Published in final edited form as:

Science. 2005 March 25; 307(5717): 1972–1975. doi:10.1126/science.1105932.

Structural Insights Into the Activity of Transcriptional Enhancer-Binding Proteins:

The structure of an activator protein in complex with an RNA polymerase subunit suggests how ATP hydrolysis is coupled to conformational changes that activate transcription in bacteria

Mathieu Rappas^{1,2}, Jorg Schumacher¹, Fabienne Beuron^{1,2}, Hajime Niwa^{1,2}, Patricia Bordes¹, Sivaramesh Wigneshweraraj¹, Catherine A Keetch³, Carol V Robinson³, Martin Buck¹, and Xiaodong Zhang^{1,2,*}

¹ Department of Biological Sciences, Imperial College London, London, SW7 2AZ, UK

² Centre for Structural Biology, Imperial College London, London, SW7 2AZ, UK

³ Department of Chemistry, Cambridge University, Lensfield Road, Cambridge, CB2 1EW, UK

Abstract

Activators of bacterial σ^{54} -RNA polymerase holoenzyme are mechanochemical proteins that use ATP hydrolysis to activate transcription. We have determined a 20 Å resolution structure of an activator, PspF₍₁₋₂₇₅₎, bound to an ATP transition state analog (ADP.AIF_x), in complex with its basal factor σ^{54} by cryo-electron microscopy. By fitting the crystal structure of apo PspF₍₁₋₂₇₅₎ at 1.75 Å into the EM map we identify two loops involved in binding σ^{54} . By comparing enhancer-binding structures in different nucleotide states and mutational analysis, we propose nucleotide dependent conformational changes that free the loops for association with σ^{54} .

Gene expression is regulated at the level of RNA polymerase (RNAP) activity. Bacterial RNAP containing the σ^{54} factor requires specialized activator proteins, referred to as bacterial Enhancer-Binding Proteins (EBPs) that interact with the basal transcription complex from remote DNA sites by DNA looping (1-4). EBPs bind Upstream Activating Sequences (UAS) via their C-terminal DNA-binding domains and form higher order oligomers that use ATP-hydrolysis to activate transcription (5, 6). EBPs' central σ^{54} -RNAP interacting domain is responsible for ATPase activity and transcription activation (7-9) and belongs to the larger AAA+ (ATPase Associated with various cellular Activities) family of proteins (10-12). Well studied EBPs include Phage Shock protein F (PspF), nitrogen fixation protein A (NifA), nitrogen regulation protein C (NtrC), and C₄-dicarboxylic acid transport protein D (DctD) (1-3, 7, 13).

PspF from *Escherichia coli* forms a stable oligomeric complex with σ^{54} at the point of ATP hydrolysis (14). PspF-ADP.AIF_x alters the interaction between σ^{54} and promoter DNA similarly to PspF hydrolyzing ATP (15), and was thus deemed a functional hydrolysis intermediate. Activator nucleotide-hydrolysis dependent events couple the chemical energy of hydrolysis to transcriptional activation. The highly conserved and EBP-specific GAFTGA amino acid motif (Fig. S1) is a crucial mechanical determinant for the successful transfer of energy from ATP hydrolysis in EBP to the RNAP holoenzyme via σ^{54} 's small N-terminal

* To whom correspondence should be addressed. xiaodong.zhang@imperial.ac.uk.

EBP-interacting domain (called Region I, ~ 56 residues and sufficient for PspF interaction) (1, 14, 16-18).

The lack of structural information has hindered progress towards understanding the basis of this energy transfer process required for transcriptional activation. We now present a structure-function analysis of one such system using: 1) a cryo-electron microscopy reconstruction of PspF's AAA+ domain (residues 1-275, PspF₍₁₋₂₇₅₎) in complex with σ^{54} at the point of ATP hydrolysis (mimicked by in-situ formed ADP.AIF_x), 2) the crystal structure of apo PspF₍₁₋₂₇₅₎ at 1.75 Å resolution, and 3) mutational analysis.

Nano-electro spray mass spectroscopy of a PspF₍₁₋₂₇₅₎- σ^{54} complex with ADP.AIF_x established that six monomers of PspF₍₁₋₂₇₅₎ are in complex with a monomeric σ^{54} , consistent with AAA+ proteins functioning as hexamers (10, 12).

The 3-dimensional reconstruction of the PspF₍₁₋₂₇₅₎-ADP.AIF_x- σ^{54} complex (~240 kDa) obtained by cryo-electron microscopy (EM) of native samples (Figs. S3 and S4) shows a PspF₍₁₋₂₇₅₎ hexamer interacting with one σ^{54} . Class averages show well-defined ring-like structures displaying six subunits of PspF₍₁₋₂₇₅₎ within a simple ring and clear extra density for σ^{54} located ~15 Å above the ring (Fig. 1A, B and C). The clear hexagonal ring structure has a diameter of ~125 Å and a central pore of ~20 Å; from the side, the ring is ~40 Å in height (Fig. 1C and Fig. S5B), dimensions consistent with other hexameric AAA+ proteins (19, 20).

Viewed from the side, the hexamer appears concave with a central depression that readily accommodates the σ^{54} density. A 90° rotation along the 6-fold axis reveals that the σ^{54} density, which runs along the ring, is elongated, bent and thicker in the middle (Fig. S5A). This 'horseshoe' shape extra density resembles the envelope model of σ^{54} (21). The estimated mass of σ^{54} is ~30 kDa, significantly higher than the 6 kDa Region I and less than the 54 kDa σ^{54} . Therefore, we postulate that though we see more than just Region I, which is sufficient for PspF binding, there are parts of σ^{54} that are not visualized in our reconstruction because they are mobile (22).

To confirm the presence and integrity of σ^{54} in the particles, we marked N and C-termini of σ^{54} using nanogold beads. Single-cysteine σ^{54} constructs (⁴⁶C σ^{54} and ⁴⁷⁴C σ^{54} , (23)) were covalently linked to nanogold beads before forming the PspF₍₁₋₂₇₅₎-ADP.AIF_x- σ^{54} complex. The negative stained samples were then analyzed using electron microscopy. Detection of nanogold beads in both experiments confirmed the presence and the integrity of σ^{54} in the complex (Fig. 1D, (1) and (2)).

When displayed at lower contour levels the 3D electron density map shows weak densities connecting the PspF₍₁₋₂₇₅₎ ring to σ^{54} (see arrow in Fig. 1C). Based on earlier biochemical results (14, 17), we postulate that the connecting densities found almost at right angle to the PspF₍₁₋₂₇₅₎ ring (Fig. 1C) mark the location of certain GAFTGA motifs within the PspF₍₁₋₂₇₅₎ hexamer in stable association with Region I of σ^{54} .

To facilitate localizing PspF, we determined the crystal structure of apo PspF₍₁₋₂₇₅₎ in space group P6₅ at 1.75 Å resolution using MAD phasing ((24), Table S1). The structure displays a typical AAA+ protein organisation, consisting of an α/β Rossmann fold followed by an α -helical domain (Fig. 2). The GAFTGA motif forms the tip of a loop (L1) inserted into helix3 of the α/β domain. Another loop (L2) consisting of residues 130 to 139 is inserted between helix4 and strand4 (Fig. 2). The extremities of both loops (L1 and L2) show high degrees of flexibility with the tip of L1 being the most flexible, as no electron density was observed for this region (residues 82-89).

Initially a monomer of PspF₍₁₋₂₇₅₎ was visually fitted into the EM map as a rigid body so that the α -helical domain sat in a 'claw' of the ring (Fig. 3A). This fitting positioned L1 and L2 loops in close proximity to one of the weak connecting densities contacting σ^{54} (Fig. 3A and B). From the fitted model we generated a hexamer and then visually readjusted individual subunits to better fit the EM map. L1 and L2 loops were either adjusted to fit the connecting densities or removed when no EM density was observed to account for them ((24) and Fig. S6A and B). It appears that at least two adjacent PspF₍₁₋₂₇₅₎ monomers contact one σ^{54} at the point of ATP hydrolysis. We infer that at the point of ATP hydrolysis, certain L1 and L2 loops extend upwards to maintain a stable interaction of PspF with σ^{54} . When successfully engaged with σ^{54} , L1 and L2 are more structured.

To investigate the link between nucleotide bound state and location of the GAFTGA-containing L1, we compared the apo PspF structure with the structures of different forms of Aquifex aeolicus NtrC1 (PDB code 1NY5 and 1NY6), that has 47 % sequence similarity to PspF in its AAA+ domain, including an ADP bound state thought to be incompetent for stable σ^{54} interaction (3). Our apo PspF₍₁₋₂₇₅₎ crystal structure and that of ATP-soaked crystals (data not shown) are similar suggesting that the apo-structure presented here is close to the ATP bound form which is σ^{54} -interaction competent (14). Aligning on the conserved P-loops resulted in overall good alignment (Fig. 4A) with differences in the relative orientation of the α -helical domain to the α/β domain as well as the position of helices 3 and 4, suggesting a nucleotide-dependent change in domain relationships (12, 25).

Closer examination reveals important differences in the properties and positions of L1 and L2 (Fig. 4A, B and C). In NtrC1 (1NY6), the branch part of L1 is at a right angle to the stem part and points into the central pore of the ring, with a twisted L2 lying in close proximity (Fig. 4B, (13)). Discrete interactions stabilize the branch part of L1 in NtrC1, notably the interactions between E212 in L1 (equivalent to E81 in PspF) and R262 (R131), K267 (Q136) in L2. Also, L263 (V132) of the tip of L2, F216 (F85) of the L1 tip, and A206 (C β of S75) of helix3 form a hydrophobic cluster which locks the GAFTGA motif in a buried and unfavourable conformation for stable σ^{54} interaction (Fig. 4B; Fig. S1). Based on our EM map, in this conformation the GAFTGA motif cannot contact σ^{54} .

In PspF₍₁₋₂₇₅₎, the relative rotation of helices 3 and 4 disrupts the hydrophobic interactions between F85 of L1, V132 of L2, and C β of S75 of helix 3 while interactions between the stem part of L1 and helix3, i.e. between H80 and S75, H92 and E76, are strengthened (Fig. 4C). We hypothesize that these changes in interactions are nucleotide-dependent. Indeed, these interactions form part of a larger network that involves residues R95, R98 of helix3 and S62 of the central β sheet which includes the Walker B motif (D107, E108) responsible for nucleotide hydrolysis (Fig. 4C). This interaction network is well suited to relay nucleotide-dependent movements of the central β sheet, which originate in the Walker B motif, to helix 3 and more importantly to L1 with its GAFTGA motif. A similar network in NtrC1 appears to be based mostly on hydrophobic interactions

To investigate the functionality of the interaction network believed to ultimately control the GAFTGA-containing L1, we conducted structure-based single amino acid mutational analyses. Disruption of the proposed network linking the ATP active site to the GAFTGA loop is predicted to cause a decrease in nucleotide-dependent activities of PspF₍₁₋₂₇₅₎. To examine this hypothesis we mutated H92, which interacts with E76 of helix3 (Fig. 4C), into either F, to mimic the overall geometry while eliminating the charge effects, or to R, to maintain the overall charge (R is also the most frequent residue in this position in EBPs, Fig. S1). R95 on the other hand, which interacts with both E76 and S62 of the central β sheet, was mutated into either A or to the more related K. PspF₍₁₋₂₇₅₎^{R95A} and PspF₍₁₋₂₇₅₎^{H92F} each fail in all measurable post-ATP-binding activities ((26), Fig. S1 and Table S2). In both

cases the mutations severely disrupt the network of interactions and thus block any communication of the conformational signal beyond the mutated residues. In contrast, PspF₍₁₋₂₇₅₎^{R95K} and PspF₍₁₋₂₇₅₎^{H92R}, retained most of the in vitro activities measured, including transcriptional activation (Table S2). The Walker B to L1 network is consistent with properties of many mutant forms of DctD and NtrC (27, 28), supporting the existence of a common EBP nucleotide-dependent communication pathway.

To determine the functionality of the hydrophobic interactions that lock the tip of L1 in the NtrC1 structure, we mutated F85 and V132. PspF₍₁₋₂₇₅₎^{V132A} retained its nucleotide-binding activity but lacked ATPase activity. PspF₍₁₋₂₇₅₎^{F85A} likewise lacked ATPase activity, but PspF₍₁₋₂₇₅₎^{F85L} and PspF₍₁₋₂₇₅₎^{F85W} showed ATP-binding and some ATPase activity with mutation to W leading to an increase in ATPase activity (Table S2). These results suggest that the stability of the L1-L2 hydrophobic interaction has direct consequences on the ATP hydrolysis cycle. PspF₍₁₋₂₇₅₎^{V132A} can engage σ^{54} in an ADP·AlF_x-dependent manner suggesting an effect upon P_i-release (Table S2). Disrupting this interaction would impair this stage of the ATPase cycle whereas strengthening it (e.g. PspF₍₁₋₂₇₅₎^{F85W}) would increase ATPase activity. PspF₍₁₋₂₇₅₎^{F85W} fails to stably bind σ^{54} , demonstrating that the integrity of the GAFTGA motif is crucial for σ^{54} interaction (14, 17). The proposed communication pathway in PspF links changes in the ATP hydrolysis site to conformational changes in the GAFTGA-containing L1 loop that remodels the σ^{54} -RNAP-holoenzyme to activate transcription.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We would like to thank the members of MB and XZ's laboratories for their help and useful discussions throughout this work. We are grateful to Ingar Leiros at ESRF and beamline scientists at Daresbury for their help in data collection, we are also grateful for the support provided by the IC-CBEM. This work is supported by BBSRC funding to XZ and MB. Coordinates for the reported structures have been deposited in the Protein Data Bank (accession codes 2BJW and 2BJV). The EM map has been deposited to EMD (accession code EMD-1109).

References and Notes

1. Rombel I, North A, Hwang I, Wyman C, Kustu S. Cold Spring Harb Symp Quant Biol. 1998; 63:157. [PubMed: 10384279]
2. Buck M, Gallegos MT, Studholme DJ, Guo Y, Gralla JD. J Bacteriol. Aug.2000 182:4129. [PubMed: 10894718]
3. Zhang X, et al. Mol Microbiol. Aug.2002 45:895. [PubMed: 12180911]
4. Wedel A, Kustu S. Genes Dev. Aug 15.1995 9:2042. [PubMed: 7649482]
5. Popham DL, Szeto D, Keener J, Kustu S. Science. Feb 3.1989 243:629. [PubMed: 2563595]
6. Cannon WV, Gallegos MT, Buck M. Nat Struct Biol. Jul.2000 7:594. [PubMed: 10876247]
7. Xu H, Hoover TR. Curr Opin Microbiol. Apr.2001 4:138. [PubMed: 11282468]
8. Morett E, Segovia L. J Bacteriol. Oct.1993 175:6067. [PubMed: 8407777]
9. Studholme DJ, Dixon R. J Bacteriol. Mar.2003 185:1757. [PubMed: 12618438]
10. Lupas AN, Martin J. Curr Opin Struct Biol. Dec.2002 12:746. [PubMed: 12504679]
11. Neuwald AF, Aravind L, Spouge JL, Koonin EV. Genome Res. Jan.1999 9:27. [PubMed: 9927482]
12. Ogura T, Wilkinson AJ. Genes Cells. Jul.2001 6:575. [PubMed: 11473577]
13. Lee SY, et al. Genes Dev. Oct 15.2003 17:2552. [PubMed: 14561776]
14. Chaney M, et al. Genes Dev. Sep 1.2001 15:2282. [PubMed: 11544185]

15. Cannon W, Bordes P, Wigneshweraraj SR, Buck M. J Biol Chem. May 30.2003 278:19815. [PubMed: 12649285]
16. Wang YK, Lee JH, Brewer JM, Hoover TR. Mol Microbiol. Oct.1997 26:373. [PubMed: 9383161]
17. Bordes P, et al. Proc Natl Acad Sci U S A. Mar 4.2003 100:2278. [PubMed: 12601152]
18. Gonzalez V, Olvera L, Soberon X, Morett E. Mol Microbiol. Apr.1998 28:55. [PubMed: 9593296]
19. Beuron F, et al. J Mol Biol. Mar 28.2003 327:619. [PubMed: 12634057]
20. Furst J, Sutton RB, Chen J, Brunger AT, Grigorieff N. Embo J. Sep 1.2003 22:4365. [PubMed: 12941689]
21. Svergun DI, Malfois M, Koch MH, Wigneshweraraj SR, Buck M. J Biol Chem. Feb 11.2000 275:4210. [PubMed: 10660585]
22. Murakami KS, Masuda S, Darst SA. Science. Jun 17.2002 296:1280. [PubMed: 12016306]
23. Burrows PC, Severinov K, Ishihama A, Buck M, Wigneshweraraj SR. J Biol Chem. Aug 8.2003 278:29728. [PubMed: 12750380]
24. See supporting online materials for details
25. Wang J, et al. Structure (Camb). Nov.2001 9:1107. [PubMed: 11709174]
26. Bordes P, Wigneshweraraj SR, Zhang X, Buck M. Biochem J. Mar 15.2004 378:735. [PubMed: 14659000]
27. Wang YK, Hoover TR. J Bacteriol. Sep.1997 179:5812. [PubMed: 9294439]
28. Li J, Passaglia L, Rombel I, Yan D, Kustu S. J Bacteriol. Sep.1999 181:5443. [PubMed: 10464219]

Fig 1. Electron Microscopy (EM) analysis of the PspF₍₁₋₂₇₅₎-ADP.AIF_x-σ⁵⁴ complex

A: 3 of the 123 final class averages used to generate the 3D cryo-EM reconstruction together with their assigned Euler angles (in white, euler angles: α , β , γ). B: Respective reprojected images of the 3D model. C: Surface renderings of the final 3D cryo-EM map of the complex; bottom and side views of the complex are displayed. Also shown in mesh is the EM map displayed at a lower threshold to highlight the ring to σ^{54} connecting densities. D: (1): Negative-stain of the PspF₍₁₋₂₇₅₎-ADP.AIF_x-⁴⁶Cσ⁵⁴ nanogold-labelled complex; arrows indicate the nanogold bead; (2): Negative-stain of the PspF₍₁₋₂₇₅₎-ADP.AIF_x-⁴⁷⁴Cσ⁵⁴ nanogold-labelled complex.

Fig 2. Crystal structure of PspF₍₁₋₂₇₅₎

P6₅ hexamer of PspF₍₁₋₂₇₅₎ as viewed down the six-fold axis. Both α/β (green) and α -helical domains (magenta) of one monomer are contoured using dashed lines. The nucleotide binding pocket is highlighted using a light yellow shade and is located in the cleft between the α/β and α -helical domain at the interface with the adjacent monomer. N and C-termini of two adjacent monomers are also shown. Colour coding: blue: helices, red: central β -sheet, orange: L1 and green: L2. The tip of the highlighted L1 is shown as a dotted line because residues 82-89 were not resolved in our crystal structure.

Fig 3. Fitting of PspF₍₁₋₂₇₅₎ crystal structure into the EM electron density map of the PspF₍₁₋₂₇₅₎-ADP.AIF_x-σ⁵⁴ complex

A: front view of the EM density -transparent grey- and the PspF₍₁₋₂₇₅₎ crystal structure - blue ribbon representation - . The fitting of the α-helical domain into one 'claw' of the hexameric ring is highlighted; the densities connecting PspF₍₁₋₂₇₅₎ to σ⁵⁴ are indicated by red arrows. B: cross-eye stereo view of the fitting of one pair of L1 and L2 loops into the connecting densities; the tip of L1 is shown as a dotted line because residues 82-89 are not resolved. Note that when positioned in the densities, the loops extend upwards almost at right angle to the plane of the hexamer.

Fig 4. Nucleotide-dependent relocation of L1 and L2 loops

A: Ribbon representation of the P-loop-superimposed AAA+ structures of NtrC1 –red- and PspF₍₁₋₂₇₅₎ –blue-. L1 loop structures are in orange, darker shade for NtrC1 L1 and the tip of PspF₍₁₋₂₇₅₎'s L1 as dotted line. L2 loop structures are in green, darker shade for NtrC1. The NtrC1 L1-GAFTGA motif is shown as sticks. Apparent relocations of the α -helical domain and helices 3 and 4 are highlighted by arrows. B: Close-up of the α/β domain of NtrC1; colour coding: yellow - Walker A, cyan - Walker B, orange - L1 and green - L2. The L1-GAFTGA motif, A206 and L263 are shown as sticks. Also shown in sticks are E212 and R262 to highlight their role in coordinating L1 and L2. C: Close-up of the α/β domain of PspF₍₁₋₂₇₅₎ using the same colour coding as mentioned above. Residues as sticks are part of the Walker B motif to L1 network. Also shown in sticks is the catalytic K42 residue of the Walker A motif. The L1-GAFTGA motif is shown as dotted line.