Published in final edited form as:
J Mol Biol. 2013 January 9; 425(1): 156–170. doi:10.1016/j.jmb.2012.10.017.

Structural mechanism of GAF-regulated σ^{54} activators from *Aquifex aeolicus*

Joseph D. Batchelor, **Peter S. Lee**, **Andrew C. Wang**, **Michaeleen Doucleff**, and **David E. Wemmer***
Graduate Group in Biophysics and Department of Chemistry, University of California, Berkeley and Physical Biosciences Division, Lawrence Berkeley National Laboratory Berkeley, CA 94720

Abstract

The σ subunits of bacterial RNA polymerase occur in many variant forms, and confer promoter specificity to the holopolymerase. Members of the σ^{54} family of σ subunits require the action of a 'transcriptional activator' protein to open the promoter and initiate transcription. The activator proteins undergo regulated assembly from inactive dimers to hexamers that are active ATPases. These contact σ^{54} directly and, through ATP hydrolysis, drive a conformational change that enables promoter opening. σ^{54} activators use several different kinds of regulatory domains to respond to a wide variety of intracellular signals. One common regulatory module, the GAF domain, is used by σ^{54} activators to sensesmall molecule ligands. The structural basis for GAF-domain regulation in σ^{54} activators has not previously been reported. Here, we present crystal structures of GAF regulatory domains for *Aquifex aeolicus* σ^{54} activators Nlh2 and Nlh1 in three functional states – an 'open', ATPase inactive state; a 'closed', ATPase inactive state; and a 'closed', ligand-bound, active ATPase state. We also present small angle X-ray scattering (SAXS) dataforNlh2 linked GAF-ATPase domains in the inactive state. These GAF domain dimers regulate σ^{54} activator proteins by holding the ATPase domains in an inactive dimer conformation. Ligand binding of Nlh1 dramatically remodels the GAF domain dimer interface, disrupting the contacts with the ATPase domains. This mechanism has strong parallels to the response to phosphorylation in some two-component regulated σ^{54} activators. We describe a structural mechanism of GAF-mediated enzyme regulation that appears to be conserved among humans, plants, and bacteria.

Keywords

transcriptional activator; sigma-54; enhancer-binding protein; structural activation mechanism; signal transduction

*Corresponding author. University of California, 508C Stanley Hall, MC-3220, Berkeley, CA 94720-3220, USA. Phone: +1 (510) 666-2683, FAX: +1 (510) 643-1255. dewemmer@berkeley.edu.
Present addresses: J. D. Batchelor, Department of Molecular Microbiology, Washington University School of Medicine, Saint Louis, MO 63110; P. S. Lee, Department of Molecular Biology, The Scripps Research Institute, La Jolla, CA 92037, USA; A. C. Wang, Division of Chemistry, California Institute of Technology and Howard Hughes Medical Institute, Pasadena, CA 91125; M. Doucleff, National Public Radio, Washington, DC 20001

Introduction

Sigma factors are the subunits of bacterial RNA polymerase that allow the holoenzyme to recognize promoter sequences.[1–3] There are two families of bacterial sigma factors, σ^{70} and σ^{54}, designated by the size of the founding members of the families. While the σ^{70} family has many subfamilies with varied molecular weights,[4] there is only a single member of the σ^{54} family. The activity of σ^{70} family members is commonly regulated by anti-σ^{70} proteins that sequester σ^{70}, preventing assembly of the holoenzyme, or by preventing the σ^{70} holoenzyme from binding promoters through competition with repressor proteins that bind DNA elements overlapping the polymerase binding site. Polymerase with a σ^{70} subunit is intrinsically competent for transcription. In contrast, σ^{54} family members do not have anti-σ proteinsor repressors that compete for the promoters. Instead σ^{54} -polymerase binds the promoter in an inactive conformation, and requires the action of an activator protein to open the promoter and initiate transcription.[5] σ^{54} promoters are generally found in only a small percentage of genes and control tightly regulated processes such as the nitrogen and sulfate limitation responses, the heavy metal toxicity response, or nitrogenase synthesis.[5–8] An exception to this rule is *Myxococcus xanthus*, where σ^{54} controls fruiting body development and the large numbers of genes are under σ^{54} control.[9]

The σ^{54} transcriptional activators are AAA+ ATPases that undergo regulated assembly to form active hexamers that contact σ^{54} and drive a conformational change that is required for promoter opening and transcription initiation. As for many other proteins with multiple functions, these activators contain multiple structural domains with distinct functions. σ^{54} activators are generally composed of an N-terminal regulatory domain (or domains), a central AAA+ ATPase domain, and a C-terminal sequence specific DNA binding domain.[10] Some activators are exceptions to this pattern, such as phage shock protein F (*E.coli* PspF), which has no regulatory domain and is regulated in trans by PspA,[11] as well as variants with PAS or other regulatory domains in single or multiple copies.[10] The regulatory domains sense intracellular signals and serve as an allosteric switch to convert the activator from an inactive dimer to an active AAA+ ATPase hexamer ring that is able to drive the conformational change in polymerase that is needed to initiate σ^{54} mediated transcription.[12–15]

The most common of regulatory modules are receiver domains, which are activated through phosphorylation by a histidine kinase (a two-component signaling system). There have been studies that detail the regulation mechanisms of σ^{54} activators with regulatory receiver domains.[12; 13; 15–25] Phosphorylation causes a conformational change in the receiver domain that can either stimulate an interaction with the ATPase domain of a neighboring activator molecule driving formation of the hexamer,[26] or cause a change in dimer interface that leads to a loss of interaction with the ATPase domains to release them to spontaneously assemble into the active hexamers.[17; 18]

The next most common regulatory elements are GAF domains. The GAF domain evolved 2 billion years ago, before plants and animals evolved from bacteria, and are found in all phyla of life[27–29] and its acronym comes from the first three classes of proteins found to contain this domain: c**G**MP-specific and regulated cyclic nucleotidephosphodiesterases, **A**denylyl cyclases, and *E. coli*σ^{54} activator **F**hlA.[27] GAF domains bind small molecule ligands,[27; 30; 31] and often serve a sensing function for N- or C-terminal effector domains including histidine kinases, enzymes, and ATPases.[29] σ^{54} activators with GAF domains have been shown to regulate acetoin and carbohydrate metabolism,[32; 33] sense nitric oxide[34; 35] and dibenzofuran,[36] serve as a nitrate/nitrile reductase sensor[37] and a formate hydrogenlyase sensor,[38] and regulate the transcription of nitrogenase and alternate nitrogenase.[39–41] Although there have been several functional studies of σ^{54} activators with

GAF domains,[35; 40; 42–45] the structural mechanism of GAF-mediated AAA+ regulation remains unknown.

Here we report structures of GAF domains from two *Aquifex aeolicus*(*A.a.*) σ^{54} activators in three functional states. The*A.a.*NifA-like homolog 2GAF domain (Nlh2-G) was solved in a GAF-closed/ATPase-inactive state. We show Nlh2's GAF domain represses ATPase activity in the absence of ligand, indicating thatNlh2is a negatively regulated σ^{54} activator, and we use small angle X-ray scattering (SAXS) to show Nlh2has an off-state analogous to σ^{54} activators with regulatory receiver domains such as *A.a.*NtrC1.Two slightly different*A.a.*NifA-like homolog 1(Nlh1) GAF domain constructs crystallized in distinct conformation states - GAF-open/ATPase inactive and GAF-closed/ATPase-active. These two structures, together with the data on Nlh2, show the mechanism by which ligand binding to GAF domains modulates the assembly of the ATPase domains. From the similarity to NtrC1 and NtrC4 we conclude that σ^{54} activators with different regulatory domains share a conserved general mechanism of repression and activation.

Results

In the extreme thermophile *Aquifex aeolicus*, six σ^{54} activators were identified to have sequence homology to other known activators. When the genome was annotated, five were given gene names NtrC1–4 and NifA, suggesting homology to the *E. coli* proteins with these names. Further work showed the NifA homolog, here referred to as NifA-like homolog 1 (Nlh1), stimulated transcriptional activity[42]. Examination of the sequences shows that NtrC1 (Aq_1117), NtrC3 (Aq_230), and NtrC4 (Aq_164) contain N-terminal receiver domains, and hence are homologs of *E.coli*NtrC (nitrogen regulatory protein C), although there is almost nothing known about the genes that are regulated by these proteins in *A.a.* Examination of the sequence of NtrC2 (Aq_1792) showed that it contains an N-terminal GAF domain, like the NifA homolog, rather than the receiver domain implied by the name. Because of this, we refer to NtrC2 here as NifA-like homolog 2 (Nlh2). The sixth activator was initially thought to be nonfunctional because of an insertion of a stop codon in the region of the required ATPase domain, however subsequent analysis indicates that it is likely a functional protein, and contains a regulatory PAS domain (Aq_093).

Crystal structure of Nlh2's GAF domain in its closed apo-state

A native construct of *A.a.*Nlh2-GAF domain (Nlh2-G, residues 1 to 172 of Nlh2) crystallized in the spacegroup $P3_2$ 21 with two molecules in the asymmetric unit. A double mutant, L36M V42M, was made to enable selenomethionine multiwavelength anomalous dispersion (MAD) phasing, as Nlh2-G's single, centrally located methionine was not sufficient for MAD experimental phase determination. The L36M V42Mdouble mutant crystallized in the spacegroup $P3_1$ 21 with one molecule in the asymmetric unit. MAD was used to obtain initial experimental phases. Molecular replacement, using the SeMet model as a search model, was performed to obtain initial phases for the native Nlh2-G diffraction data. We refined the native model at a resolution of 1.70 Å with an R_{work} of 16.3% and an R_{free} of 20.7%.

Nlh2-G crystallized as a homodimer with noncrystallographic symmetry relating the two GAF domains in the asymmetric unit (Fig. 1a). We found that Nlh2-G is dimeric in solution, as assessed using gel filtration chromatography(data not shown). The asymmetric unit likely contains the physiologically relevant dimer, as this dimerization interface buries 1150 $Å^2$ per subunit with a predicted $\Delta\Delta G°$ of dimerization of −16.0 kcal/mol (http://www.ebi.ac.uk/msdsrv/ prot_int/). The fold of Nlh2-G is similar to that in previously solved GAF domains.[31; 46–51] A five stranded antiparallel beta sheet is located between a three helix bundle and two long linker regions with a single helix. These linker regions (L1 and L2,

approximately 15 and 20 residues in length) connect β2 to α3 and β3 to β4, respectively, and have no regular secondary structure. L1 and L2 form a 'mouth' which encloses the GAF domain's activating ligand in previously solved structures, though there is no ligand in Nlh2's ligand binding pocket.

Crystal structure of *A.a.*Nlh1's GAF domain in its open, apo state

A.a. Nlh1-GAF domain (apo Nlh1-G1, residues 1–175) crystallized in the spacegroup $P4_12_12$ with two molecules in the asymmetric unit (Fig. 1b). Molecular replacement, using Nlh2-G as a search model, was performed to obtain initial phases. We refined the model at a resolution of 3.05 Å with an R_{work} of 20.74% and an R_{free} of 24.22%.We expect that the asymmetric unit contains the physiologically relevant dimer interface, as this interface buries 1067Å^2 per subunit with a predicted ΔΔG° of dimerization of -18.5 kcal/mol (http://www.ebi.ac.uk/msd-srv/prot_int/). Like Nlh2-G, Nlh1-G1 is composed of a five-stranded antiparallel beta sheet with a three helix bundle at one face and at the opposing side of the beta sheet, two long linker regions, L1 and L2 and an additional helix. Unlike Nlh2-G, most of L1 and L2 are completely disordered (Fig. 1b), and beta strand 2 of apoNlh1-G1, which contacts L1 in Nlh2-G, is substantially displaced compared to the Nlh2-Gpocket.

Crystal structure of Nlh1's GAF domain in its closed ligand-bound state

A slightly shorter construct of *A.a.* Nlh1-G (ligand-bound Nlh1-G2,residues 1–171) crystallized in the spacegroup $P4_32_12$ with one molecule in the asymmetric unit (Fig. 1c). Molecular replacement, using Nlh2-G as a search model, was performed for initial phases. We refined the model at a resolution of 2.80 Å with an R_{work} of 20.42% and an R_{free} of 23.70%. Ligand-bound Nlh1-G2 is globally similar to Nlh2-G, except that α1 is completely unstructured and not visible in the electron density. Because α1 is disordered, ligand-bound Nlh1-G2 cannot form the extensive three-helix bundle parallel homodimer interface observed in Nlh1-G1's off state and Nlh2-G. Instead, Nlh1-G2's homodimer interface consists of a 2-helix bundle with a ~50° rotation, relative to Nlh1-G1's offstate or Nlh2-G. This new homodimer interface buries 771Å^2 per subunit with a predicted ΔΔG° of dimerization of −16.9 kcal/mol (http://www.ebi.ac.uk/msd-srv/prot_int/).

The refined structure of Nlh1-G2 (1–171) showed clear density in the ligand binding pocket (Fig. 2). Since no ligand was added during purification the bound molecule must have copurified after expression. The density corresponds to ~5 or 6 C/N/O atoms. A larger region of density at one end which is proximal to an arginine residue in the GAF domain suggest that one end of the ligand may have a carboxyl group. At the current resolution of the structure an unambiguous identification of the ligand is not possible. Previous work has shown that the *A.a.*NifA homolog 1 binds to the promoter region of the GlnB gene (Aq_109), suggesting a role in nitrogen regulation (not nitrogen fixation as the name of the homolog might imply), however this is not sufficient information to suggest the regulatory ligand.

A SAXS derived model of Nlh2-GC resembles the NtrC1-RC off-state

To determine the organization of Nlh2 GAF domain in the off-state with the ATPase domain present, we analyzed SAXS data of Nlh2-GC. SAXS can be used to accurately determine the radius of gyration and the maximum dimensions of the molecule, and is sensitive to the global shape of a protein, which together can be used to test structural models.[52; 53] Experimental scattering data closely matched predicted scattering resulting from a model based on the off-state of NtrC1-RC (Fig. 3a). This model was created by swapping the receiver domain of NtrC1-RC with the structure of Nlh2-G(Fig. 3). These data show that in the absence of ligand,Nlh2's GAF and central ATPase domains resemble the off-state of NtrC1's receiver and ATPase domains.

Nlh2's GAF domain represses activity of its central ATPase domain

We performed ATPase activity assays on Nlh2-GCD (full length) and Nlh2-CD (GAF domain deletion). Measurements of the ATP hydrolysis rates of Nlh2-GCD and Nlh2-CD shows that the GAF domain represses ATPase activity. At 70° C, Nlh2-GCD hydrolyzed ATP at a rate of 0.13±0.03 min^{-1}, while Nlh2-CD hydrolyzed ATP at a rate of 0.70±0.13 min^{-1}, a 5.5 fold increase. This is similar to the levels of repression seen in other negatively regulated σ^{54} activators such as NtrC1,[20] and is also comparable to the suppression of activity in PDE2A by its GAF domains (~4 fold at half saturating ligand).[54]

Discussion

The Nlh2-G structure and activity support a model of negative regulation

The *A.a.* Nlh2 GAF domain has a fold that is very similar to previously described GAF domains, as expected from sequence homology. Size exclusion chromatography showed that Nlh2-G is dimeric in solution, and there is a clear dimer interface between proteins in the crystal that is consistent with dimer interfaces identified in other GAF domain structures. One particular feature of interest is the fact that the C-terminal helices of the two monomers pack together at the center of the interface (Fig. 1a). The C-terminal site chosen for expression of the construct we have studied was based on homology to other GAF domains, and indeed the expected fold terminates very close to the last residue in the construct. When a secondary structure prediction is done for the amino acid sequence for the full length protein, an extended C-terminal helix is predicted that was not included in our construct, which suggests a signaling helix connecting to the ATPase domains.[55]

σ^{54} activators are classified as being either positively or negatively regulated. Positively regulated σ^{54} activators, such as *E. coli* NtrC, cannot assemble and become active as ATPases without the aid of an activated regulatory domain. For NtrC it has been shown that the phosphorylated receiver domain binds to the ATPase domain of a neighboring monomer, stabilizing the assembled, active ATPase hexamer ring[26]. Negatively regulated σ^{54} activators, such as DctD, NtrC1 and NtrC4, have ATPase domains that are intrinsically competent for assembly and ATP hydrolysis, and are repressed from assembling by interactions of unactivated receiver domains.[17; 18; 20] Upon activation by phosphorylation of the receiver domains, a conformational change disrupts the receiver-ATPase interface, freeing the ATPase domains for assembly. Our measurements of Nlh2 ATPase activity show that a construct lacking the GAF domain, Nlh2-CD, is an active ATPase, and the presence of a GAF domain represses ATPase activity and ATPase assembly. We observed a 5.5 fold increase in specific activity in a construct lacking the GAF regulatory domain. This increase agrees well with previously studied negatively receiver domain regulated σ^{54} activators. An approximate 20 fold increase in transcriptional activation was seen *in vivo* for *A.a.* NtrC1.[20] This *in vivo* work was done in *E. coli* on a heterologous promoter. Positively regulated σ^{54} activators have no ATPase activity in the absence of activated regulatory domains. Thus these results clearly indicate thatNlh2 is a negatively regulated σ^{54} activator.

From our understanding of the regulation of *A.a.* receiver domain containing activators DctD, NtrC1 and NtrC4, it seemed likely that C-terminal signaling helices from Nlh2-G domains could hold the ATPase domains in the inactive face-to-face conformation. We tested this hypothesis by creating a model with extended signaling helices, whose length corresponded to the helix length predictions from the sequence connecting the GAF domain dimer, and positioning it with the ATPase dimer from the crystal structure of NtrC1. The SAXS curves from this model fit extremely well with the experimental SAXS data collected for the GAF-ATPase domain construct, Fig.3. These data further support the model that the

'off-state' (no ligand bound) GAF domain of Nlh2 maintains the ATPase domains in an inactive conformation, repressing transcriptional activation.

Nlh2-G has an empty ligand binding pocket

The Nlh2-G protein has a binding pocket in the expected location based on homology to other solved GAF domain structures. The L1 and L2 loops that define the lid of the pocked are in a closed conformation (Fig. 4) as seen in other published structures[31; 46–51] completely separating the interior of the pocket from the exterior solvent region. In our structure there is no ligand density in the region of the ligand-binding pocket. The binding pocket is too small to enclose cyclic nucleotides that are common ligands for GAF domains. Bioinformatic analysis of *Aquifex aeolicus*indicates that the Nlh2 gene is likely part of an operon, but the functions of the proteins are not known and hence provide no clues regarding possible ligands for the Nlh2GAF domain.Nlh2 is not near any of the identified σ^{54} promoters in the *A.a.* genome.

The apo-Nlh1 GAF domain structure is similar to Nlh2-G but with open ligand binding pockets

The structure determined for the longer of the two Nlh1 GAF domain constructs, Nlh1-G1, is quite similar to that of the Nlh2-G(Fig.1a and 1b). In particular the dimer interface contains three helices from each monomer brought together, with the C-terminal helix extending away from the core of the GAF domain, bringing the two copies into contact. This is again consistent with formation of a signaling helix pair that would hold the ATPase domains in an inactive conformation. The surprising finding in this structure is that the ligand binding pocket is open, with the loop segment, which is folded down over the pocket in the Nlh2-G structure, disordered (Fig.1b and Fig 4a, b). It seems likely that these loops are flexible for the protein in solution, and may well fluctuate between this kind of open structure, and the closed empty structure as observed for Nlh2-G, which could be transiently populated but due to the absence of ligand could not drive further conformational rearrangements discussed below.

The Nlh1 GAF domain undergoes substantial conformation rearrangements upon ligand binding

The structure of the Nlh1 GAF domain construct, truncated four residues earlier at the C-terminal end, Nlh1-G2, crystalized with an adventitious ligand bound (Fig. 3). A comparison of this structure with that of the longer version shows that the Nlh1GAF domain undergoes a substantial conformation change upon binding its activating ligand (Fig. 4). Upon ligand binding, the loops surrounding Nlh1-G2's ligand binding pocket, L1 and L2, close around the ligand, β2 moves in5 Å towards the ligand, and α1 becomes unstructured (Fig. 4a). α1 forms a large part of the homodimer interface of Nlh1's off-state, so destabilization of α1 results in an alternate Nlh1-G dimer interface created by α2 and α4 (Fig. 4b). This alternate Nlh1 dimer is expected to be the active state, as it lacks α4 contacts required for repression of the ATPase domain. Though it is formally possible these conformational changes result from the removal of four residues from Nlh1-G's C-terminus, we strongly believe these are caused by ligand binding, as ligand binding causes similar conformational changes in homologous GAF-mediated systems.[54; 56] This is further supported by a crystal structure of Nlh1-G2 in essentially the same conformation as Nlh1-G1 (data not shown).

These gross structural changes result from an altered set of side chain contacts. Notably, a proline at position 28 is present in the second turn of α2, which undoubtedly weakens this helix. Indeed, helical contacts N-terminal to Pro 28 are perturbed in Nlh1's on state but are present in Nlh1's off state. β2, which undergoes a 5 Å conformational change upon ligand binding, is likely responsible for stabilizing the secondary structure of residues N-terminal to

proline-28 and ultimately stabilizing α1 (Fig. 4c). In the absence of ligand, Leu 23 of α2 makes stabilizing hydrophobic contacts with Val 57 of β2 (Fig. 4c). Upon ligand binding, Val 57 instead contacts Val 102 of L2, which is on the opposite side of the strand (Fig. 4d). This swaps the positions of Val 57 and Arg 58 and results in an unusual beta strand, with sequential sidechains (Arg 58 and Ala 59) on the same side of the strand (Fig. 4e). In addition to destabilizing α2 through loss of hydrophobic contacts between Val 57 and Leu 23, Arg 58 at this position would come within 5Å of Arg 163 of α3 in Nlh1 off-state's signal helix. Though the ligand-bound Nlh1 construct's C-terminus was residue 171, there was no visible density past residue 160. This likely is due to Arg 58 repelling Arg 163 and directly destabilizing the signal helix α4.Thus, ligand binding results in Arg 58 contacting α2 in place of Val 57, destabilization of α2 N-terminal to Pro 28 and direct destabilization of the signal-helix α4.The specific residues which play a role in *A.a.* Nlh1 activation are present in a subset of GAF-regulated σ^{54} activators, including NifA homologs from rhizosphere bacteria*Azospirillum brasilense* and *Azospirillum lipoferum*, but do not appear to be present in other widely studied GAF-regulated σ^{54} activators including NifA from *K. pneumoniae* and *A. vinlandii* or NorR from *E. coli* (Fig. 5).

Ligand binding results in an alternate Nlh1-G dimer interface

The most significant change upon ligand binding is the destabilization of α1. α1 forms the majority of homodimer contacts in both Nlh1-G1's inactivated state and also in Nlh2-G. Surprisingly, destabilization of α1 in ligand-bound Nlh1-G2 leads to a completely new homodimer interface created from α2 and α4 contacts (Fig. 1b and 1c). This on-state interface is predicted to be nearly as favorable as the off-state interface, with a predicted ΔΔG° of dimerization of −16.9 kcal/mol compared to the predicted off-state homodimer interface with a ΔΔG° of dimerization of −18.5 kcal/mol. This ligand-induced homodimer interface is created by a rotation of ~50° of the two subunits, about an axis roughly perpendicular to the plane of the dimer interface. Such an interface is also observed in human phosphodiesterases[54] as discussed below. This rotation could not occur in the presence of α1, and results in a complete disruption of the homodimer coiled-coil contacts of the C-terminal helix, α4. This C-terminal helix is a 'signaling helix',[55] a coiled-coil connecting the GAF domain to the central AAA+ ATPase domain. The change in dimer interface induced by ligand binding to the*A.a.*Nlh1GAF domain results in the disruption of the α4 signal helix interactions and releases theNlh1ATPase domains to assemble into the active ATPase hexamer (Fig. 6).

GAF-regulated σ^{54} activators use diverse regulatory mechanisms

The regulatory mechanism identified here of ligand binding resulting in an altered GAF homodimer interface which disrupts the signaling helix required for repression of ATPase assembly contrasts with the regulatory mechanism previously identified for a GAF-mediated σ^{54} activator from *E. coli*, NorR. NorR, which responds to nitric oxide,[35] assembles into an inactive hexamer in the absence of activation.[57] This preassembly is thought to allow the cell to rapidly respond to NO stress. Rather than disrupting assembly, NorR employs an intramolecular mechanism in which the GAF domain directly targets the ATPase domain's σ^{54} interaction surface. This preassembled hexamer state is possible because NorR has an unstructured GAF-ATPase linker region, unlike the coiled-coil signaling helices joining the GAF and ATPase domain of both *A.a.* Nlh1 and Nlh2. In the case of Nlh1 and Nlh2, the coiled-coil linker presumably results in a strongly dimeric off-state, and GAF activation disrupts this off-state to allow for hexameric assembly.

Many GAF domains show a conserved mechanism of ligand-mediated activation

The GAF domain is a common sensory domain that regulates many biological activities in addition to σ^{54} activators, occurring in many diverse pathways. Four structurally

characterized systems regulated by GAF domains include vertebrate phosphodiesterases (PDEs), the Ile responsive global regulatory CodY, bacterial quorum signaling proteins, and plant phytochromes.

The behavior of the ligand binding pocket loops in Nlh1 described above is likely relevant to the *Agrobacterium tumefaciens* quorum-sensing regulator TraR. TraR consists of a ligand binding domain that is structurally similar to a GAF domain and a DNA binding domain.[58–60] TraR activates gene expression upon binding *N*-3-oxooctanoyl-l-homoserine lactone. In the absence of activating ligand, apo-TraR is rapidly proteolysed *in vivo* and *in vitro*, presumably due to the presence of unfolded regions,[61–63] however in the presence of ligand it is stable. A likely explanation for this behavior is that the L2 loop in TraR's ligand binding domain is open and unstructured in the absence of ligand, analogous to the loops in Nlh1's GAF domain. In this state the loops would be susceptible to intracellular proteases, but when ligand binds the loops become structured around it and hence would be much less likely to be cleaved.

Another well characterized GAF regulated system is CodY, a global regulator of stationary phase and virulence in gram positive bacteria.[46] In *Bacillus subtilius*, CodY is a global regulator of up to 200 genes.[64] Recent reports[65; 66] have presented structures of CodY's regulatory GAF domain in its apo and ligand-bound states. CodY and Nlh1 appear to have different mechanisms of regulation. Unlike Nlh1, CodY's apo and liganded states have nearly identical dimer interfaces, and the mechanism of signal transduction for CodY's GAF domain is unknown. Both Nlh1's and CodY's L1 and L2 regions undergo large conformational rearrangements, but Nlh1 undergoes more dramatic conformation changes, including disorder in portions of L2 in the apo state, and α1 in the ligand-bound state. Thus, different GAF domains may have different activation mechanisms, as seen in receiver domain regulated σ^{54} activator proteins.[14; 20; 67]

Plant and cyanobacterial phytochromes also have significant similarities to GAF regulated σ^{54} activators. Phytochromes sense the quality of light by existing in red (Pr) and farred (Pfr) conformations, mediating plant and cyanobacterial shade avoidance responses [68]. Phytochromes are composed of interlocking N-terminal PAS, GAF, and PHY domains connected to a C-terminal histidine kinase by a long signaling helix. The GAF domain has a covalently bound bilin chromophore[69; 70] that undergoes a reversible conformational change between the Pr and Pfr states. The change in the chromophore geometry alters the protein conformation, ultimately modulating the activity of the histidine kinase. Yang et al. showed that the conformational change leads to alteration of the interface between the two equivalent GAF domains, leading to a reorientation of the signaling helix that connects to the output domains. This process corresponds to an approximately 30°rotation of the signaling helix upon dark activated photoconversion of the GAF domain's bilin chromophore to the Pfr state[47] compared to the Pr state,[71; 72] which is similar to the change in the *A.a.* Nlh1 dimer, a ~60°rotation upon GAF ligand binding. However in the phytochrome the α1 helix is still present in the 'on' signaling state, so that the detailed mechanism is different although the gross conformational alteration is similar.

PDE5 and PDE6 are important enzymes for secondary messenger regulated signal transduction in vertebrate smooth muscle cells and photoreceptors, respectively. These phosphodiesterases have a regulatory domain composed of two N-terminal GAF domains, denoted GAF-A and GAF-B. GAF-A of PDE5 and PDE6 has been shown to undergo a conformation change upon binding cGMP.[50; 51; 73] PDE6's GAF-A domain was shown to be substantially more thermodynamically stable upon binding cGMP, changing the T_{melt} by ~20°C,[50] and [^{1}H, ^{15}N] TROSY-HSQC spectra of a PDE5 GAF AB construct had many missing amide peaks in the apo state, presumably from regions that become unstructured,

which were present in the liganded state. These results are consistent with our more unstructured ‘open’ and more structured ‘closed’ Nlh1-G structures.

The structural mechanism of GAF-mediated regulation has been extensively characterized in PDE2A, as structures exist for both the liganded tandem regulatory GAF domains[56] and for the unliganded GAF domains in a construct also containing the C-terminal phosphodiesterase domain.[54] These structures show PDE2A GAF-B undergoes a conformational change upon ligand binding that is rather similar to the conformation change upon ligand binding observed for Nlh1. Like Nlh1, ligand binding of PDE2A GAF-B results in the ordering of loops L1 and L2 around the bound ligand, and a more subtle propagated conformational change that leads to repacking of the interface between the two GAF-B domains in the dimer. For PDE2A the result is a displacement of helix α1 from the dimer interface, as occurs for Nlh1-GAF, but rather than becoming disordered this helix flips nearly 180° and becomes part of the signaling helix connecting the GAF domains to phosphodiesterase catalytic domain. In both PDE2A and Nlh1ligand binding results in a ~30° rotation of α5, the C-terminal helix that is at the dimer interface.

GAF activation, followed by rotation of α5 has been observed in human phosphodiesterases, plant phytochromes, and now bacterial σ^{54} activators. We postulate α5 rotation is a commonly used feature of GAF-mediated regulation. This rotation destabilizes contacts, which allows regulatory GAF domains to repress their effector domains, resulting in activation of C-terminal effector domains. As this structural mechanism of GAF-mediated repression and activation is conserved among bacterial, plant, and human GAF domains, it may well have been present when the first primordial GAF domains evolved 2 billion years ago. However, it should be noted that not all GAF domains are activated in this manner, notably the GAF domain of NorR which, as discussed previously, has an unstructured GAF-ATPase linker region and represses activity using an intramolecular mechanism rather than controlling oligomerization.

Diverse regulatory folds use a similar repression mechanism

A coiled-coil signaling helix links the regulatory and ATPase domains in many different proteins, including σ^{54} activators.[55] In the activators, repression of assembly relies on the docking of this dimeric coiled coil signaling helix into the linked ATPase domains, holding them in a face to- face orientation that is inactive. Upon receiving a signal - phosphorylation for NtrC1[18; 20] and DctD,[21; 74] or ligand binding for Nlh1 and presumably Nlh2 -the signaling helix is disrupted and the ATPase domains can reorient to pack in an active front-to-back orientation. Association of three dimers forms the complete AAA+ ring assembly that is active to stimulate σ^{54} mediated transcriptional initiation, as shown schematically in Fig.6.Thus, although the details of the signaling helix disruption are completely different, the functional roles and basic mechanics of the receiver domains and GAF domains are very similar in regulating activators, indicating a general structural mechanism of σ^{54} activator regulation – disruption of the coiled-coil linker. Biochemical studies have previously shown that diverse regulatory domains generally repress assembly of the active oligomer in this class of protein and this work suggests that many of these regulatory domains will have a common mechanism.

Materials and Methods

Protein expression and purification

Nlh2-G (residues 1–172), Nlh2-GC (residues 1–372), Nlh2-CD (residues 131–442), Nlh2-GCD (residues 1–442), Nlh1-G1(residues 1-175), and Nlh1-G2(residues 1–171) were subcloned into a PSKB3 plasmid with a His6 tag. All proteins were expressed in *E. coli*

BL21 (DE3) with a Rosetta. pLysS plasmid using Studier's auto induction protocol (Studier 2005). Cells were resuspended in 300mM NaCl, 50mM tris pH 8.2, 1mM DTT, and 0.1mM PMSF with one Roche Complete EDTA-free antiproteolysis tablet (buffer A). Cells were harvested by sonication and lysates were heated for 30 minutes at 70° C. Lysate was spun at 30,000 r.p.m. to remove cellular debris. The supernatant was run over a Ni-agarose column, washed with buffer A, eluted with buffer A plus 300 mM imidazole, and then dialyzed overnight into buffer A. To remove the N-terminal His tag, Nlh2-G, Nlh1-G1, and Nlh1-G2 were incubated with 0.01 mg/mL TEV protease for 16 hours at 4° C.

Activity assays

Samples (300 μl) of Nlh2-GCD and Nlh2-CD at 10 μ M were incubated at 70°C for 5 min. Then, 3 μl of 60 mM ATP was added to the samples for a final ATP concentration of 1 mM. At various time points, 50-μl aliquots were quenched with 350 μl of 0.88 M HNO^3. A color developing solution (350 μl) consisting of 44 am $Bi(NO_3)_2$, 31 μm $(NH_4)6\ Mo_7\ O_{24}$, and 0.11% ascorbic acid was added to the quenched reactions. Exactly 3 min after the addition of color-developing solution, absorbance at 700 nm was measured, which is proportional to the amount of free phosphate in solution.

Nlh2-G crystallization

Crystals of Nlh2-G (residues 1–172) and SeMet Nlh2-G (L36M V42M) were grown at room temperature using hanging-drop vapor diffusion. One micro liter of protein solution (20 mg/ml) was mixed with an equal volume of well solution containing 22% PEG 550 monomethyl ether, 150mM $MgCl_2$, and 100mM HEPES pH 7.0. Crystals of both native Nlh2-G and the double methionine SeMet mutant appeared after 1 day. They grew to 1.0 mm × 0.5 mm × 0.5 mm after 1 day and were hexagon-shaped. Twenty-four hours before data collection, 5 M NaCl was added to the wells to a final concentration of 1.0 M NaCl for cryoprotection. Crystals were flash-cooledin liquid nitrogen prior to data collection.

Apo Nlh1-G crystallization

Crystals of Nlh1-G1 (residues 1–175) were grown at room temperature using hanging-drop vapor diffusion. One microliter of protein solution (20 mg/ml) was mixed with an equal volume of well solution containing 32.5% PEG 3350, 250 mM NaCl, and 50 mM MES pH 6.2. Crystals appeared after 1 day. They grew to 0.4 mm × 0.4 mm × 0.1 mm after 1 week and were diamond-shaped. Crystals were flash-cooledin liquid nitrogen prior to data collection.

Ligand-bound Nlh1-G crystallization

Crystals of Nlh1-G2 (residues 1– 171) were grown at room temperature using hanging drop vapor diffusion. One microliter of protein solution (20 mg/ml) was mixed with an equal volume of well solution containing 0.2M Na/K phosphate, 0.1M Bis Tris propane pH 7.5, and 20% PEG 3350. Crystals appeared after 2 weeks. They grew to 0.2 mm × 0.2 mm × 0.1 mm and were diamond-shaped. Crystals were flash-cooledin liquid nitrogen prior to data collection.

Structure determination and refinement

Data for the three structures were collected at beamline 8.3.1 of the Advanced Light Source at Lawrence Berkeley National Laboratory. The data were processed and scaled with MOSFLM[75] and the CCP4 program suite[76] using the ELVES interface[77]. The initial Nlh2 SeMet phases were obtained from MAD phasing with SOLVE[78] in PHENIX.RESOLVE was used to build an Nlh2-G SeMet model.[79; 80] Molecular replacement using the PHASER[81] implementation of phenix.automr, with the SeMet model as a search model, was

performed to obtain initial the higher resolution Nlh2-G native phases. Molecular replacement using PHASER with the native Nlh2-G structure as a search model was performed to obtain the initial apo Nlh1-G1 phases. Molecular replacement using phenix.automr with a partly refined apo Nlh1-G1 structure as a search model was performed to obtain initial ligand-bound Nlh1-G2 phases. Density modification and partial automated building of all three structures was performed using RESOLVE in PHENIX.[79; 80] The three amino-acid sequences were fitted to the electron density map using Coot[82]. Refinement was performed with phenix.refine. Data collection and refinement statistics for Nlh2-G and the Nlh1-G proteins are presented in Tables 1 and 2, respectively.

SAXS data collection and rigid body modeling

Small angle X-ray scattering (SAXS) data of a C-terminal domain truncation construct of Nlh2-GC (residues 1–373) were collected at ALS beamline 12.3.1. Buffer (50 mM Tris, 200 mM NaCl [pH 8.5]) scattering intensities were subtracted from the sample scattering data. PRIMUS[83] was used for initial SAXS data processing. SAXS data were collected at Nlh2-GC protein concentrations of 10mg/mL, 5mg/mL, and 2.5mg/mL. 10mg/mL data was used in our analysis. We modeled these data using the program BUNCH[84] to fit atomic resolution structures to our SAXS data. We assumed P2 symmetry because of the dimeric state of the GAF domains, and all known off-state σ54 activators have an approximate 2-fold symmetry axis. To ensure an adequate sampling, we averaged 100 BUNCH runs using DAMAVER,[85] and aligned the 23 solutions which corresponded to the average fit (Fig2a). The theoretical scattering of our rigid body model fit well to our SAXS data, with an average chi value of 4.6 ± 0.5.

The coordinates for the structures have been deposited in the PDB with IDs: **4G3V** (Nlh2-G); **4G3K** (Nlh1-G1 apo); **4G3W** (Nlh1-G2 ligand bound).

Acknowledgments

We would like to thank Prof. Sydney Kustu for helpful discussions. X-ray diffraction data were collected at the Advanced Light Source at Lawrence Berkeley National Lab, which is supported by the Director, Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. This work was supported by grant R01 GM62163 from the National Institutes of Health.

Abbreviations used

A.a.	*Aquifex aeolicus*
SAXS	small angle X-ray scattering
MAD	multiwavelength anomalous dispersion
AAA+	ATPases Associated with diverse cellular Activities
PDE	phosphodiesterase

References

1. Burgess RR, Travers AA, Dunn JJ, Bautz EKF. Factor Stimulating Transcription by RNA Polymerase. Nature. 1969; 221:43–46. [PubMed: 4882047]
2. Wosten MM. Eubacterial sigma-factors. FEMS Microbiol. Rev. 1998; 22:127–150. [PubMed: 9818380]
3. Gross CA, Chan C, Dombroski A, Gruber T, Sharp M, Tupy J, Young B. The Functional and Regulatory Roles of Sigma Factors in Transcription. Cold Spring Harbor Symposia on Quantitative Biology. 1998; 63:141–156.

4. Paget MS, Helmann JD. The sigma70 family of sigma factors. Genome Biol. 2003; 4:203. [PubMed: 12540296]
5. Hirschman J, Wong PK, Sei K, Keener J, Kustu S. Products of nitrogen regulatory genes ntrA and ntrC of enteric bacteria activate glnA transcription in vitro: evidence that the ntrA product is a sigma factor. Proc Natl Acad Sci U S A. 1985; 82:7525–7529. [PubMed: 2999766]
6. Garcia E, Bancroft S, Rhee SG, Kustu S. The product of a newly identified gene, glnF, is required for synthesis of glutamine synthetase in Salmonella. Proc Natl Acad Sci U S A. 1977; 74:1662–1666. [PubMed: 16262]
7. Habe H, Kouzuma A, Endoh T, Omori T, Yamane H, Nojiri H. Transcriptional regulation of the sulfate-starvation-induced gene sfnA by a sigma-54-dependent activator of Pseudomonas putida. Microbiology. 2007; 153:3091–3098. [PubMed: 17768252]
8. Leonhartsberger S, Huber A, Lottspeich F, Böck A. The hydH/G genes from Escherichia coli code for a zinc and lead responsive two-component regulatory system. Journal of Molecular Biology. 2001; 307:93–105. [PubMed: 11243806]
9. Jakobsen JS, Jelsbak L, Jelsbak L, Welch RD, Cummings C, Goldman B, Stark E, Slater S, Kaiser D. ?54 Enhancer Binding Proteins and Myxococcus xanthus Fruiting Body Development. Journal of Bacteriology. 2004; 186:4361–4368. [PubMed: 15205438]
10. Studholme DJ, Dixon R. Domain architectures of sigma54-dependent transcriptional activators. J Bacteriol. 2003; 185:1757–1767. [PubMed: 12618438]
11. Joly N, Burrows PC, Engl C, Jovanovic G, Buck M. A Lower-Order Oligomer Form of Phage Shock Protein A (PspA) Stably Associates with the Hexameric AAA+ Transcription Activator Protein PspF for Negative Regulation. Journal of Molecular Biology. 2009; 394:764–775. [PubMed: 19804784]
12. Volkman BF, Nohaile MJ, Amy NK, Kustu S, Wemmer DE. Three-dimensional solution structure of the N-terminal receiver domain of NTRC. Biochemistry. 1995; 34:1413–1424. [PubMed: 7827089]
13. Nohaile M, Kern D, Wemmer D, Stedman K, Kustu S. Structural and functional analyses of activating amino acid substitutions in the receiver domain of NtrC: evidence for an activating surface. J Mol Biol. 1997; 273:299–316. [PubMed: 9367763]
14. Kern D, Volkman BF, Luginbuhl P, Nohaile MJ, Kustu S, Wemmer DE. Structure of a transiently phosphorylated switch in bacterial signal transduction. Nature. 1999; 402:894–898. [PubMed: 10622255]
15. Volkman BF, Lipson D, Wemmer DE, Kern D. Two-state allosteric behavior in a single-domain signaling protein. Science. 2001; 291:2429–2433. [PubMed: 11264542]
16. Batchelor JD, Sterling HJ, Hong E, Williams ER, Wemmer DE. Receiver Domains Control the Active-State Stoichiometry of Aquifex aeolicus Ï*f*54 Activator NtrC4, as Revealed by Electrospray Ionization Mass Spectrometry. Journal of Molecular Biology. 2009; 393:634–643. [PubMed: 19699748]
17. Batchelor JD, Doucleff M, Lee C-J, Matsubara K, De Carlo S, Heideker J, Lamers MH, Pelton JG, Wemmer DE. Structure and regulatory mechanism of Aquifex aeolicus NtrC4: Variability and evolution in bacterial transcriptional regulation. J. Mol. Biol. 2008; 384:1058–1075. [PubMed: 18955063]
18. Lee SY, De La Torre A, Yan D, Kustu S, Nixon BT, Wemmer DE. Regulation of the transcriptional activator NtrC1: structural studies of the regulatory and AAA+ ATPase domains. Genes Dev. 2003; 17:2552–2563. [PubMed: 14561776]
19. Hastings CA, Lee SY, Cho HS, Yan D, Kustu S, Wemmer DE. High-resolution solution structure of the beryllofluoride-activated NtrC receiver domain. Biochemistry. 2003; 42:9081–9090. [PubMed: 12885241]
20. Doucleff M, Chen B, Maris AE, Wemmer DE, Kondrashkina E, Nixon BT. Negative regulation of AAA + ATPase assembly by two component receiver domains: a transcription activation mechanism that is conserved in mesophilic and extremely hyperthermophilic bacteria. J Mol Biol. 2005; 353:242–255. [PubMed: 16169010]

21. Nixon BT, Yennawar HP, Doucleff M, Pelton JG, Wemmer DE, Krueger S, Kondrashkina E. SAS solution structures of the apo and Mg2+/BeF3(−)- bound receiver domain of DctD from Sinorhizobium meliloti. Biochemistry. 2005; 44:13962–13969. [PubMed: 16229485]
22. Chen B, Doucleff M, Wemmer DE, De Carlo S, Huang HH, Nogales E, Hoover TR, Kondrashkina E, Guo L, Nixon BT. ATP Ground- and Transition States of Bacterial Enhancer Binding AAA+ ATPases Support Complex Formation with Their Target Protein, sigma-54. Structure. 2007; 15:429–440. [PubMed: 17437715]
23. Sallai L, Tucker PA. Crystal structure of the central and C-terminal domain of the [sigma]54-activator ZraR. Journal of Structural Biology. 2005; 151:160–170. [PubMed: 16005641]
24. Chen B, Sysoeva TA, Chowdhury S, Guo L, De Carlo S, Hanson JA, Yang H, Nixon BT. Engagement of Arginine Finger to ATP Triggers Large Conformational Changes in NtrC1 AAA+ ATPase for Remodeling Bacterial RNA Polymerase. Structure. 2010; 18:1420–1430. [PubMed: 21070941]
25. Sterling HJ, Batchelor JD, Wemmer DE, Williams ER. Effects of Buffer Loading for Electrospray Ionization Mass Spectrometry of a Noncovalent Protein Complex that Requires High Concentrations of Essential Salts. Journal of the American Society for Mass Spectrometry. 2010; 21:1045–1049. [PubMed: 20226685]
26. De Carlo S, Chen B, Hoover TR, Kondrashkina E, Nogales E, Nixon BT. The structural basis for regulated assembly and function of the transcriptional activator NtrC. Genes Dev. 2006; 20:1485–1495. [PubMed: 16751184]
27. Aravind L, Ponting CP. The GAF domain: an evolutionary link between diverse phototransducing proteins. Trends in Biochemical Sciences. 1997; 22:458–459. [PubMed: 9433123]
28. Martinez SE, Beavo JA, Hol WGJ. GAF Domains: Two-Billion- Year-Old Molecular Switches that Bind Cyclic Nucleotides. Molecular Interventions. 2002; 2:317–323. [PubMed: 14993386]
29. Zoraghi R, Corbin JD, Francis SH. Properties and Functions of GAF Domains in Cyclic Nucleotide Phosphodiesterases and Other Proteins. Mol Pharmacol. 2004; 65:267–278. [PubMed: 14742667]
30. Anantharaman V, Koonin EV, Aravind L. Regulatory potential, phyletic distribution and evolution of ancient, intracellular small-molecule-binding domains. Journal of Molecular Biology. 2001; 307:1271–1292. [PubMed: 11292341]
31. Ho Y-SJ, Burden LM, Hurley JH. Structure of the GAF domain, a ubiquitous signaling motif and a new class of cyclic GMP receptor. EMBO J. 2000; 19:5288–5299. [PubMed: 11032796]
32. Deutscher J, Francke C, Postma PW. How Phosphotransferase System-Related Protein Phosphorylation Regulates Carbohydrate Metabolism in Bacteria. Microbiology and Molecular Biology Reviews. 2006; 70:939–1031. [PubMed: 17158705]
33. Kruger N, Steinbuchel A. Identification of acoR, a regulatory gene for the expression of genes essential for acetoin catabolism in Alcaligenes eutrophus H16. The Journal of Bacteriology. 1992; 174:4391–4400.
34. Pohlmann, Anne; Cramm, Rainer; Schmelz, Karin; Friedrich, Bärbel. A novel NO-responding regulator controls the reduction of nitric oxide in Ralstonia eutropha. Molecular Microbiology. 2000; 38:626–638. [PubMed: 11069685]
35. D'Autreaux B, Tucker NP, Dixon R, Spiro S. A non-haem iron centre in the transcription factor NorR senses nitric oxide. Nature. 2005; 437:769–772. [PubMed: 16193057]
36. Iida T, Waki T, Nakamura K, Mukouzaka Y, Kudo T. The GAF-Like- Domain-Containing Transcriptional Regulator DfdR Is a Sensor Protein for Dibenzofuran and Several Hydrophobic Aromatic Compounds. The Journal of Bacteriology. 2009; 191:123–134.
37. Fedtke I, Kamps A, Krismer B, Gotz F. The Nitrate Reductase and Nitrite Reductase Operons and the narT Gene of Staphylococcus carnosus Are Positively Controlled by the Novel Two-Component System NreBC. The Journal of Bacteriology. 2002; 184:6624–6634.
38. Schlensog V, Bock A. Identification and sequence analysis of the gene encoding the transcriptional activator of the formate hydrogenlyase system of Escherichia coli. Molecular Microbiology. 1990; 4:1319–1327. [PubMed: 2280686]
39. Joerger RD, Jacobson MR, Bishop PE. Two nifA-like genes required for expression of alternative nitrogenases by Azotobacter vinelandii. The Journal of Bacteriology. 1989; 171:3258–3267.

40. Dixon R, Kennedy C, Kondorosi A, Krishnapillai V, Merrick M. Complementation analysis of Klebsiella pneumoniae mutants defective in nitrogen fixation. Molecular and General Genetics MGG. 1977; 157:189–198.
41. Buchanan-Wollaston V, Cannon MC, Beynon JL, Cannon FC. Role of the nifA gene product in the regulation of nif expression in Klebsiella pneumoniae. Nature. 1981; 294:776–778. [PubMed: 6119619]
42. Studholme DJ, Wigneshwereraraj SR, Gallegos M-T, Buck M. Functionality of Purified σ^N(σ^{54}) and a NifA-Like Protein from the Hyperthermophile *Aquifex aeolicus*. Journal of Bacteriology. 2000; 182:1616–1623. [PubMed: 10692367]
43. Little R, Dixon R. The amino-terminal GAF domain of Azotobacter vinelandii NifA binds 2-oxoglutarate to resist inhibition by NifL under nitrogen-limiting conditions. J Biol Chem. 2003; 278:28711–28718. [PubMed: 12759352]
44. Martinez-Argudo I, Little R, Dixon R. Role of the amino-terminal GAF domain of the NifA activator in controlling the response to the antiactivator protein NifL. Mol Microbiol. 2004; 52:1731–1744. [PubMed: 15186421]
45. Martinez-Argudo I, Little R, Shearer N, Johnson P, Dixon R. The NifL-NifA System: a multidomain transcriptional regulatory complex that integrates environmental signals. J Bacteriol. 2004; 186:601–610. [PubMed: 14729684]
46. Sonenshein AL. CodY, a global regulator of stationary phase and virulence in Gram-positive bacteria. Current Opinion in Microbiology. 2005; 8:203–207. [PubMed: 15802253]
47. Yang X, Kuk J, Moffat K. Crystal structure of Pseudomonas aeruginosa bacteriophytochrome: Photoconversion and signal transduction. Proceedings of the National Academy of Sciences. 2008; 105:14715–14720.
48. Asen I, Djuranovic S, Lupas AN, Zeth K. Crystal Structure of SpoVT, the Final Modulator of Gene Expression during Spore Development in Bacillus subtilis. Journal of Molecular Biology. 2009; 386:962–975. [PubMed: 18996130]
49. Podust LM, Ioanoviciu A, Ortiz de Montellano PR. 2.3 Å X-ray Structure of the Heme-Bound GAF Domain of Sensory Histidine Kinase DosT of Mycobacterium tuberculosis. Biochemistry. 2008; 47:12523–12531. [PubMed: 18980385]
50. Heikaus CC, Stout JR, Sekharan MR, Eakin CM, Rajagopal P, Brzovic PS, Beavo JA, Klevit RE. Solution Structure of the cGMP Binding GAF Domain from Phosphodiesterase 5: Insights into Nucleotide Specificity, Dimerization, and cGMP-Dependent Conformational Change. Journal of Biological Chemistry. 2008; 283:22749–22759. [PubMed: 18534985]
51. Martinez SE, Heikaus CC, Klevit RE, Beavo JA. The Structure of the GAF A Domain from Phosphodiesterase 6C Reveals Determinants of cGMP Binding, a Conserved Binding Surface, and a Large cGMP-dependent Conformational Change. J. Biol. Chem. 2008; 283:25913–25919. [PubMed: 18614542]
52. Petoukhov MV, Svergun DI. Analysis of X-ray and neutron scattering from biomacromolecular solutions. Current Opinion in Structural Biology. 2007; 17:562–571. [PubMed: 17714935]
53. Putnam CD, Hammel M, Hura GL, Tainer JA. X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. Quarterly Reviews of Biophysics. 2007; 40:191–285. [PubMed: 18078545]
54. Pandit J, Forman MD, Fennell KF, Dillman KS, Menniti FS. Mechanism for the allosteric regulation of phosphodiesterase 2A deduced from the Xray structure of a near full-length construct. Proc Natl Acad Sci U S A. 2009; 106:18225–18230. [PubMed: 19828435]
55. Anantharaman V, Balaji S, Aravind L. The signaling helix: a common functional theme in diverse signaling proteins. Biology Direct. 2006; 1:25. [PubMed: 16953892]
56. Martinez SE, Wu AY, Glavas NA, Tang X-B, Turley S, Hol WGJ, Beavo JA. The two GAF domains in phosphodiesterase 2A have distinct roles in dimerization and in cGMP binding. Proceedings of the National Academy of Sciences. 2002; 99:13260–13265.
57. Bush M, Ghosh T, Tucker N, Zhang X, Dixon R. Nitric oxideresponsive interdomain regulation targets the ?54-interaction surface in the enhancer binding protein NorR. Molecular Microbiology. 2010; 77:1278–1288. [PubMed: 20624215]

58. Zhang LH, Murphy PJ, Kerr A, Tate ME. Agrobacterium Conjugation and Gene-Regulation by N-Acyl-L-Homoserine Lactones. Nature. 1993; 362:446–448. [PubMed: 8464475]
59. Zhang RG, Pappas T, Brace JL, Miller PC, Oulmassov T, Molyneauz JM, Anderson JC, Bashkin JK, Winans SC, Joachimiak A. Structure of a bacterial quorum-sensing transcription factor complexed with pheromone and DNA. Nature. 2002; 417:971–974. [PubMed: 12087407]
60. Vannini A, Volpari C, Di Marco S. Crystal structure of the quorumsensing protein TraM and its interaction with the transcriptional regulator TraR. Journal of Biological Chemistry. 2004; 279:24291–24296. [PubMed: 15044488]
61. Zhu J, Winans SC. The quorum-sensing transcriptional regulator TraR requires its cognate signaling ligand for protein folding, protease resistance, and dimerization. Proceedings of the National Academy of Sciences of the United States of America. 2001; 98:1507–1512. [PubMed: 11171981]
62. Zhu J, Winans SC. Autoinducer binding by the quorum-sensing regulator TraR increases affinity for target promoters in vitro and decreases TraR turnover rates in whole cells. Proceedings of the National Academy of Sciences of the United States of America. 1999; 96:4832–4837. [PubMed: 10220379]
63. Pinto, Uelinton M.; Winans, Stephen C. Dimerization of the quorumsensing transcription factor TraR enhances resistance to cytoplasmic proteolysis. Molecular Microbiology. 2009; 73:32–42. [PubMed: 19432796]
64. Molle V, Nakaura Y, Shivers RP, Yamaguchi H, Losick R, Fujita Y, Sonenshein AL. Additional Targets of the Bacillus subtilis Global Regulator CodY Identified by Chromatin Immunoprecipitation and Genome-Wide Transcript Analysis. J. Bacteriol. 2003; 185:1911–1922. [PubMed: 12618455]
65. Levdikov VM, Blagova E, Joseph P, Sonenshein AL, Wilkinson AJ. The Structure of CodY, a GTP- and Isoleucine-responsive Regulator of Stationary Phase and Virulence in Gram-positive Bacteria. J. Biol. Chem. 2006; 281:11366–11373. [PubMed: 16488888]
66. Levdikov VM, Blagova E, Colledge VL, Lebedev AA, Williamson DC, Sonenshein AL, Wilkinson AJ. Structural Rearrangement Accompanying Ligand Binding in the GAF Domain of CodY from Bacillus subtilis. Journal of Molecular Biology. 2009; 390:1007–1018. [PubMed: 19500589]
67. Lee J, Owens JT, Hwang I, Meares C, Kustu S. Phosphorylation induced signal propagation in the response regulator ntrC. J Bacteriol. 2000; 182:5188–5195. [PubMed: 10960104]
68. Rockwell NC, Su Y-S, Lagarias JC. Phytochrome Structure and Signaling Mechanisms. Annual Review of Plant Biology. 2006; 57:837.
69. Wu S-H, Lagarias JC. Defining the Bilin Lyase Domain: Lessons from the Extended Phytochrome Superfamilyâ€. Biochemistry. 2000; 39:13487–13495. [PubMed: 11063585]
70. Lagarias JC, Rapoport H. Chromopeptides from phytochrome. The structure and linkage of the PR form of the phytochrome chromophore. Journal of the American Chemical Society. 1980; 102:4821–4828.
71. Wagner JR, Brunzelle JS, Forest KT, Vierstra RD. A lightsensing knot revealed by the structure of the chromophore-binding domain of phytochrome. Nature. 2005; 438:325–331. [PubMed: 16292304]
72. Yang X, Stojković EA, Kuk J, Moffat K. Crystal structure of the chromophore binding domain of an unusual bacteriophytochrome, RpBphP3, reveals residues that modulate photoconversion. Proceedings of the National Academy of Sciences. 2007; 104:12571–12576.
73. Francis SH, Bessay EP, Kotera J, Grimes KA, Liu L, Thompson WJ, Corbin JD. Phosphorylation of isolated human phosphodiesterase-5 regulatory domain induces an apparent conformational change and increases cGMP binding affinity. Journal of Biological Chemistry. 2002; 277:47581–47587. [PubMed: 12359732]
74. Park S, Meyer M, Jones AD, Yennawar HP, Yennawar NH, Nixon BT. Two-component signaling in the AAA + ATPase DctD: binding Mg2+ and BeF3- selects between alternate dimeric states of the receiver domain. Faseb J. 2002; 16:1964–1966. [PubMed: 12368235]
75. Powell H. The Rossmann Fourier autoindexing algorithm in MOSFLM. Acta Crystallographica Section D. 1999; 55:1690–1695.

76. Bailey S. The CCP4 suite: programs for protein crystallography. Acta Crystallogr D Biol Crystallogr. 1994; 50:760–763. [PubMed: 15299374]

77. Holton J, Alber T. Automated protein crystal structure determination using ELVES. Proc Natl Acad Sci U S A. 2004; 101:1537–1542. [PubMed: 14752198]

78. Terwilliger TC, Berendzen J. Automated MAD and MIR structure solution. Acta Crystallogr D Biol Crystallogr. 1999; 55:849–861. [PubMed: 10089316]

79. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung L-W, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallographica Section D. 2010; 66:213–221.

80. Terwilliger T. Maximum-likelihood density modification. Acta Crystallographica Section D. 2000; 56:965–972.

81. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. Journal of Applied Crystallography. 2007; 40:658–674. [PubMed: 19461840]

82. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr. 2004; 60:2126–2132. [PubMed: 15572765]

83. Konarev PV, Volkov VV, Sokolova AV, Koch MHJ, Svergun DI. PRIMUS: a Windows PC-based system for small-angle scattering data analysis. Journal of Applied Crystallography. 2003; 36:1277–1282.

84. Petoukhov MV, Svergun DI. Global Rigid Body Modeling of Macromolecular Complexes against Small-Angle Scattering Data. Biophysical journal. 2005; 89:1237–1250. [PubMed: 15923225]

85. Volkov VV, Svergun DI. Uniqueness of ab initio shape determination in small-angle scattering. Journal of Applied Crystallography. 2003; 36:860–864.

Highlights

- GAF domains regulate many different proteins by responding to small molecule ligands
- a sigma54 activator GAF domain represses ATPase domain assembly using an extended helical pair
- sigma54activator GAF domain binding pocket loops can open in the absence of ligand
- ligand binding changes the dimer interface between GAF domains that allows ATPase assembly
- new activator GAF structures have provided insight into their detailed regulatory mechanism

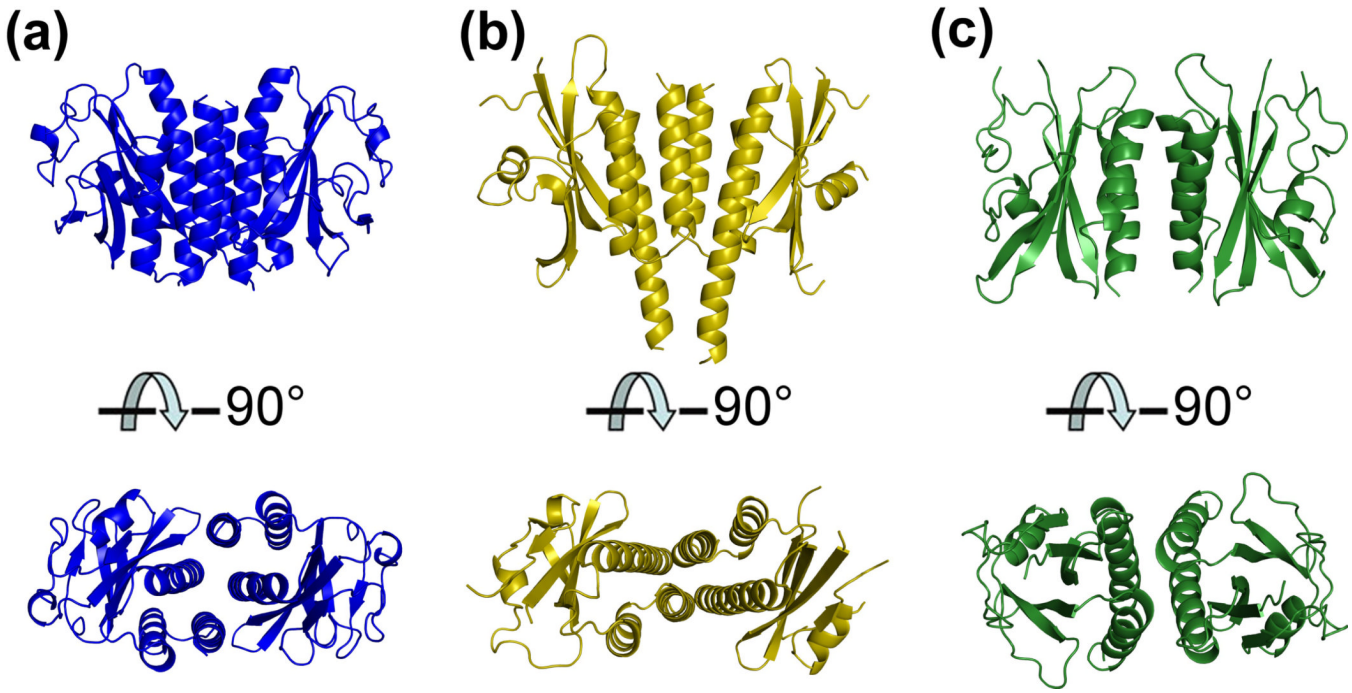

Fig. 1.

Overview of the three structures. (a) Apo *A.a.* Nlh2-G, in blue has loops closed over an empty ligand-binding pocket with a homodimer interface created by three helices. (b) Apo *A.a.* Nlh1-G1, in yellow has loops in an open conformation, and a homodimer interface is created by three helices. (c) Ligand-bound *A.a.* Nlh1-G2, in green, has loops in a closed conformation like Nlh2-G, and has a homodimer interface created by two helices.

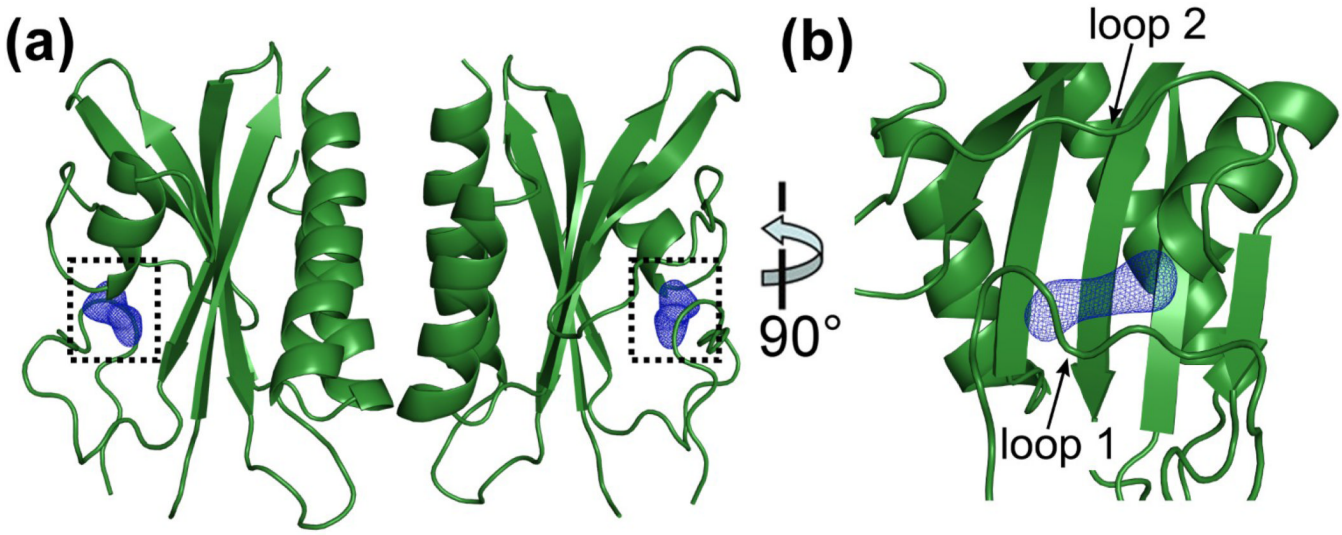

Fig. 2.

Ligand density present in ligand-bound Nlh1-G2.(a) Difference density is clearly seen in an F_o-F_c difference map, contoured at 3.5 σ and represented as a blue mesh. (b) A frontal view, two loops enclose the Nlh1 ligand.

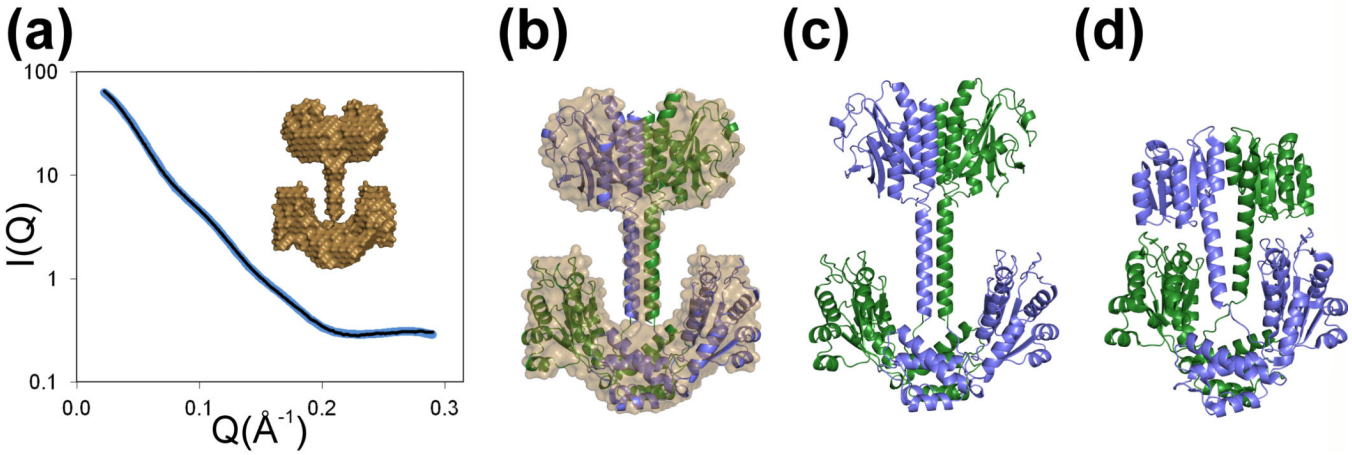

Fig. 3.

SAXS of Nlh2-GC shows a dimeric off-state with a similar architecture to the receiver domain regulated activator protein NtrC1. (a) Experimental data (black) are in excellent agreement with predicted scattering of a rigid-body model (blue and shown in tan) where two GAF domains sit above two ATPase domains. (b) Locations of Nlh2 GAF and ATPase domains, shown in blue and green, within the SAXS model, shown in tan. This model for Nlh2 (c) is similar to the NtrC1 off-state, mediated by receiver domains (d).

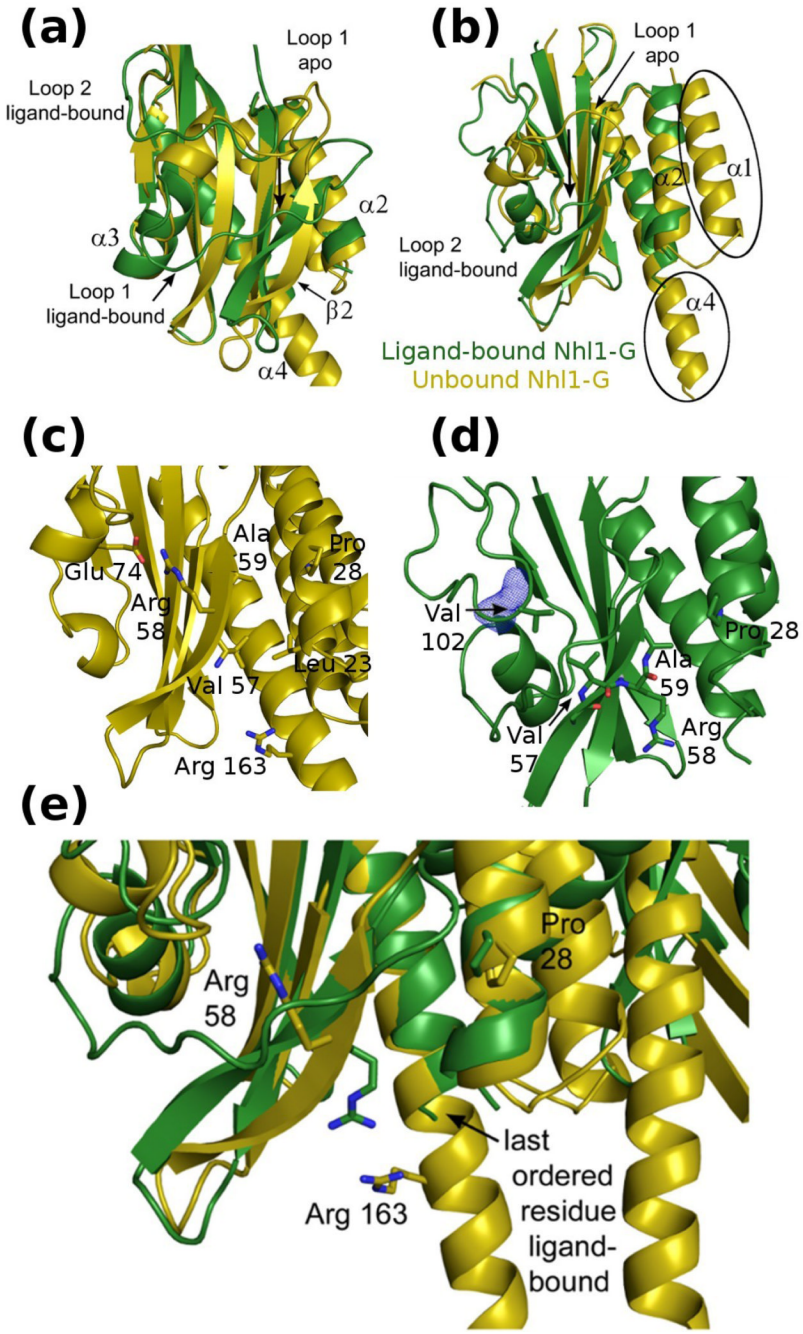

Fig. 4.

A superposition of unbound apo Nlh1-G1(yellow) and ligand-bound Nlh1-G2 (green). (a) Ligand binding results in a large movement of loop 1, and stabilizes the loop 2. (b) Nlh1 ligand binding results in destabilizing the α1 helix and the C-terminal region of α4. (c)Nlh1 ligand binding flips Arg 58 180° across β2, directly preventing formation of a signaling helix seen in apo Nlh1 due to clashes with Arg 163.

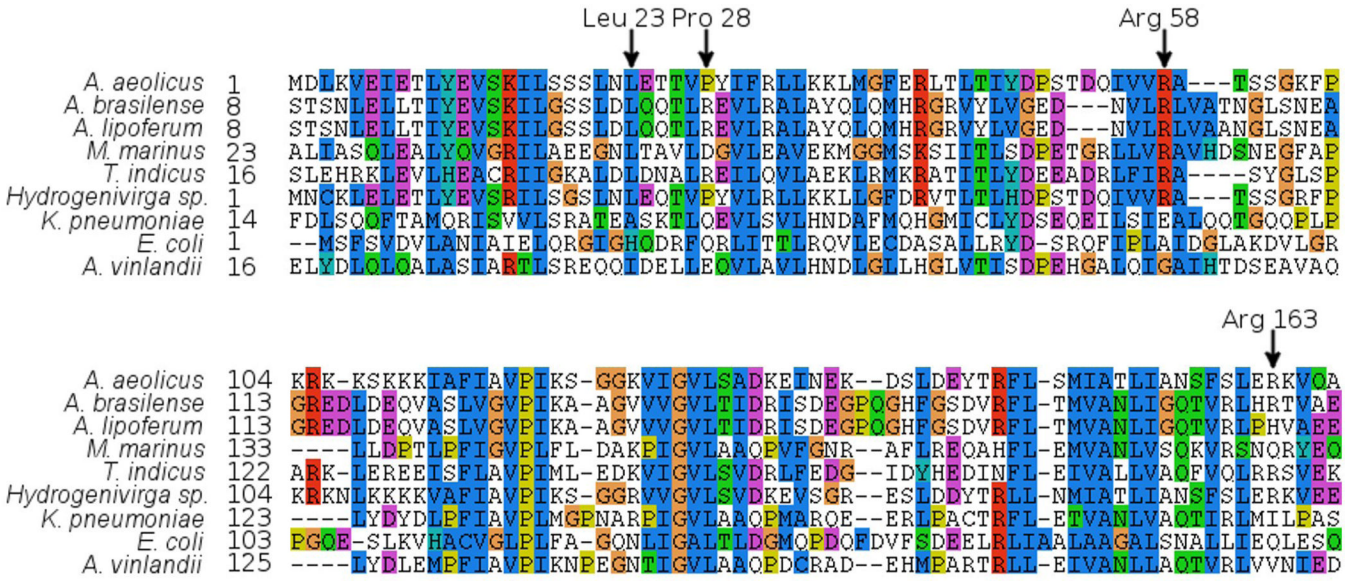

Fig. 5.

The residues which play a role in *A.a.* Nlh1 activation are present in a subset of GAF-regulated σ^{54} activators. An alignment of *A. aeolicus* Nlh1 with NifA homologs from *Azospirillum brasilense*, *Azospirillum lipoferum*,*Magne to coccus marinus*,*Thermodesulfator indicus*, *Hydrogenivirgna sp.*, *Klebsiella pneumoniae*, *E. coli*, and *Azotobacter vinelandii*.

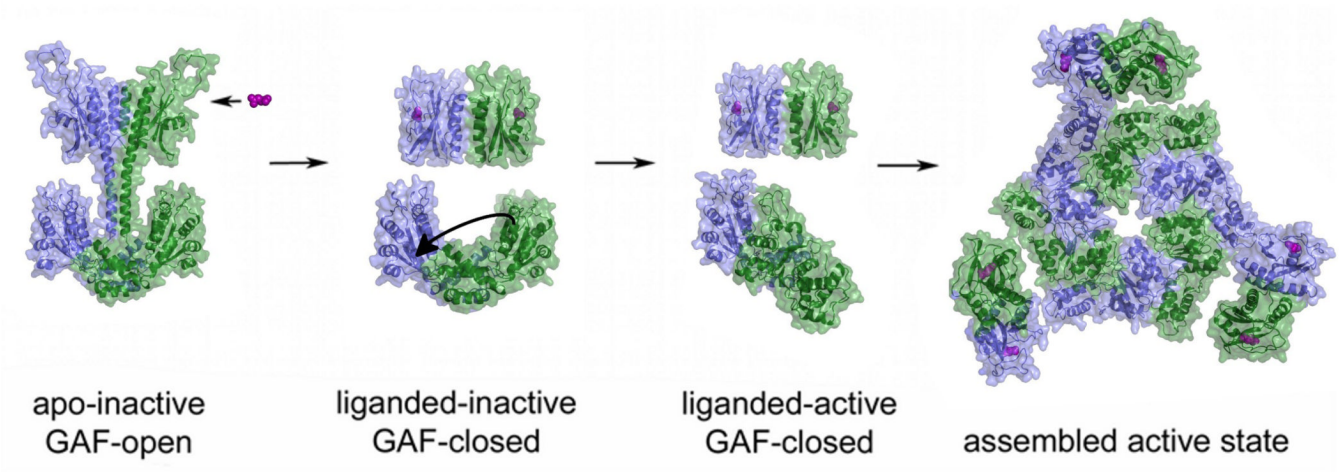

Fig. 6.

A model for GAF-regulated σ^{54} activators. Monomers are shown in blue and green. In the absence of activating ligand, GAF domains exist as off-state dimers, as in the apo-open Nlh1-G1 structure. Binding of the activating ligand (pink) results in a conformation change in the GAF domain, represented by the ligand-bound Nlh1-G2 structure where the signaling helix linking the GAF and AAA+ domains has been disrupted. Without a stable signaling helix, the GAF domain is then unable to repress assembly, and the σ^{54} activator is converted to its active hexamer state.

Table 1

Data collection and refinement statistics for Nlh2-G

	Native	SeMet Peak	Se-Met High Remote
Data collection			
Space group	$P3_221$	$P3_121$	$P3_121$
Cell dimensions			
a, *b*, *c* (Å)	65.5 × 65.5 × 128.6	64.7 × 64.7 × 63.0	64.7 × 64.7 × 63.0
α, β, γ (°)	90 × 90 × 120	90 × 90 × 120	90 × 90 × 120
Resolution (Å)	20–1.70(1.74–1.70)	63.0–2.12(2.23–2.12)	63.0–2.09 (2.20–2.09)
R_{sym}	.042(1.63)	.083 (0.974)	.081 (0.947)
$I/\sigma/$	29.29(1.55)	18.7(2.8)	18.9(2.9)
Completeness (%)	99.7(100)	100(100)	100(100)
Redundancy	10.7(10.3)	16.4(15.3)	16.3(14.9)
Refinement			
Resolution (Å)	20–1.70 A		
No. Reflections	35719		
R_{work} / R_{free}	16.26/20.65		
No. Atoms			
Protein	2601		
Water	169		
B-factors			
Protein	45.13		
Water	51.50		
R.m.s. deviations			
Bond lengths	0.007		
Bond angles (°)	0.969		

*Values in parentheses are for highest-resolution shell.

Data were collected on a single crystal for each dataset This structure has been deposited into the PDB, code **4G3V**

Table 2

Data collection and refinement statistics for Nlh1-G1 and Nlh1-G2

	Nlh1-G1 apo	Nlh1-G2ligand-bound
Data collection		
Space group	$P42_12$	$P4_32_12$
Cell dimensions		
a, *b*, *c* (Å)	119.35 × 119.35×60.29	59.8 × 59.8 × 106.24
α, β, γ (°)	90 × 90 × 90	90 × 90 × 90
Resolution (Å)	20–3.05(3.13–3.05)	20–2.70 (2.80–2.70)
R_{sym}	.0101 (1.39)	.0125 (0.752)
I/σ/	19.19 (1.7)	14.8 (2.41)
Completeness (%)	99.3 (99.8)	97.1 (97.2)
Redundancy	8.7 (8.7)	11.7(10.1)
Refinement		
Resolution (Å)	20–3.05 A	20–2.70
No. Reflections	8681	5539
R_{work} / R_{free}	20.74/24.22	20.42/23.70
No. Atoms		
Protein	2307	1086
Water	3	18
B-factors		
Protein	91.7	53.82
Water	57.7	38.37
R.m.s. deviations		
Bond lengths (Å)	0.002	0.016
Bond angles (°)	0.582	1.653

*Values in parentheses are for highest-resolution shell.

Data were collected on a single crystal for each dataset These structures have been deposited into the PDB, codes **4G3K** (G1) and **4G3W**(G2).