR</td>
<td>G266</td>
<td>C/S/Q/M</td>
<td>Partial “escape” from repression in vivo; shows WT-like activities</td>
<td>36</td>
</tr>
<tr>
<td>NorR</td>
<td>G266</td>
<td>K/R/P/A/E/L/V/T/H/I/W/Y/F</td>
<td>Fails to activate transcription in vivo</td>
<td>36</td>
</tr>
<tr>
<td>NtrC</td>
<td>A220</td>
<td>T</td>
<td>Fails to activate transcription despite little reduction in ATPase activity</td>
<td>154</td>
</tr>
<tr>
<td>NtrC</td>
<td>A220</td>
<td>V</td>
<td>Does not activate transcription in vivo but shows “hyperactivity” at low concentrations in vitro</td>
<td>131</td>
</tr>
<tr>
<td>NifA</td>
<td>A310</td>
<td>S</td>
<td>&lt;20% activity in vivo</td>
<td>86</td>
</tr>
<tr>
<td>NifA</td>
<td>A310</td>
<td>N/D/G</td>
<td>Fails to activate transcription in vivo</td>
<td>86</td>
</tr>
<tr>
<td>DctD</td>
<td>A225</td>
<td>T</td>
<td>Fails to activate transcription in vivo/in vitro; reduced ATPase</td>
<td>214</td>
</tr>
<tr>
<td>DmpR</td>
<td>A315</td>
<td>T</td>
<td>Fails to activate transcription in vivo; WT ATPase activity</td>
<td>222</td>
</tr>
</tbody>
</table>

\* WT, wild type; NTD, N-terminal domain.
F85C, and F85Q variants retained the ability to hydrolyze ATP, which is explained by their ability to self-associate. However, they were unable to interact with σ54 to form “trapped” complexes in the presence of ADP.ΔΔF54, indicating that this residue is critical for bEBP-σ54 contact. Since some variants, e.g., F85Q, showed a significant decrease in ATPase activity, it was suggested that the F85 residue communicates with the ATP hydrolysis site. In contrast, the F85A, F85E, and F85R variants showed <10% of the ATPase activity of the wild-type protein, and gel filtration indicated that this was due to a defect in their ability to form higher-order oligomers. This suggests that there is a structural and functional link between the phenylalanine and the distant interface of self-association. The PspF F85Y variant was the only variant that was able to interact with σ54 to form the ADP.ΔΔF54 “trapped” complex. However, it too was unable to activate transcription, and this can be explained by the inability of the protein to form activator-σ54-DNA complexes using promoter DNA probes with a mismatch at the −11/−12 position (20, 234). Importantly, the G4L substitution in region I of σ54 (58) can rescue this σ54-DNA interaction defect (234). Therefore, the conserved phenylalanine of the GAFTGA motif plays a role in “sensing” the conformation of DNA at the −12 promoter position, in agreement with a model based on recent cryo-EM reconstructions (23). Overall, studies using different bEBPs suggest that the conserved phenylalanine has multiple, interrelated roles during transcriptional activation. The conserved threonine has been shown to contact region I of σ54 (see below), and the adjacent phenylalanine is also critical for this interaction. Previous studies suggested that F85 stabilizes L1-region I interactions indirectly through the positioning of T86 (21), and therefore, its major role is likely to be in contacting the promoter DNA rather than σ54.

The role of the conserved threonine in the GAFTGA motif is well understood (Table 2). Substitutions of the T218 residue of NtrC, the T308 residue of NifA, and the T223 residue of DctD abolish the ability of the bEBP to activate transcription (86, 131, 214). The same is true for the T85A and T85V substitutions in PspF, but T86S remains partially active (49). Strong evidence for a direct interaction between the threonine residue of the GAFTGA motif and region I of σ54 was provided by the identification of substitutions (e.g., G4L in region I) that specifically suppress the defects of the partially active T86S variant (20, 49, 58). Significantly, the G4L substitution allows the T86S variant to interact with but not activate the Eσ54 (G4L) complex when the promoter DNA is “premelted.” This supports a role for the GAFTGA motif in the “sensing” of the promoter DNA conformation downstream of the −10 position. It has been suggested that a communication of this information to σ54 via region I might allow Eσ54 to establish contact with single-stranded DNA, which is required for open complex formation (58).

The role of the second glycine of the GAFTGA motif has been less well studied. The G219K variant of NtrC showed only a 50% reduction in activity but failed to initiate open complex formation. In agreement with this, the equivalent substitution in NorR (G266K) was unable to activate transcription in vivo (36). Surprisingly, The NtrC G219K variant showed improved DNA binding properties (154). In contrast, the G219C and G266C variants of NtrC and NorR, respectively, were competent to activate transcription (36, 131), and for NtrC, increased ATPase activity was demonstrated, which is most likely due to increased oligomerization (131). However, as is the case for the A216C variant of NtrC, binding to enhancer DNA prevents the G219C variant of NtrC from activating transcription, and the ability of the variant to activate transcription is restored in a form of the bEBP defective for DNA binding (229). Taken together, the data for NtrC suggest that the binding of the C-terminal DNA binding domain to enhancers influences substrate remodeling by the GAFTGA motif. A more comprehensive mutational analysis of this conserved glycine residue was conducted with NorR, where most substitutions of the equivalent residue (G266) prevented the bEBP from activating transcription in vivo. However, the replacement of G266 with either C, S, Q, or M resulted in variants that were competent to activate transcription in response to the cognate signal (nitric oxide [NO]). Interestingly, the G266D and G266N substitutions gave rise to high levels of deregulated transcription in vivo. The discovery of such variants led to the proposal of a novel mechanism for the regulation of bEBP activity (see below) (36).

In common with the other residues of the GAFTGA motif, studies of NtrC, NifA, DctD, and DmpR revealed that the second alanine residue is critical for σ54-dependent transcription. The A220T and A220V variants of NtrC (131, 154); the A310N, A310D, and A310G variants of NifA (86); the A225T variant of DctD (214); and the A315T variant of DmpR (222) all fail to activate transcription. Only the A310S variant of NifA exhibits activity, although this is less than 20% of the activity of the wild-type protein (86). Many of the variants at this position are still able to hydrolyze ATP, and it is therefore likely that these substitutions destabilize the interaction between loop 1 and region I of σ54 that forms at the regulatory center in the closed complex.

**Coupling of ATP Hydrolysis to the Activation of Transcription**

ATP hydrolysis is coupled to open complex formation via conformational changes in the AAA" domain that ultimately lead to the relocation of bEBP-containing L1 and L2. A number of crystal structures of bEBP AAA" domains have been reported, providing detailed information about the nucleotide binding pocket and other key determinants (51, 128, 172, 181). In order to examine the structure of the AAA" domain at discrete stages of the nucleotide cycle, a variety of structural and biochemical techniques have been employed (22). These techniques include the soaking of transiently stable crystals of the PsPF AAA" domain (residues 1 to 275) in the presence of different nucleotides to obtain various nucleotide-bound structures (173). High-resolution crystal structures of NtrC1 bound to both ATP and ADP have also been reported (51, 128). Lower-resolution techniques, including cryo-EM and SAXS (small-angle X-ray scattering)/WAXS (wide-angle X-ray scattering), used in conjunction with nucleotide analogues have provided information about the larger, macromolecular conformational changes that occur in the bEBP as ATP is hydrolyzed (50, 52, 62, 172). However, caution should be taken in the analysis of all such structures, since the coordination of ATP hydrolysis between bEBP protomers is thought to involve heterogeneous nucleotide-bound states (117).

**The ground state of hydrolysis.** The ATP-bound “ground” state of the nucleotide cycle has been examined by soaking crystals of PsPF1−275 with ATP, either in the absence of Mg2+ to prevent hydrolysis or by using a hydrolysis-defective variant of the bEBP (22, 29, 173). The structures reveal the residues responsible for binding ATP within the nucleotide binding pocket (Fig. 8A). The
Walker B glutamate (E108 in PspF) senses the γ-phosphate of ATP and forms a strong interaction with a nearby highly conserved asparagine (N64 in PspF). The adjacent aspartate (D107 in PspF) coordinates the position of a water molecule for nucleophilic attack on this γ-phosphate (173). In NtrC1, the conserved R finger from the adjacent protomer (R299 in NtrC1 and R168 in PspF) appears to engage with the γ-phosphate, stabilizing the binding of ATP (51). Significantly, the ATP-bound structures of PspF and NtrC1 indicate that L1 and L2 are in a raised conformation, consistent with low-resolution SAXS-derived structures using the ground-state analogue ADP-BeFx in NtrC1 (50, 51, 173). Biochemical experiments confirmed that both PspF and NtrC1 can establish contact with \( \sigma^{54} \) in the presence of this ground-state analogue (22, 50). Taken together, these data indicate that the binding and sensing of the nucleotide cause significant conformational changes in the bEBP AAA\(^+\) domain. The combined evidence to date suggests that GAFTGA-containing L1, assisted by L2, is released to make an initial, unstable interaction with \( \sigma^{54} \) (Fig. 9).

The transition state of hydrolysis. Studies using the transition-

FIG 8 (A and B) Structure of monomeric PspF<sub>1–275</sub> bound to ATP (PDB accession number 2C96/2C9C) (A) and ADP (PDB accession number 2C98/2C9C) (B). Important motifs are highlighted: Walker A (brown), Walker B residue E108 (cyan), sensor I residue T148 (magenta), and “switch” residue N64 (yellow). Loop 1 (L1) and loop 2 (L2) are labeled, with the nonresolved fold of L1 indicated with a red dotted line. (C) Switching mechanism of the Walker B E108 residue. In the ATP-bound state, E108 interacts with N64 (indicated by a black dotted line). In the ADP-bound state, there is a 90° rotation in the N64 side chain so that E108 interacts with sensor I residue T148 via a water molecule (indicated by a red dotted line).
state analogue ADP.AlF₄ have established a clear role for the GAFTGA motif of loop 1 in contacting region I of H9268 prior to the remodeling of the closed complex (21, 39, 49). Most significantly, a stable interaction between GAFTGA-containing L1 and H9268 is formed in the presence of the transition-state analogue, as revealed by the fitting of the PspF₁₋₋₂₇₅ crystal structure into the cryo-EM reconstruction of the activator in complex with H9268 (172, 173). In agreement with this, low-resolution structures of NtrC and NtrC₁ bound to ADP.AlF₄ reveal an electron density above the plane of the oligomeric bEBP corresponding to a raised position of L1 and L2 (50, 62). Significantly, the ADP.AlF₄ complex of NtrC₁ was shown to be more stable than the ADP-BeF₄ complex, indicating that ATP hydrolysis strengthens the unstable interaction between the bEBP and H9268 that forms in the ground state (50).

Similar results have been obtained for PspF (33). In addition, this form of the activator is able to initiate the early stages of promoter “melting” (34). Moreover, in the transition state, when bound to ADP.AlF₄, PspF is competent to induce the Eo₄⁵₄ holoenzyme to initiate RNA synthesis on single-stranded promoter DNA, identifying the prehydrolysis state of bEBPs as being functionally important (32). However, the “trapped” forms of PspF are largely defective for nucleotide hydrolysis, do not promote DNA melting, and are incapable of activating transcription on double-stranded DNA (dsDNA) (32, 33, 49), suggesting that the continuation of the cycle of nucleotide-driven conformational changes is vital for the formation of the open complex.

**The postnucleotide hydrolysis state.** Crystal structures of various bEBPs in their ADP-bound forms have provided extensive information regarding the nature of the “posthydrolysis” state (128, 172, 173, 181). Comparisons of the ADP-bound forms with ground- and transition-state structures have shed light on the conformational rearrangements that occur upon ATP hydrolysis (Fig. 8 and 10). Structures of PspF indicate that the release of the γ-phosphate to form ADP causes a 90° rotation of the glutamate side chain of the Walker B motif (E108 in PspF). As a result, the interaction between the glutamate and the conserved asparagine (N64 in PspF) is broken, resulting in the Walker B glutamate interacting instead with the sensor I threonine residue (T148 in PspF) via a water molecule (Fig. 8C). Communication of the altered position of Walker B by the asparagine residue leads to conformational changes in helix 3 (H3) and helix 4 (H4) that are translated to loop 1 (L1) and loop 2 (L2) via strategically placed residues in the central domain. The functional significance of these key residues has now been assessed, and it has emerged that both intra- and intersubunit interactions have a role in modulating the conformation of the σ₄⁵₄ interaction surface (Fig. 10) (110).

The structure of ADP-bound PspF₁₋₋₂₇₅ suggests that the Walker B-asparagine “switch” results in the disruption of an interaction
between the R131 residue of L2 and the E97 residue of H3 (Fig. 10A) (173). Single substitutions at these positions abolish ATPase activity and interactions, while residue swaps (R131E/E97R) partially restore activity, demonstrating the importance of this polar interaction (110). Following this, the E97 residue was proposed to interact with R91, while the R131 residue is thought to contact L1 residue E81. Variants of R91 in PspF were unaltered for ATPase activity, and although they showed only a slight decrease in the ability to contact $\sigma^{33}$, they were significantly less able to form open promoter complexes, indicating that this residue is
important for substrate remodeling after the initial interaction (110). This arginine residue does not show a high level of conservation in the bEBP subfamily of AAA\(^+\) proteins, but in many cases, the adjacent residue may serve a similar function. Together, these new interactions result in the compaction of the loops down toward the surface of the AAA\(^+\) domain (Fig. 10), enabling \(\sigma^{54}\) relocation, which is crucial to the conversion from the closed to the open complex (22, 50, 173).

Interestingly, it has now been shown that inter-subunit (in trans) interactions confer cooperativity in the nucleotide-dependent substrate remodeling of PspF (Fig. 10B) (110). In the ATP-bound state, the E130 residue at the base of L2 was proposed to interact in trans with R98 of H5 from the adjacent subunit. Upon the release of the γ-phosphate, the E130 residue is expected to interact instead with R95 of H3, also from the adjacent protomer. Substitutions made at these positions cause the uncoupling of ATPase activity and substrate remodeling, since these variants are able to oligomerize and hydrolyze ATP but are not competent to contact \(\sigma^{54}\). Therefore, the “switching” of this in trans interaction between protomers is thought to contribute to the coordination of L1 and L2 during nucleotide hydrolysis (110). These residues do not show strict conservation in the bEBP subfamily, but similar interactions may play a significant role in the coupling of ATP hydrolysis to substrate remodeling. For example, in NtrC1, the equivalent residues (F226, L229, and Y261) are thought to facilitate trans hydrophobic contacts (51).

The Walker B-asparagine switch. Analyses of active-site structures reveal that the glutamate “switch” residue of the Walker B motif (E108 in PspF) is a common feature of hydrolysis in the majority of AAA\(^+\) proteins (237). Whereas substitutions of the adjacent Walker B aspartate (D107) have severe effects on various aspects of PspF activity, in line with a key role for this residue in ATP hydrolysis, a substitution of the adjacent glutamate (E108 in PspF) has only moderate effects on activity. Such variants have been used to effectively study intermediate states en route to open complex formation and confirm the pivotal role of the Walker B glutamate in substrate remodeling (110). This arginine residue does not show a high level of conservation in the bEBP subfamily, but similar interactions may play a significant role in the coupling of ATP hydrolysis to substrate remodeling. For example, in NtrC1, the equivalent residues (F226, L229, and Y261) are thought to facilitate trans hydrophobic contacts (51).

The Walker B-asparagine switch. Analyses of active-site structures reveal that the glutamate “switch” residue of the Walker B motif (E108 in PspF) is a common feature of hydrolysis in the majority of AAA\(^+\) proteins (237). Whereas substitutions of the adjacent Walker B aspartate (D107) have severe effects on various aspects of PspF activity, in line with a key role for this residue in ATP hydrolysis, a substitution of the adjacent glutamate (E108 in PspF) has only moderate effects on activity. Such variants have been used to effectively study intermediate states en route to open complex formation and confirm the pivotal role of the Walker B glutamate in transmitting nucleotide-dependent conformational changes (116). Sequence alignments of bEBPs indicate that the glutamate switch-interacting asparagine (N64 in PspF, labeled as “Switch Asn” in Fig. 6) is strictly conserved, in agreement with the hypothesis that this residue plays a key role in these specialized activators (112). In PspF, variants of the N64 residue have altered oligomeric states and ATPase activities, suggesting that the conserved asparagine plays a significant role in the organization of the active site. In line with this, the positions of Mg\(^{2+}\)-bound ATP and the water molecule involved in the nucleophilic attack of the γ-phosphate correspond with a catalytic role of the asparagine in ATP hydrolysis (173). Significantly, in the absence of the glutamate or asparagine side chains (i.e., in the E108A or N64A variant), PspF was shown to form a stable complex with \(\sigma^{54}\), indicating that the interaction between the activator and \(\sigma^{54}\) is not strictly dependent on the Walker B glutamate or switch asparagine residues (112, 116). However, the N64A variant was significantly less able to form open complexes than the wild-type activator. This defect was suppressed when premelted DNA was used, suggesting that the conserved asparagine is also required for the melting of the promoter DNA and the associated loading of the template into the RNAP active site (112). Despite the Walker B-interacting asparagine not being present in all members of the AAA\(^+\) superfamily, structural alignments of AAA\(^+\) proteins suggest that the distance between the glutamate and this residue is conserved. Consequently, it was suggested that the positioning of these two residues is important for their communication with each other and for forming a fully functional catalytic site. In the case of the specialized bEBP family, this communication helps to control the positioning of GAFTGA-containing L1. In AAA\(^+\) proteins that do not contain the GAFTGA insertion, it is likely that the interaction between these residues controls similar nucleotide-dependent conformational changes that regulate functionality in the oligomeric ring (112).

The arginine finger-directed “switch.” A recent report of a high-resolution crystal structure of the ATP-bound NtrC1 central domain (NtrC\(^{54}\)) and comparison with the ADP-bound form identified an alternative mechanism for the coupling of hydrolysis to substrate remodeling (31, 51, 128). In order to examine the configuration of an ATP-bound activator, the Walker B glutamate was replaced with alanine, allowing NtrC1 to bind but not to turn over the nucleotide. This led to the first structure of a bEBP in which the highly conserved arginine (R) finger (R299 in NtrC1) is seen to contact the γ-phosphate (Fig. 11). This is in contrast to the ATP-bound structure of monomeric PspF1–275, where the alternative R finger residue R162 (R293 in NtrC1), rather than R168 (R299 in NtrC1), is predicted to be in close proximity to the γ-phosphate of the adjacent protomer (173). A comparison of the ATP- and ADP-bound NtrC1 structures indicates that the engagement of the γ-phosphate by the R299 R finger stimulates a rearrangement of interaction networks in the same protomer (in cis), at the R finger side of the interprotomer interface. It has been proposed that the interaction of this R finger with the γ-phosphate causes helical distortions that ultimately lead to the transition of L1 and L2 to a raised conformation. The K250 residue appears to be particularly important in this transition; the side chain exists in distinct environments in the two nucleotide-bound states (51). This model is in contrast to prior studies of PspF, which indicated that the Walker B-asparagine “switch” is responsible for modulating the conformation of the \(\sigma^{54}\) interaction surface in the protomer that contains the Walker A and B residues of the ATP hydrolysis site (112, 116).

Model of nucleotide-driven conformational change. Overall, structural and biochemical studies have shown that the coupling of ATP hydrolysis to open complex formation by Eta\(^{54}\) is dependent upon conformational changes in the AAA\(^+\) domain (Fig. 9) (22, 51, 173). These changes center on the “sensing” of the γ-phosphate by either the asparagine “switch” residue that detects changes in the Walker B motif upon ATP hydrolysis (173) or the highly conserved R finger of the adjacent protomer (51). Depending upon the stage of the ATP hydrolysis cycle and the position of the “switch” or R finger, GAFTGA-containing L1 and L2 adopt various conformations (51, 110). Upon nucleotide binding, the loops are in an extended conformation, and the GAFTGA motif forms an unstable interaction with \(\sigma^{54}\). ATP hydrolysis strengthens this interaction, and a remodeling of the holoenzyme can then occur to enable open complex formation. Upon phosphate release, rearrangements of both in cis and in trans interactions cause the loops to disengage the σ factor. The cycle can then complete with the exchange of ADP for ATP. The GAFTGA motif thus performs a crucial role in the “power stroke” of bEBPs in the coupling of ATP hydrolysis to conformational rearrangements of the \(\sigma^{54}\)-RNA polymerase.
Coordination of ATP Hydrolysis

Asymmetry in the bEBP oligomer. Structural studies of the activator bound to different nucleotides have helped establish the conformational changes within the AAA⁺ domain that couple nucleotide hydrolysis to σ54 contact (173). The ADP.AIß-bound cryo-EM structure of PspF1–275 indicates that not all protomers within the AAA⁺ hexamer contact σ54 during the transition state of ATP hydrolysis, indicating asymmetry (172). In line with this, a number of heterohexameric AAA⁺ proteins exist (e.g., eukaryotic MCM2 to MCM7), comprising up to six different proteins, strongly suggesting that each subunit may have a distinct role in the activity of the hexamer (17,77). Studies of the homohexameric bEBP PspF have subsequently confirmed that an asymmetric configuration is a key requirement for open complex formation (111). The introduction of the GAFTGA substitution T86A into single-chain forms of PspF with two or three subunits allowed Joly and Buck (111) to examine the minimal requirements for σ54 contact and substrate remodeling. This substitution was shown to uncouple the ATPase and oligomerization activities of the bEBP from its ability to contact and remodel σ54. The minimal configuration for a stable interaction with σ54 was two adjacent functional subunits, revealing that more than one GAFTGA-containing L1 loops is required to mediate σ54 interactions and open complex formation in turn reveals that only a subset of ATP hydrolysis sites is required for bEBP activity (111). In support of this, work carried out to determine how ATP hydrolysis between protomers is coordinated suggested that ATPase activity in PspF is partially sequential (117). In the AAA⁺ family of proteins, two models of nucleotide occupancy may explain how hydrolysis is coordinated (Fig. 12)(2). Homogeneous nucleotide occupancy has been observed for a number of AAA⁺ protein crystal structures (79,129, 236), supporting a model of concerted/synchronized hydrolysis (Fig. 12D), in which subunits of the AAA⁺ ring simultaneously hydrolyze ATP. Other AAA⁺ structures show mixed nucleotide occupancy with ATP, ADP, or no nucleotide bound at the catalytic sites between subunits (18, 512).
This supports a model of sequential or rotational hydrolysis (Fig. 12B and C), in which heterogeneous nucleotide occupancy is coordinated between protomers. Studies using the bEBP PspF revealed that either ATP or ADP stimulates the oligomerization of the activator and that physiological ADP concentrations stimulate the ability of the protein to hydrolyze ATP. This suggests that in PspF, ADP binding promotes the formation of the stable hexamer and causes structural changes leading to increased ATPase activity in adjacent protomers. Furthermore, where a nonoptimal binding of nucleotides occurs, there are negative homotropic effects (184). High ATP concentrations at which every catalytic site in the hexamer is likely to be in the ATP-bound form inhibit ATP hydrolysis and the activation of transcription in vitro. Taken together, these data strongly suggest that heterogeneous nucleotide occupancy, coordinated between protomers in the hexameric ring, plays a crucial role in the activation of σ70-dependent transcription by bEBPs. Therefore, the concerted/synchronized model of hydrolysis can be discounted (Fig. 12D). Because of the “ATP inhibition” and “ADP stimulation” of PspF ATPase activity, nucleotide hydrolysis in bEBPs is unlikely to be a stochastic process in which each catalytic site is independent (Fig. 12A). Rather, the arrangement of the hydrolysis site and in vitro data suggest that cooperativity exists between protomers of the hexamer (117). Indeed, R-finger residues have been shown to function in trans, coordinating the bound nucleotide in the adjacent subunit of the hexamer (27, 51, 87, 235). Furthermore, recent mutagenesis studies of PspF have identified non-R-finger residues involved in interprotomer interactions that help to coordinate the positions of L1 and L2 during ATP hydrolysis (110). Subsequently, a model of sequential or rotational hydrolysis is favored (Fig. 12B and C). These mechanisms would create an asymmetry in the exposure of GAFTGA motifs in the hexamer, which has been shown to be important for the mechanical action of the activator (111, 117). However, the exact number of nucleotides bound and the conformation of individual protomers at discrete steps of hydrolysis still remain unclear. Determinations of high-resolution crystal structures of mixed nucleotide-bound hexamers would therefore be beneficial.

**Oligomerization State of bEBPs**

Since the nucleotide hydrolysis site is formed through interactions between residues of adjacent protomers, the oligomerization of the AAA+ domain is required to form a bEBP that is competent to...
Magnesium structure determined for an activator of meric state of bEBPs has become a matter of debate. The first NtrC4(RC, C, CD) bEBP Homologue of NtrC /H11001 RuvB AAA ZraRCD bEBP Zinc-responsive transcriptional p97 AAA hydrolyze the nucleotide) shows that it is competent to form a EM of the Walker B mutant derivative (that can bind but not firms a heptameric arrangement, and furthermore, negative-stain reported crystal structure of the ATP-bound form of NtrC1 con- height and a diameter of 40 Å and 124 Å, respectively. The recently (128). The 3.1-Å structure revealed a heptameric ring with a height and a diameter of 40 Å and 124 Å, respectively. The recently reported crystal structure of the ATP-bound form of NtrC1 con- firms a heptameric arrangement, and furthermore, negative-stain EM of the Walker B mutant derivative (that can bind but not hydrolyze the nucleotide) shows that it is competent to form a complex with $\sigma^{54}$ (31, 51). Additionally, low-resolution WAXS/ SAXS NtrC1 structures reveal a seven-membered ring (50, 51), indicating that the heptamer is not simply an artifact of crystallization. However, other bEBP structures have revealed hexameric arrangements (Table 3). When the crystal structure of the isolated ATPase domain of PspF ($\text{PspF}_{1-275}$) is fitted into the cryo-EM structure of the activator in complex with dsDNA, the electron density can accommodate six monomers (172). Indeed, electrospray ionization-mass spectrometry (ES-MS) shows that six monomers of $\text{PspF}_{1-275}$ form a complex with monomeric $\sigma^{54}$, consistent with bEBPs functioning as hexamers (133, 155). The 3-Å X-ray struc- ture of the zinc-responsive protein ZraR from $\text{Salmonella enteritica}$ serovar Typhimurium lacking the N-terminal regulatory domain also reveals a hexameric arrangement (181). These more recent studies have called into question whether the heptameric configuration of NtrC1 (128) represents a physiologically relevant form of the bEBP. In the absence of the C-terminal (DNA binding) and N-terminal (regulatory) domains, the stoichiometry of the full-length NtrC1 protein is unknown. In addition, the odd number of

<table>
<thead>
<tr>
<th>Protein</th>
<th>Class</th>
<th>Function</th>
<th>Organism(s)</th>
<th>Technique(s)*</th>
<th>Oligomeric structure(s)</th>
<th>Reference(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PsP&lt;sup&gt;54&lt;/sup&gt;</td>
<td>bEBP</td>
<td>Transcriptional activator that regulates phage shock response</td>
<td>E. coli</td>
<td>Electron microscopy</td>
<td>Hexamer</td>
<td>172</td>
</tr>
<tr>
<td>NtrCR&lt;sup&gt;RC&lt;/sup&gt;C&lt;sup&gt;C&lt;/sup&gt;</td>
<td>bEBP</td>
<td>Transcriptional activator that regulates nitrogen metabolism</td>
<td>Salmonella Typhimurium</td>
<td>SAXS/WAXS and electron microscopy</td>
<td>Heptamer (activated)</td>
<td>62</td>
</tr>
<tr>
<td>ZraR&lt;sup&gt;CD&lt;/sup&gt;</td>
<td>bEBP</td>
<td>Zinc-responsive transcriptional activator</td>
<td>Salmonella Typhimurium</td>
<td>X-ray crystallization</td>
<td>Hexamer</td>
<td>181</td>
</tr>
<tr>
<td>NtrC1&lt;sup&gt;RC&lt;/sup&gt;C&lt;sup&gt;C&lt;/sup&gt;</td>
<td>bEBP</td>
<td>Homologue of NtrC (Salmonella enteritica)</td>
<td>Aquifex aeolicus</td>
<td>X-ray crystallization</td>
<td>Heptamer (ADP bound and ATP bound)</td>
<td>51, 128</td>
</tr>
<tr>
<td>NtrC4&lt;sup&gt;RC&lt;/sup&gt;C&lt;sup&gt;C&lt;/sup&gt;C&lt;sup&gt;CD&lt;/sup&gt;</td>
<td>bEBP</td>
<td>Homologue of NtrC (Salmonella enteritica)</td>
<td>Aquifex aeolicus</td>
<td>ES-MS</td>
<td>Full-length and activated</td>
<td>12</td>
</tr>
<tr>
<td>RuvB</td>
<td>AAA&lt;sup&gt;+&lt;/sup&gt;</td>
<td>ATP-dependent motor for branch migration in homologous recombination</td>
<td>Thermus thermophilus</td>
<td>Electron microscopy</td>
<td>Hexamer in the presence of dsDNA; heptamer in the absence of dsDNA</td>
<td>142</td>
</tr>
<tr>
<td>MCM</td>
<td>AAA&lt;sup&gt;+&lt;/sup&gt;</td>
<td>Orthologue of eukaryotic replicative helicase candidate</td>
<td>Methanothermobacter thermautotrophicus</td>
<td>Electron microscopy</td>
<td>Hexamer in the presence of dsDNA; heptamer in the absence of dsDNA and in the presence of nucleotide</td>
<td>56, 57, 232</td>
</tr>
<tr>
<td>ClpB</td>
<td>AAA&lt;sup&gt;+&lt;/sup&gt;</td>
<td>Chaperone in protein-disaggregating machinery</td>
<td>E. coli</td>
<td>Sedimentation equilibrium/sedimentation velocity and electron microscopy</td>
<td>Hexamer predominant in the presence of ATPγS and ADP; heptamer predominant under low-ionic-strength conditions</td>
<td>3, 126</td>
</tr>
<tr>
<td>HslU</td>
<td>AAA&lt;sup&gt;+&lt;/sup&gt;</td>
<td>Part of HslUV two-component protease</td>
<td>E. coli</td>
<td>Electron microscopy; STEM of cross-linked protein; cross-linking/EMSA</td>
<td>Mixture of hexameric and heptameric rings</td>
<td>178</td>
</tr>
<tr>
<td>Lon protease</td>
<td>AAA&lt;sup&gt;+&lt;/sup&gt;</td>
<td>ATP-dependent protease</td>
<td>E. coli and Saccharomyces cerevisiae</td>
<td>Electron microscopy</td>
<td>Hexamer (E. coli); heptamer (Saccharomyces)</td>
<td>159, 198</td>
</tr>
<tr>
<td>Magnesium chelatase</td>
<td>AAA&lt;sup&gt;+&lt;/sup&gt;</td>
<td>Drives insertion of Mg&lt;sup&gt;2+&lt;/sup&gt; into protoporphyrin (chlorophyll biosynthesis)</td>
<td>Synechocystis (cyanobacterium) and Rhodobacter capsulatus (proteobacterium)</td>
<td>Electron microscopy</td>
<td>Hexamer (Rhodobacter); heptamer (Synechocystis)</td>
<td>174, 223</td>
</tr>
<tr>
<td>p97</td>
<td>AAA&lt;sup&gt;+&lt;/sup&gt;</td>
<td>Homotypic membrane fusion</td>
<td>Mammalian</td>
<td>X-ray crystallization, electron microscopy</td>
<td>Hexamer (full length); heptamer (C-terminal D2 only)</td>
<td>61, 236</td>
</tr>
</tbody>
</table>

* STEM, scanning transmission electron microscopy.
subunits does not match up with the dimeric arrangement of the receiver domains in both active and inactive forms (66, 128). To resolve the effect that different domains have on the oligomerization state of σ^{54}-dependent activators, a more thorough analysis was conducted by using different domain combinations of the NtrC4 protein from *Aquifex aeolicus* (12). As is the case for NtrC1, the ATPase activity of NtrC4 is subject to negative regulation. The assembly of the active oligomer is repressed by the receiver domain, and phosphorylation is likely to remove this repression (11). Unlike NtrC1, NtrC4 has a partially disrupted receiver-AAA\(^+\) domain interface and can assemble into active oligomers at high protein concentrations, independent of phosphorylation (11). ES-MS experiments showed that full-length NtrC4 (NtrC4\(^{RCD}\)) and activated NtrC4 lacking the DNA binding domain (NtrC4\(^{RC}\)) form hexamers. In contrast, the isolated ATPase domain (NtrC4\(^{α}\)), nonactivated NtrC4 lacking the DNA binding domain (NtrC4\(^{RC}\)), and NtrC4 lacking the regulatory domain (NtrC4\(^{C}\)) all form heptamers. A heptamer arrangement for the central ATPase domain in isolation is consistent with the heptamer observed when this domain of NtrC1 is crystallized (128). Therefore, it seems that for the extreme thermophile *Aquifex aeolicus*, a heptamer is the most stable arrangement for the AAA\(^+\) domain in the absence of regulatory and DNA binding domains. This is in contrast to the central domain of PspF, which forms hexamers when in isolation (172), and therefore, despite the high level of conservation of AAA\(^+\) domains, the PspF and NtrC1/NtrC4 central domains must have some differences. Interestingly, when the regulatory domain of NtrC4 is absent, a heptamer is formed, but when it is present and activated by phosphorylation, hexamerization occurs. Since the activated receiver domain stabilizes the hexameric form of NtrC4, it appears that an intermediate mechanism of regulation exists, somewhere between the negative mechanism of NtrC1/DctD and the positive mechanism of NtrC (11, 12). Overall, studies examining the oligomeric state of NtrC4 have conclusively shown that truncated or nonactivated proteins may have a propensity to exhibit altered stoichiometries. Although structural studies of ZraR suggested the possibility that the AAA\(^+\) domain is held in a hexameric configuration by the DNA binding domains, which are dimeric in nature (181), a truncated form of NtrC4 containing both the central domain and the C-terminal domain (CTD) is heptameric. Due to the difficulty in activating full-length NtrC1, the oligomeric state of this construct cannot be assessed, but the similarity of NtrC4 to NtrC1 suggests that the activated form of NtrC1 is likely to be hexameric. In addition to bEBPs, heptamer arrangements have also been observed frequently for other members of the AAA\(^+\) family (Table 3), such as MCM (56, 57, 232), RuvB (142), ClpB (3, 126), magnesium chelatase (174), HslU (178), Lon (198), and the C-terminal domain of p97 (61). Significantly, hexamers have also been observed for each of these proteins (56, 57, 142, 159, 232, 236). Various explanations exist for the existence of two different isoforms of the same AAA\(^+\) protein. In the case of RuvB from *Thermus thermophilus* and the MCM protein from *Methanothermobacter thermautotrophicus*, heptamers form in the absence of DNA, but hexamerization occurs when DNA is present (56, 57, 142, 232). It is possible that the heptamer facilitates the loading of DNA into the central channel of the protein ring before the loss of a single subunit results in the hexamer (232). The bacterial protein-disaggregating chaperone ClpB forms heptamers in the absence of a nucleotide but undergoes rearrangements to form hexamers when ATP or ADP binds (3, 126). This implies that during the ATP hydrolysis cycle (as ATP binds and is hydrolyzed and ADP is released), there is “switching” between hexameric and heptameric states. This partial ring dissociation has been suggested to facilitate the “prying apart” of aggregated substrates (3). In the case of the ATPase HslU, rings of 7-fold and 6-fold symmetry have apparently been observed under the same conditions (178). Here, it is unclear as to whether the heptameric form of the protein is competent to associate with the partner protease, HslV. However, it was suggested that the symmetry mismatch between a heptameric HslU and a hexameric HslV may facilitate the loading of the substrate into the proteolytic chamber, as was suggested for ClpAP (125). Unlike the ATP-dependent proteases HslUV and ClpAP, the Lon protease combines both proteolytic and ATPase functions within a single subunit. The Lon protease from *Saccharomyces cerevisiae* has been reported to consist of seven flexible subunits (198), possibly reflecting the requirement for a “mis-match symmetry” of the HslUV and ClpAP systems. However, electron microscopy of the Lon protease from *Escherichia coli* indicates a hexameric arrangement (159). Likewise, the magnesium chelatase subunit BchI from the proteobacterium *Rhodobacter capsulatus* has 6-fold symmetry (223), but the equivalent subunit in the cyanobacterium *Synechocystis* has 7-fold symmetry (174). This suggests that it may also be possible for the same AAA^+ protein to have different oligomeric states in different organisms or evolutionary groups. Finally, as described above for the bEBP NtrC4 (12), proteins may be heptameric when expressed as isolated AAA^+ domains but hexameric in their intact forms. For example, when crystallized, the isolated C-terminal (D2) domain of p97 reveals a 7-fold symmetry (61), but cryo-EM studies of the full-length form indicate that, in fact, the protein has a hexameric arrangement (236).

### ROLE OF THE N-TERMINAL REGULATORY DOMAIN

#### Signal Sensing

Many bEBPs contain an N-terminal or regulatory (R) domain that responds to various environmental signals and regulates the activity of the central AAA^+ domain as a result (186). There are three main ways in which the regulation of the bEBP is achieved by the regulatory domain in response to an environmental cue: (i) phosphorylation, (ii) ligand binding, and (iii) protein-protein interactions (Table 1). Depending on the method of bEBP activation, different domains are found in the regulatory region of the protein (examples of which are shown in Fig. 5) (200). The signals perceived by many of these are now well understood. However, there are a large number of different bEBP architectures in the Pfam database (http://pfam.sanger.ac.uk/), and many contain other types of N-terminal domains (including Prp_N, RtcR, FHA, CBS, PILZ, TIM-barrel, DapB_N, PocR, Fer, NAD binding, CHASE, HDOD, ICL, and LMWPC) whose functions in signal perception and AAA^+ domain regulation are currently unknown.

#### Phosphorylation

Many bEBPs are part of two-component systems (TCs) that couple an external stimulus to an internal response (199). Such systems are commonly composed of a histidine protein kinase (HK) with a conserved kinase core domain and a response regulator (RR) protein with a conserved regulatory domain. Extracellular stimuli are sensed by the HK to modulate its activity in phosphotransfer. The HK transfers a phosphoryl group to a conserved aspartate in the RR (a reaction catalyzed by the RR),...
and the phosphorylated RR is able to activate a downstream effector domain that elicits a specific response in the bacterial cell. For example, the bEBPs NtrC, NtrC1, NtrC4, DctD, ZraR, and FlgR all have RR domains (Fig. 5) that are phosphorylated by specific sensor kinases. The best-studied RR in this group is the bEBP NtrC, which is phosphorylated at the conserved D54 residue by the sensor kinase NtrB in response to the nitrogen status of the cell (175). Briefly, the phosphorylation cascade is controlled by the uridyltransferase (GlnD), which transmits the nitrogen status to the NtrB protein via the PI1 protein GlnB. Under nitrogen-limiting conditions, NtrB phosphorlates NtrC, activating it as a bEBP. Under nitrogen-excess conditions, the phosphatase activity of NtrB prevents NtrC activation (reviewed in reference 65). The NtrC1 and NtrC4 bEBPs from the thermophile A. aeolicus, which have been studied extensively at the structural level, are both classed as NtrC family members based on their high amino acid similarity (59%), but their cognate HKs and the signals control-classed as NtrC family members based on their high amino acid similarities (158). The ZraR RR is another example of a group I bEBP, which is phosphorylated by the HK ZraS in response to high Zr2+ concentrations. The activation of the ZraR effector domain results in the expression of a periplasmic Zn2+ binding protein, Zrap (130). In contrast, the FlgR RR from Helicobacter pylori is a group V bEBP (Fig. 5), which lacks a DNA binding domain (discussed below). This activator is phosphorylated by the sensor HK FlgS, leading to the transcription of genes required for flagellar biosynthesis (26, 197).

**Ligand binding.** Other σ54 activators have a regulatory domain that binds small effector molecules (Fig. 5). The direct binding of aromatic compounds to a vinyl 4 reductase (V4R) domain activates the group II bEBPs DmpR and XylR (161, 189, 194). In response to small-ligand binding, DmpR activates the expression of the dmp operon, which encodes enzymes involved in the catabolism of phenol and methyphenols (191–193). XylR binds to tol- uene, m-xylene, and p-xylene to activate transcription at the Pu promoter of the TOL plasmid, allowing Pseudomonas putida to grow on toluene and related hydrocarbons (64). N-terminal-domain swaps between the XylR and DmpR proteins confirm that the specificity of the response is conferred by the regulatory domain (193). Another regulatory domain frequently found in bEBPs is the GAF (cyclic GMP [cGMP]-specific and -stimulated phosphodiesterases, Anabaena adenylate cyclases, and E. coli FhlA) domain (Fig. 5), a member of a large and diverse domain family that is found in all kingdoms of life (5). FhlA contains two GAF domains that bind formate to activate the transcription of the formate hydrogen lyase system (100). NorR contains a single GAF domain, which binds NO to activate the transcription of the norV and norW genes, enabling NO detoxification (60, 106). The Per, ARNT, and Sim (PAS) domain (168) often detects signals via a bound cofactor such as heme or flavin (204) and, like the GAF domain, is present in many bEBPs. Although the PAS and GAF domains share little sequence similarity, they have similar structures and may share an ancestor (97, 168). Finally, the aspartokinase, chorismate mutase, and TyrA (ACT) domain is common in metabolic enzymes that are regulated by amino acid concentrations. The bEBP-like protein TyrR contains both a PAS domain and an ACT domain (Fig. 5) and facilitates the activation or repression of the transcription of genes involved in aromatic amino acid biosynthesis and transport, although it is not an activator of σ54-dependent transcription (164, 165, 209). The ACT domain is likely the binding site for the aromatic amino acid tyrosine, phenylalanine, or tryptophan, whereas the PAS domain has been suggested to have a role in contacting the αCTD of RNAP (165).

**Protein-protein interactions.** A further group of bEBPs regulates the activity of the AAA+-domain through protein-protein interactions with another protein called an antiactivator. In the nitrogen-fixing organism Azotobacter vinelandii, the bEBP NifA is bound by the antiactivator NifL to prevent the transcription of nif genes under conditions that are inappropriate for nitrogen fixation (136, 137, 145). In addition, the NifA protein itself contains a regulatory GAF domain. The binding of 2-oxoglutarate (2-OG) to the GAF domain antagonizes the influence of adenine nucleotides on the NifL/NifA interaction to ensure that the bEBP is not inhibited by NifL under nitrogen-fixing conditions (132, 135).

Some bEBPs, such as PspF, lack an N-terminal regulatory domain altogether (200). PspF is instead negatively regulated by PspA in trans (70, 73, 114). Initially, PspA was suggested to be an escaped regulatory domain of PspF, but phylogenetic analyses have placed PspF into a distinct clade of response regulator bEBPs (200). The pspABCDE operon encodes several proteins that help maintain membrane integrity. It has been suggested that upon proton motive force (PMF) dissipation, PspB and PspC act as positive regulators of transcription by binding PspA and relieving the inhibition of PspF. This enables PspF to activate σ54-dependent transcription from the psp promoter (1, 59, 114, 143). Like PspF, the related HrpR and HrpS activators are group IV bEBPs that lack an N-terminal regulatory domain (Fig. 5). The HrpR/ HrpS system regulates transcription from the hrp (hypersensitive response and pathogenicity) gene cluster (105, 188), which encodes plant pathogenicity genes, including components and effectors of a type III secretion pathway in Pseudomonas syringae (4, 54). The extracytoplasmic function (ECF) σ factor (141) HrpL is the primary transcription factor that controls the expression of the hrp gene cluster. HrpR and HrpS have been shown to interact, forming a stable heteromeric complex that activates the σ54-dependent transcription of hrpR (105). In a manner analogous to that of the regulation of PspF, HrpS (but not HrpR) is specifically bound by another protein, HrpV, to repress the activity of the heterohexamer (122). However, there is no apparent homology between HrpV and PspA (170, 200).

**Global regulatory signals.** In addition to the regulation of transcription via the signal-sensing functions of bEBP regulatory domains, the output from σ54-dependent promoters can additionally be controlled in response to global regulatory signals (190). This regulation could be mediated by factors that counteract the binding either of the holoenzyme at the promoter or of the activator (exclusion). In Pseudomonas sp. strain ADP, the bEBP NtrC has been shown to activate the σ54-dependent transcription of the atzR gene from solution (169). The AtzR protein acts as a regulator of cyanoic acid metabolism but also autoregulates its own expression by binding to a site that overlaps the atzR promoter to inhibit the formation of the closed complex by the σ54-RNA polymerase (169). In Klebsiella aerogenes, NtrC activates transcription from the σ54-dependent nac promoter under nitrogen-limiting conditions. In this case, Nac negatively autoregulates its own ex-
pression by a mechanism of antiactivation; it binds within the intergenic region (between the upstream NtrC enhancer sites and the promoter) to prevent productive interactions between the activator and the RNA polymerase holoenzyme. Both exclusion and antiactivation mechanisms seem to be utilized in the regulation of the activation of transcription by DctD at the dctA-dependent dctA promoter. Cyclic AMP (cAMP)-bound CRP (cAMP receptor protein) is able to bind to two sites that overlap the binding sites of the bEBP, but it is also able to interact directly with the holoenzyme to inhibit transcription. In vitro data suggest that the cAMP-CRP interaction with the holoenzyme results in an alternative closed complex that slowly converts into one that is subject to bEBP activation.

Controlling the Activity of the Central AAA⁺ Domain

Negative regulation as a dominant mechanism of control. As discussed above, the N-terminal regulatory (R) domain allows the bEBP to regulate transcription at σ54-dependent promoters in response to environmental cues. However, bEBPs have developed different methods for the transduction of this signal from the R domain (the site of detection) to the catalytic C domain. Generally, this transduction can be subject to (i) positive control or (ii) negative control. Assaying the activity of the bEBP (i.e., ATP hydrolysis, oligomerization, or activation of transcription) in a truncated form that lacks the N-terminal R domain has become the standard method for determining the mechanism of control. The first indication that the R domains of bEBPs may function in the repression of AAA⁺ domain activity came from the identification of semiconstitutive variants of XylR and DmpR that had substitutions in either the R domain, the C domain, or the interdomain linker (often referred to as either the B linker, Q linker, or L₁ linker). Furthermore, N-terminally truncated forms of these proteins that lacked the R domain exhibited constitutively active phenotypes in vivo, indicating that the C domain is subject to repression by the R domain. For DmpR, constitutive ATPase activity in vitro was demonstrated in the absence of this negative control. These observations led to a model of interdomain repression in which the R domain represses the ATPase activity of the C domain in the absence of a signaling molecule. Ligand binding to the R domain is then expected to cause derepression, allowing the bEBP to hydrolyze ATP and activate transcription (Fig. 13A). More recently, this mechanism of negative control has been identified in the response regulator bEBPs DctD and NtrC1. Here, phosphorylation and not effector binding relieves interdomain repression (66, 128). The removal of

FIG 13 Negative (A) and positive (B) control of AAA⁺ domain activity. In the more common mechanism of negative control, ligand binding (or phosphorylation) relieves the repression of the regulatory (R) domain on the central (C) domain, which is intrinsically competent to hydrolyze ATP. The AAA⁺ domain is then able to carry out ATP hydrolysis. Accordingly, when the R domain is removed, the bEBP is active irrespective of the presence or absence of a signaling molecule (shown as a purple triangle) or an available kinase. In positive control, ligand binding or phosphorylation has a genuine stimulatory function. The phosphorylated or ligand-bound form of the R domain activates the C domain, which is not intrinsically competent to hydrolyze ATP. The AAA⁺ domain is then able to carry out ATP hydrolysis. Accordingly, when the R domain is removed, the bEBP is inactive irrespective of the presence or absence of a signaling molecule or an available kinase.
the R domain and the L₁ linker in DctD produced an active protein without the need for phosphorylation (88). Similar results were obtained with NtrC₁, where activity was repressed both in vivo and in vitro in the presence of the R domain and the L₁ linker but was derepressed in their absence (128). The founder member of the σ^{54}-dependent class of transcription factors, NtrC, in contrast, is positively regulated (Fig. 13B). The deletion of the R domain to give a form of the activator that can no longer be phosphorylated by NtrB results in a constitutively inactive form of the protein, indicating that the AAA⁺ domain is subject to positive regulation (69, 216). Here, the phosphorylation of the R domain has a genuine stimulatory, rather than a derepressing, function. Therefore, despite sharing ~60% sequence similarity, the NtrC and NtrC₁ bEBPs have evolved entirely different mechanisms of regulation.

In the absence of a transcriptional assay, it is not known whether ZrrA is subject to positive or negative regulation, but on the basis of structural similarities, it is likely to belong to the negatively regulated NtrC₁/DctD group (181). Overall, the relative advantages of protein-protein interactions, phosphorylation, or effector binding as a control mechanism are not understood, but it seems that whatever the mechanism of sensing, negative regulation is the dominant mechanism of control (189).

**Functions of the C domain targeted by the R domain.** In order for the output of the bEBP to be regulated, the sensory domain must respond to the detection of an environmental or metabolic signal by controlling the activity of the AAA⁺ domain that is indispensable and often sufficient for σ^{54}-dependent transcription (15, 121, 222, 226). Irrespective of whether the R domain regulates the C domain positively or negatively, it has been shown to target three different aspects of AAA⁺ activity: (i) the oligomerization of the AAA⁺ domain, (ii) the ATPase activity of the AAA⁺ domain, and (iii) the interaction with σ^{54} (Fig. 14). These targets will now be considered.

(i) Controlling AAA⁺ oligomerization. As noted above, a self-association of the AAA⁺ domains of the bEBP must occur in order to form the functional activator (171, 235). Therefore, the oligomeric determinants of the C domain represent an ideal target for the N-terminal regulatory domain for either a positive or negative mechanism of control (Fig. 15). Structural studies of full-length and truncated forms of NtrC₁ and DctD from *Aquifex aeolicus* and *Sinorhizobium meliloti*, respectively, indicated that the N-terminal R domain targets the oligomeric determinants of the AAA⁺ C domain in the mechanism of negative control (Fig. 15B) (52, 66, 128, 140, 158). The crystal structure of a form of NtrC₁ comprising the regulatory domain joined to the central AAA⁺ domain by linker L₁ (R-L₁-C) reveals a dimeric structure in which the arrangement of the subunits is incompatible with AAA⁺ ring assembly (128). Here, the unphosphorylated receiver domains form a homodimer that holds the AAA⁺ protomers in an inactive front-to-front configuration via interactions involving the coiled-coil structure of linker L₁. Structures of the activated regulatory domain indicate that phosphorylation results in an alternative homodimer configuration that disrupts the repressive interaction between the R and C domains, allowing the reorientation of the AAA⁺ protomers into a front-to-back configuration (66). This phosphorylation-dependent rearrangement allows self-association to take place to form an oligomer competent to hydrolyze ATP. In line with this, the crystal structure of the isolated AAA⁺ domain of NtrC₁ reveals a heptameric arrangement in the absence of the R domain (128). It appears that the coiled coil of the L₁ linker is critical for holding the central domain in an inhibitory configuration; its presence has become indicative of this type of regulation (66). Indeed, for the ligand binding XylR and DmpR proteins, mutational analysis has shown that the integrity of the linker between the regulatory and central domains is crucial for the repression of activity (83, 156). The mechanism of regulation in NtrC also targets the oligomeric determinants of the AAA⁺ domain but is in stark contrast to those in NtrC₁ and DctD (Fig. 15A) (52, 62, 66). A truncated form of the protein that lacks the R domain is constitutively inactive, indicating a genuine stimulatory rather than a derepressive role for phosphorylation. The activation of oligomerization occurs upon phosphorylation, which exposes a hydrophobic patch on the R domain allowing it to bind to the N-terminal region of the central AAA⁺ domain. Recent X-ray solution scattering (SAXS/WAXS) and electron microscopy studies indicated that the R domain interacts with the C domain of an adjacent protomer on the outside edge of the AAA⁺ ring, promoting self-association and contributing to the stability of the resulting hexamer (62). Put simply, in the positive regulation of NtrC, the phosphorylation of the regulatory domain creates a new inter-

**FIG 14** Possible targets of regulatory domain-mediated regulation. (A) Most commonly, the regulatory domain represses oligomerization, e.g., NtrC₁ and DctD (66, 128), or promotes self-association in response to the signal, e.g., NtrC (62). (B) In PspF, negative regulation directly targets the nucleotide hydrolysis machinery (115). While the binding of ATP releases L₁ and L₂ to establish a weak interaction with σ^{54}, hydrolysis is required to produce a strong interaction that results in the remodeling of the holoenzyme (173). (C) NorR may represent a newly identified mechanism of control in which the interaction with σ^{54} is the target of the regulatory domain. Where oligomerization is not the target, the preassembly of a hexamer prior to activation may have a physiological advantage, e.g., a rapid response to stress. The mechanisms of regulation that target ATP hydrolysis, oligomerization, and σ^{54} interactions are likely to be highly interconnected, with the enzymatic activity of the AAA⁺ domain being the ultimate target of regulation.
action that leads to oligomerization, whereas in negative regulation, typified by NtrC1 and DctD, phosphorylation releases an interaction that leads to the formation of the functional oligomer (Fig. 15). Interestingly, sequence analysis reveals a correlation between the mechanism of negative control in NtrC1 and DctD and the presence of a structured linker between the R and C domains (linker L1) and an unstructured linker between the C and D domains (linker L2). Conversely, the positively regulated NtrC contains an unstructured L1 linker and a structured L2 linker. In both classes, the structured linker seems to play a significant role in the stabilization of the inactive dimer, and examinations of the L1 and L2 sequences of other bEBPs may help identify whether self-association is subject to positive or negative control (66). For example, a mutation of the linker between the R and C domains in NtrC does not affect its activity (225), in agreement with a function for the regulatory domain in positive rather than negative control.

The bEBP NtrC4 from *Aquifex aeolicus* has a partially disrupted receiver-AAA domain interface and can assemble into active oligomers at high protein concentrations independent of phosphorylation, a process that does not occur with NtrC1 (11). The activated receiver domain has been shown to stabilize the hexameric form of NtrC4, thus functioning as an intermediate between the negative mechanism of NtrC1/DctD and the positive mechanism of NtrC (11, 12).

(ii) Controlling ATPase activity. Ultimately, the target of R-domain-mediated regulation is the enzymatic activity of the bEBP. Where the R domain targets the oligomeric determinants, the effect is to promote or prevent the formation of an oligomer that is capable of hydrolyzing ATP. However, the regulatory domains of some bEBPs may specifically target the nucleotide hydrolysis machinery without influencing the oligomerization state (Fig. 14B). This has been shown for PspF, which is regulated in trans through direct interactions between the activator and the negative regulator PspA (Fig. 16) (70, 72, 73, 114). Here, oligomerization is driven by the binding of ADP and ATP to the individual protomers (117), whereby the “DE” residues of the Walker motif prevent nucleotide-independent hexamer formation (116). PspA has been shown to negatively regulate the ATPase activity of PspF through the formation of an interaction that is dependent on a surface-exposed tryptophan residue (W56 of

---

**FIG 15** Models of bEBP activation by phosphorylation through the promotion of oligomerization by stimulatory (A) and derepressing (B) functions of the response regulator (RR) domain. In activated NtC, the DNA binding domain is hidden underneath the hexamer ring. For DctD and NtrC1, no information is available to define the positions of DNA binding domains. R, regulatory domain; L1, linker 1; C, central domain; L2, linker 2; D, DNA binding domain. Models were built by using published structures of NtrC fragments R (off state, PDB accession number 1KRW; on state, accession number 1KRX) and L2-D (PDB accession number 1NTC) and NtrC1 fragments R (PDB accession number 1ZY2), R-L1-C (PDB accession number 1NY5), and L1-C (PDB accession number 1NY6). (Adapted from reference 66 with permission from Elsevier.)
Recently, it was shown that the PspA-mediated inhibition of PspF ATPase activity is likely to involve the repositioning of the conserved asparagine (N64 in PspF) required for the sensing of the γ-phosphate during nucleotide hydrolysis. Substitutions of this asparagine in PspF do not prevent PspA binding, but ATPase activity is not significantly decreased by the presence of PspA as in the wild-type activator (112). Consequently, a model has been proposed to link the binding of PspA to the inhibition of ATP hydrolysis in PspF. Upon the dissipation of the PMF, PspA inhibition is prevented (possibly facilitated by PspB and PspC), and ATP hydrolysis can occur, strengthening the σ54 interaction and leading to substrate remodeling. The removal of the γ-phosphate leads to a 90° rotation of the E108 side chain, breaking the interaction with N64. This change is translated to GAFTGA-containing L1 (orange) via helix 3 (H3) (yellow), and the loops compact back downwards. (Adapted from reference 112 with permission of the publisher.)

(iii) Controlling σ54 interactions. For a BEP to activate transcription, it must ultimately make contact with the Eσ54 holoenzyme bound at the promoter. As described above, this interaction is facilitated by the highly conserved GAFTGA motif within the PspF (72, 73, 114). Recently, it was shown that the PspA-mediated inhibition of PspF ATPase activity is likely to involve the repositioning of the conserved asparagine (N64 in PspF) required for the sensing of the γ-phosphate during nucleotide hydrolysis. Substitutions of this asparagine in PspF do not prevent PspA binding, but ATPase activity is not significantly decreased by the presence of PspA as in the wild-type activator (112). Consequently, a model has been proposed to link the binding of PspA to the inhibition of ATP hydrolysis in PspF. The binding of PspA is detected via the W56 residue, which relays this information to N64 via β-sheet 2. This leads to the repositioning of the N64 side chain, altering the distances between ATP, the conserved asparagine, and the Walker B glutamate (E108 in PspF) (112). These distances are thought to be critical for ATP hydrolysis and the coordination of the resulting conformational changes in the AAA+ domain. Significantly, it has been demonstrated that the inactive regulatory complex consists of approximately six PspA subunits and six PspF subunits (113). Therefore, in contrast to the bEBPs NtrC1 and DctD, the negative regulation of PspF activity is unlikely to target the oligomeric determinants (66, 128). In addition, PspA does not inhibit the interaction between PspF and σ54, suggesting that negative regulation does not target the σ54 interaction surface of PspF. The PspA-PspF regulatory complex is instead expected to have an altered arrangement in the key ATP hydrolysis determinants that form the catalytic site at the interprotomer interfaces of the PspF hexamer. Potentially, the inhibition of a preassembled PspF hexamer by PspA allows the cell to rapidly respond to membrane damage (113, 114).
central AAA\(^+\) domain of the hexamer, in response to the binding and hydrolysis of ATP (21). Therefore, it is feasible that in some activators, the targets of the regulatory domain could be the determinants of \(\sigma^{54}\) interactions (Fig. 14C). Upon the receipt of a signal, the R domain could either promote the formation of new contacts to form the interaction surface (positive control) or remove inhibitory contacts that enable \(\sigma^{52}\) contact (negative control). Previous work identified the nitric oxide (NO)-responsive transcription factor NorR as a member of the bEBP subfamily of AAA\(^+\) proteins, which are subject to negative regulation. When the N-terminal regulatory (GAF) domain was removed, the truncated form of NorR was shown to be constitutively active (81, 167), suggesting a role for the R domain in interdomain repression. Electron paramagnetic resonance (EPR) spectroscopy has since shown that NO binds to the nonheme iron center of the N-terminal GAF domain of NorR to form a mononitrosyl \{Fe(NO)\}\(^2\) \((S = 3/2)\) species, triggering conformational changes that relieve the repression exerted by the regulatory domain upon the AAA\(^+\) domain (60). NorR is then able to hydrolyze ATP, leading to open complex formation and the expression of the nor\(VW\) genes, the products of which enable \(E.\ coli\) to reduce NO to the less toxic \(N_2O\) (82, 85, 106). More recently, a number of AAA\(^+\) variants of NorR that escape GAF-mediated repression have been identified and are located in a key region of the central domain that undergoes significant conformational changes as ATP is hydrolyzed (36). Significantly, two bypass mutations were identified in the GAFTGA motif, which is absolutely required for \(\sigma^{54}\) dependent transcription. The G266D and G266N variants (GAFTGA) enabled complete escape from the repression exerted by the regulatory domain. This invokes a model whereby the GAF domain enabled complete escape from the repression exerted by the regulatory domain, which is absolutely required for the AAA\(^+\) domain, including the key GAFTGA motif (Fig. 17). This mode of regulation of NorR, as in the case of PspF, may reflect the requirement for a rapid stress response, facilitated by the preformation of the bEBP hexamer. In the case of those “rare” bEBPs that contain a naturally occurring aspartate or asparagine residue at the second glycine of this motif, it is anticipated that the AAA\(^+\) domain of such proteins is still subject to regulation by the N-terminal domain. Indeed, the FlgR protein of \(H.\ pylori\) contains a “GAFTDA” motif but still requires phosphorylation by FlgS in order to activate transcription (26). Therefore, these bEBPs are not likely to be regulated by targeting the \(\sigma^{54}\) interaction surface as in the case of NorR; instead, control may be at the point of oligomerization or ATP hydrolysis.

**ROLE OF THE C-TERMINAL DOMAIN**

The C-terminal or DNA binding (D) domain contains a helix-turn-helix (HTH) motif that directs the binding of the bEBP to enhancer sites typically 80 to 150 bp upstream of the promoter (200, 227). The D domain has the capacity to perform four different roles: (i) directing the binding of the activator to DNA targets to ensure a specific response, (ii) facilitating the formation of oligomers, (iii) stabilizing the hexamer, and (iv) maintaining the regulatory fidelity of the R domain.

**Directing DNA Binding**

The specificity of the DNA interaction is maintained by well-conserved enhancer binding sites (upstream activator sequences [UAS\(S\)]), mediated by the second (recognition) helix of the HTH motif (55). All sites exhibit dyad symmetry, and it is therefore unsurprising that the majority of bEBPs bind to DNA as dimers. This is supported by the crystal structure of ZraR and the NMR structure of NtrC, which reveal that the dimerization determinants of their D domains are located in an \(\alpha\)-helix, similar to that found in the FIS (factor for inversion stimulation) protein (160, 181). All bEBPs bind to at least one enhancer site, and as many as three enhancer sites have been identified upstream of the target promoter. FleQ, a regulator of flagellar biosynthesis from \(Pseudomonas\ aeruginosa\), is an atypical bEBP in that DNA binding can occur either upstream or downstream depending on the target promoter. When binding upstream from a distance, FleQ activates transcription as a typical bEBP via DNA looping, but when binding downstream in the vicinity of the promoter, FleQ may activate transcription via a novel mechanism (123).

**Facilitating Oligomerization**

Oligomerization has been shown to be DNA dependent in the bEBPs XylR (162), NtrC (8, 179), and NorR (207). Where more than one UAS site is present, the binding of multiple bEBP dimers to enhancer DNA may lead to an increase in the local concentration of the activator, thereby facilitating oligomerization. Enhancer DNA is clearly important for the activation of NorR as a transcription factor (207). In the absence of the regulatory domain, the requirement for the NO signal is bypassed, but DNA containing the three NorR binding sites is still required for NorR to hydrolyze ATP and activate transcription. Furthermore, each of the three sites has been shown to be essential for NorR activity both in vivo and in vitro (207). While multiple enhancer sites for bEBPs are not uncommon, an absolute dependency on more than one binding site is unusual. In NorR, this may reflect the requirement of DNA to act as a scaffold to facilitate oligomerization prior to the receipt of the NO signal. In contrast, NtrC dimers bind to two enhancer sites and recruit a third dimer from solution to form...
the functional hexamer upon the phosphorylation of the R domain (62). Since the release of the repression mechanism exerted by the regulatory domain does not stimulate self-association in NorR, DNA binding has instead evolved to drive the process of oligomerization. Interestingly, the number of enhancer sites is not strictly conserved between different norR-containing proteobacteria. For example, in the aerobic, soil-dwelling organism Azotobacter vinelandii, only two predicted NorR enhancer sites exist upstream of the hmp gene. It would be beneficial to study the control of such promoters in these bacteria in order to provide insight into the conservation of the NorR-mediated regulation of gene transcription.

At high concentrations, some bEBPs have been shown to activate transcription without binding to enhancer DNA. Indeed C-terminally truncated forms of the activators PspF, NtrC, NifA, and DctD have been shown to be active in vivo and in vitro (14, 102, 103, 121, 147, 153). Intriguingly, some bEBPs, such as Chlamydia trachomatis CtcC and Helicobacter pylori FlgR, naturally lack the C-terminal DNA binding domain that is present in most other bEBPs (13) (Fig. 5). An N-terminally truncated version of FlgR (lacking the regulatory receiver domain) is competent to activate σ54-dependent transcription from a promoter that naturally contains no upstream or downstream enhancer sites (26). FlgR and CtcC (127) are the only activators of σ54-dependent transcription in H. pylori and C. trachomatis, respectively, negating the need for enhancer binding in these organisms. The energy savings gained by the use of such activators are likely to be small given the regulatory potential of the use of multiple DNA binding bEBPs. Consistent with this, C. trachomatis and H. pylori have only limited biosynthetic capabilities (26). In contrast, FleT also lacks the C-terminal domain but is not the sole activator of σ54-dependent transcription in Rhodobacter sphaeroides. Here, specificity is achieved through multiple σ54 paralogues that function at different sets of promoters (13,166). For example, one of the paralogues, RpoN1, functions with NifA to regulate the expression of the nif genes, while RpoN2 is required for the transcription of the flagellar genes.

**Stabilizing the Oligomer**

SAXS/WAXS structures and cryo-EM reconstructions of full-length, activated NtrC indicate a role for the DNA binding domains in the stabilization of the oligomer (62). EM reconstructions of the bEBP bound to different nucleotides reveal significant changes in the positions of the D domains during the ATP hydro-
lysis cycle (Fig. 18) (62). In the ADP.AlF$_4$-bound transition state, in which $\sigma^{54}$ contact is strengthened by interactions involving the GAFTGA motif, the DNA binding domains pack closely against the bEBP ring and are therefore likely to distort enhancer DNA. Upon phosphate release and the disengagement of the GAFTGA loop, the DNA binding domains appear to lose their tight association with the ATPase ring. Such conformational changes may stabilize the hexameric arrangement and/or facilitate the interaction between the bEBP and $\sigma^{54}$ (62). Interestingly, studies of NorR indicated that the enhancer DNA itself contributes to the stability of the bEBP oligomer (207). ATP hydrolysis by NorR was stimulated in vitro in the presence of a minimal 66-bp DNA fragment that contained each of the three enhancer sites. However, significantly higher ATPase activity was observed when a longer 266-bp DNA fragment was used, implying that the DNA flanking the enhancer sites has a role in stabilizing the oligomer. In agreement with this hypothesis, electrophoretic mobility shift assays (EMSSAs) revealed significantly higher affinity and cooperativity of binding for the longer fragment. Importantly, negative-stain electron microscopy revealed the formation of protein-DNA complexes with the expected size of a NorR hexamer only in the presence of the 266-bp and not the 66-bp fragment (207). These results suggest that flanking DNA forms extensive interactions with the NorR hexamer that may involve wrapping around the oligomer to form a stable nucleoprotein complex. Structural data for NorR bound to DNA may help to elucidate the importance of DNA wrapping in bEBP oligomer formation.

Maintaining Regulatory Fidelity

Recently, a role for the C-terminal domain (CTD) in determining the specificity of the response by the regulatory domain has emerged. To maintain the regulatory specificity of two-component systems (TCSs), it is important to reduce cross-phosphorylation between noncognate RR-HK pairs or cross talk by low-molecular-weight phosphodonorors. In the case of the FlgSR TCS, the bEBP FlgR must be phosphorylated via its cognate HK FlgS to activate the $\sigma^{54}$-dependent transcription of genes involved in flagellar synthesis. As noted above, in some species, FlgR does not contain a DNA binding domain. However, in the case of Campylobacter jejuni FlgR, which does possess a CTD but lacks the conserved HTH motif, it was proposed that this domain functions to prevent phosphotransfer from acetyl phosphate, thereby increasing the specificity of the response to phosphorylation by FlgS. Accordingly, a truncated derivative of C. jejuni FlgR lacking the CTD is able to activate $\sigma^{54}$-dependent transcription in a FlgS strain, in contrast to wild-type FlgR (119). Presumably, in the absence of the CTD, a noncognate HK or alternative phosphodonor is able to mediate the phosphorylation of the RR, suggesting an alternative role for the CTD in limiting cross talk. The DNA binding function of the C. jejuni FlgR CTD does not significantly contribute to the activity of the bEBP under physiological conditions. Despite FlgR being able to bind DNA in vitro (19), a $\text{flgR}_{\text{ACTD}}$ mutant showed no reduction in the activation of gene expression compared to the wild type (119). Moreover, the deletion of the flgDE2 promoter sequence upstream of position 29 did not significantly reduce expression levels in either the wild type or the $\text{flgR}_{\text{ACTD}}$ mutant, indicating that upstream enhancers are not required for promoter activation (19). The suggestion that C. jejuni FlgR does not require Enhancer sites in vivo is analogous to the situation with Helicobacter pylori, where FlgR naturally lacks a DNA binding domain (26). In line with this, FlgR is the sole activator of $\sigma^{54}$-dependent transcription in both bacterial species. Transposon mutagenesis using the $\Delta$flgS $\text{flgR}_{\text{ACTD}}$ strain confirmed acetyl phosphate as the likely metabolite responsible for the nonspecific activation of FlgR. Furthermore, $\text{flgR}_{\text{ACTD}}$ could be “reprogrammed” to respond to the nutritional status of the cell by increasing acetyl phosphate concentrations by altering the acetogenesis pathway in vivo, although the FlgS-dependent activation of FlgR appears important for the correct assembly of the flagellar apparatus (19). In the absence of crystal structures, the question of how the C-terminal domain of C. jejuni FlgR limits cross talk via acetyl phosphate remains. Possibly, interdomain interactions between the N- and C-terminal domains (9) or long-distance conformational changes limit nonspecific phosphotransfer.

CONCLUDING REMARKS

Since the regulations of $\sigma^{54}$- and $\sigma^{70}$-dependent transcription are so different, it is pertinent to examine the evolutionary advantages of the regulation of the transcription of genes through the activation of $\sigma^{54}$-dependent promoters by bEBPs. Due to the requirement of an activator, transcription is tightly regulated, and $\sigma^{54}$-dependent transcription occurs primarily in response to cellular and extracellular signals that regulate the activity of the AAA$^{+}$ protein. As a result, the activation of transcription occurs rapidly and specifically. This is important since $\sigma^{54}$ commonly binds to the promoters of genes associated with the bacterial stress response. Recent work has shed light on the mechanisms utilized by the variety of response regulator and sensory domains that regu-
late the activity of the enzymatic AAA⁺ domain responsible for the isomerization of the closed promoter complex. Examples of both positive and negative control have been demonstrated, in which the regulatory domain either stimulates or represses the activity of the central domain. Much effort has been expended to characterize the route by which these bEBPs couple signal sensing to substrate remodeling. Structural studies of PspF suggested that there is significant in cis communication between the σ54 interaction determinants on the surface of the bEBP ring and the ATP hydrolysis machinery (173) as well as in trans interactions between subunits of the hexamer (110). Therefore, the mechanisms of regulation that target ATP hydrolysis, oligomerization, and σ54 interactions are likely to be highly interconnected. Considering the interrelatedness of the different control mechanisms, the evolutionary advantage of each one remains unclear. Regulation at the level of ATP hydrolysis or σ54 contact could potentially allow for the assembly of the higher-order oligomer prior to activation, and this may confer a physiological advantage by enabling a rapid stress response. In PspF, the inhibition of a preassembled PspF hexamer by a PspA complex may allow the cell to rapidly respond to the dissipation of the proton motive force (PMF) at the cell membrane (113). In the case of NorR, the preassembly of a NorR hexamer, “poised” as a nucleoprotein complex at the enhancer sites, may enable the cell to rapidly respond to nitrosative stress.

It is important to note that the activator-dependent Eσ54 system of transcription does not remove the competition that exists between the plethora of sigma factors for the core enzyme (107). An important question that remains is whether the alternative sigma factor σ54 reduces the level of the cell’s available polymerase to form closed complexes at target promoters in the absence of the relevant signal(s). While this would presumably bypass the need for recruitment and speed up the transcriptional response upon signal sensing, it seems unlikely that the cell would go to such lengths to provide a global “insurance” mechanism, just in case a particular “stress” is encountered. Moreover, at the time of writing, single-molecule experiments have indicated that the closed complex is dynamic, with RNA polymerase frequently deassociating and rebinding to the promoter (78). This might create a system in which bound RNAP can more rapidly respond to bEBP activation while allowing the exchange of sigma factors within the cellular pool, thus enabling alternative promoter occupancy.

Another unique feature of the σ54 system is the manner in which the activator contacts the enhancer DNA via the “bottom” face of the hexamer and utilizes the “top” GAFTGA loop-containing surface to approach Eσ54 from the unbound face of the promoter DNA. This results in the promoter DNA being sandwiched between the activator hexamer and σ54 (217) and presents opportunities for the activator to participate directly in the DNA-melting event (234). Despite the plethora of biochemical and structural data on bEBPs that has emerged in recent years, we still do not understand how energy-driven interactions between the activator and Eσ54 within this nucleoprotein structure lead to the transition to the open promoter complex. Although a picture is beginning to emerge, whereby only a subset of GAFTGA loops within the hexamer engages with Eσ54 and remodels the closed promoter complex, we still do not know, for example, how many rounds of ATP hydrolysis are required to convert the closed promoter complex into the open promoter complex. Furthermore, although more than one GAFTGA loop is likely to contact the sigma factor, the residues in σ54 that participate in these interactions are unknown. As the ATP hydrolysis cycle in bEBPs is likely to be sequential, which loop interaction initiates the σ54 engagement and/or remodeling activities? There are many mechanistic questions such as these that remain to be answered. We look forward to insightful kinetic and structural analyses in the future, which will help to resolve the molecular details of the relationship between energy coupling and DNA melting in this important protein family.

ACKNOWLEDGMENTS

This work was supported by grants BB/J004561 and BB/D009588 from the Biotechnology and Biological Sciences Research Council and by the John Innes Foundation.

REFERENCES


Drummond MH, Contreras A, Mitchell LA. 1990. The function of isolated domains and chimaeric proteins constructed from the transcrip-
70. Dworkin J, Jovanovic G, Model P. 2000. The PsP A protein of Esche-
richtia coli is a negative regulator of sigma(54)-dependent transcription. J. Bacteriol. 182:311–319.
71. Dworkin J, Jovanovic G, Model P. 1997. Role of upstream activation se-
76. Fernandez S, de Lorenzo V, Perez-Martin J. 1995. Activation of the transcriptional regulator XylR of Pseudomonas putida by release of re-
77. Forsburg SL. 2004. Eukaryotic MCM proteins: beyond replication ini-
79. Gai D, Zhao R, Li D, Finkielstein CV, Chen XS. 2012. Activation of the AAA family transcription factor sigma 70 by release of re-
81. Hoglund E, Douclet M, Wemmer DE. 2009. Structure of the RNA poly-
merase core-binding domain of sigma(54) reveals a likely conforma-
82. Hoover TR, Santero E, Porter S, Kustu S. 1990. The integration host factor stimulates interaction of RNA polymerase with NIF A, the transcrip-
83. Hopper S, Bock A. 1995. Effector-mediated stimulation of ATPase ac-
tivity by the sigma 54-dependent transcriptional activator FHA A from Escherichia coli. J. Bacteriol. 177:2978–2983.
85. Huela E, Ausebel FM. 1989. The central domain of Rhizobium meliloti NifA is sufficient to activate transcription from the R. meliloti nifH pro-
87. Hualong YX, et al. 2006. Protein-induced DNA bending clarifies the architec-
90. Ishihama A. 2000. Functional modulation of Escherichia coli RNA poly-
94. Joly N, Buck M. 2011. Single chain forms of the enhancer binding protein PsP F provide insights into geometric requirements for gene ac-
97. Joly N, Buck M. 2011. Single chain forms of the enhancer binding protein PsP F provide insights into geometric requirements for gene ac-
99. Joly N, Buck M. 2006. Heterogeneous nucleotide occu-
pancy stimulates functionality of phage shock protein F, an AAA + transcrip-
100. Joly N, Buck M. 2012. Coupling AAA protein


Moret E, Cannon W, Buck M. 1988. The DNA-binding domain of the transcriptional activator protein NIFA resides in its carboxyl terminus, recognises the upstream activator sequences of nif promoters and can be separated from the positive control function of NIFA. Nucleic Acids Res. 16:11469–11488.


Matthew Bush first conducted research into the bacterial enhancer binding protein family at the University of York, where he studied Biochemistry. He received his Ph.D. in Molecular Microbiology at the University of East Anglia in 2011, where he carried out research into the regulation of the sigma54-dependent transcriptional activator NorR. He is currently a postdoctoral research fellow at the John Innes Centre, Norwich, United Kingdom.

Ray Dixon is Head of the Department of Molecular Microbiology at the John Innes Centre in Norwich, United Kingdom. He studied microbiology for his undergraduate degree at the University of Reading and obtained a D.Phil. in microbial genetics at the University of Sussex in 1973. He remained at the Unit of Nitrogen Fixation, University of Sussex, to lead a research group focusing on the molecular biology of nitrogen fixation, before moving to his current position in Norwich in 1995. The major aim of his current research is to understand how signals are perceived and integrated to regulate the activity of bacterial enhancer binding proteins.