Please cite this article in press as: Kim and Porter, Functional and Regulatory Roles of Fold-Switching Proteins, Structure (2020), https://doi.org/10.1016/j.str.2020.10.006

Structure

CellPress

Minireview

Functional and Regulatory Roles of Fold-Switching Proteins

Allen K. Kim[1,2] and Lauren L. Porter[1,2,3,*]
[1]National Library of Medicine, National Institutes of Health, Bethesda, MD 20894, USA
[2]National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD 20892, USA
[3]Lead Contact
*Correspondence: lauren.porter@nih.gov
https://doi.org/10.1016/j.str.2020.10.006

SUMMARY

Fold-switching proteins respond to cellular stimuli by remodeling their secondary structures and changing their functions. Whereas several previous reviews have focused on various structural, physical-chemical, and evolutionary aspects of this newly emerging class of proteins, this minireview focuses on how fold switching modulates protein function and regulates biological processes. It first compares and contrasts fold switchers with other known types of proteins. Second, it presents examples of how various proteins can change their functions through fold switching. Third, it demonstrates that fold switchers can regulate biological processes by discussing two proteins, RfaH and KaiB, whose dramatic secondary structure remodeling events directly affect gene expression and a circadian clock, respectively. Finally, this minireview discusses how the field of protein fold switching might advance.

INTRODUCTION

Protein structures are dynamic (Kay, 1998). At a minimum, highly stable proteins undergo thermal fluctuations in their loops, linkers, and termini. These picosecond-to-microsecond fluctuations include functionally relevant conformational changes and sampling of active-site loops (Henzler-Wildman et al., 2007). Thus, protein dynamics introduce heterogeneity into the structures of even the most stable folded proteins.

The structural heterogeneity of naturally occurring proteins spans a wide spectrum (Kulkarni et al., 2018; van der Lee et al., 2014). On one extreme are single-fold proteins that exist in one of two states: either fully folded or unfolded (Figure 1). When subjected to various physiological perturbations, such as changes in pH or salt concentration, these two-state single-fold proteins often—but not always (Wand et al., 1986)—maintain their secondary and tertiary structures (Li and Woodward, 1999). This structural stability frequently allows them to maintain their functions in the wake of changing cellular conditions, although different environments can affect protein stability and activity (Monteith et al., 2015; Motlagh et al., 2014). On the other extreme of the structural heterogeneity spectrum are intrinsically disordered proteins (IDPs) (Wright and Dyson, 2015). These flexible proteins lack a stable structure in isolation and thus assume many different forms that bind disparate partners (Oldfield et al., 2008) and couple diverse biological processes (Wright and Dyson, 2015). Furthermore, the highly solvent-exposed amino acids in IDPs are readily accessible to post-translational modifications that regulate both their structural ensembles and their functions (Mylona et al., 2016).

Several types of proteins that do not switch folds fall between the extremes of the structural heterogeneity spectrum (Figure 1). For example, some proteins combine folded domains with intrinsically disordered regions (IDRs), such as protein kinase A (PKA), which contains both a phosphorylating folded domain and an IDR. Whereas the functional role of the folded domain is well established, the role of the IDR was unclear until recently (Smith et al., 2013). A combination of electron microscopy and functional assays showed that this IDR binds to a tethering protein that localizes kinase activity within a "radius of action." Consistent with this observation, altering the length of the IDR affected the enzymatic efficiency of the folded domain *in situ*. Other proteins form molten globules with native-like secondary structure that lacks the hydrophobic packing of a fully folded protein. Crammer, a cathepsin inhibitor involved in long-term memory formation in *D. melanogaster*, likely exists in the low-pH environment of the lysosome *in situ*. NMR studies suggest that at lysosomal pH, Crammer exists in a monomeric molten globule state that binds cathepsin before it assumes a stable fold (Tseng et al., 2012). Still other single-fold proteins exist as partially folded intermediates, like CylR2 (Jaremko et al., 2013), whose cryogenically cooled conformation unfolds locally. Similar local unfolding events appear to enable amyloid formation (Karamanos et al., 2019).

Fold-switching proteins also fall between the extremes of the structural heterogeneity spectrum (Figure 1) (Bryan and Orban, 2010; Dishman and Volkman, 2018; Goodchild et al., 2011; Kulkarni et al., 2018; Murzin, 2008). Unlike single-fold proteins, fold switchers remodel their secondary structures in response to cellular stimuli, and unlike isolated IDPs, they assume stable folds. This combination of environmental

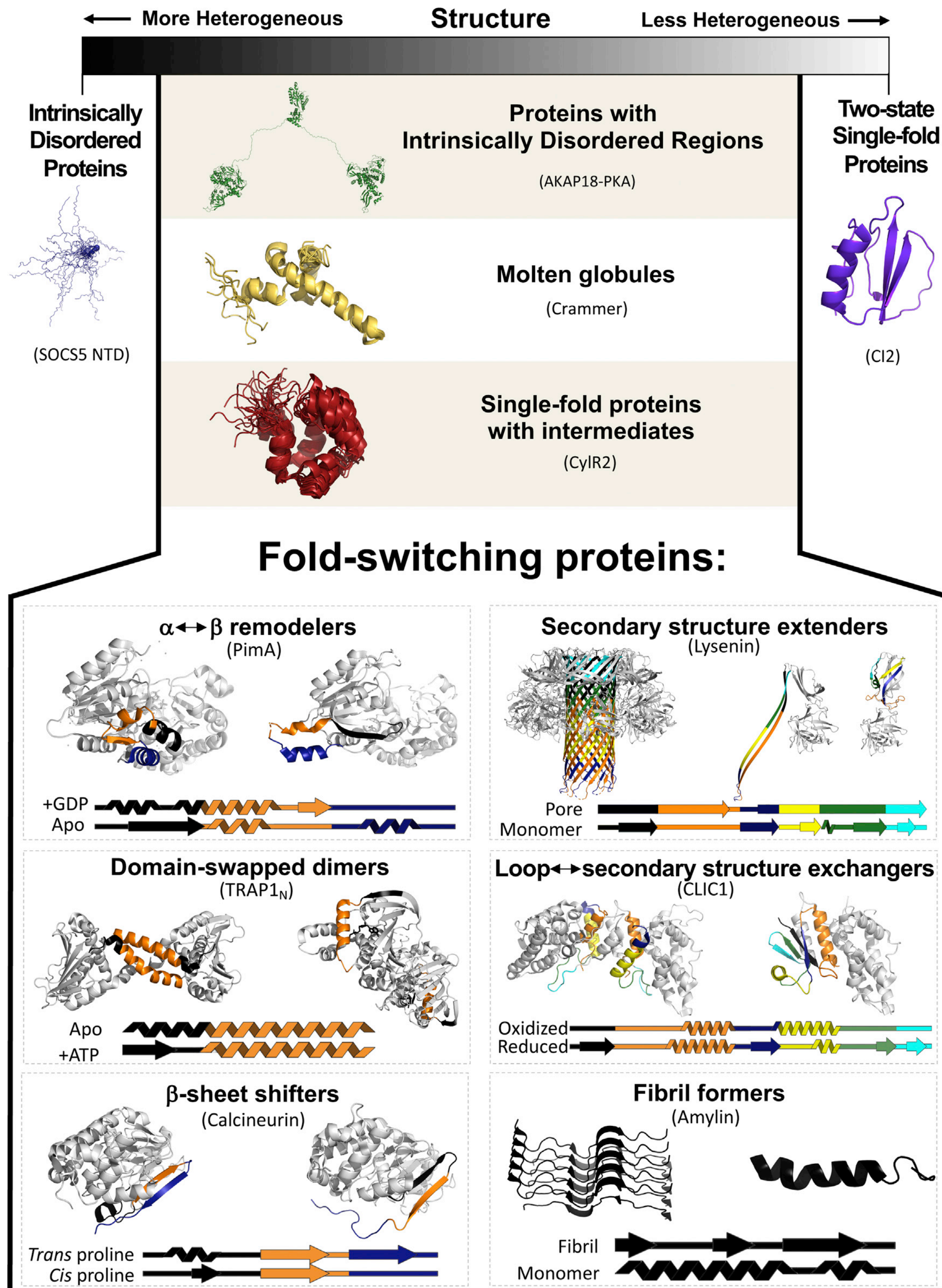

Figure 1. A Structural Heterogeneity Scale for Proteins

IDPs/two-state single-fold proteins are the most/least structurally heterogeneous and lie on the left/right extremes. Proteins between the extremes are not ordered by where they fall on the spectrum. Color choices for single-fold proteins and IDPs are not significant, but fold-switching proteins are colored in segments from N to C terminus as follows: black, orange, dark blue, yellow, green, and cyan; gray regions do not switch folds. Top/bottom secondary structure diagrams correspond to left/right protein structures. PDB IDs: SOCS5, 2N34; CI2, 1YPC; AKAP18-PKA, 3J4Q; Crammer, 2L95; CylR2, 2LYS; PimA, 4N9W and 4NC9;

(legend continued on next page)

sensitivity and structural stability allows cellular conditions to modulate the functions of fold switchers (Porter and Looger, 2018).

Secondary structure remodeling is the key point of difference between fold switchers and proteins with IDRs, molten globules, and single-fold proteins with intermediates. Proteins with IDRs can take on many different unstable conformations. By contrast, fold switchers often take different stable secondary structure configurations under different conditions. Molten globules adopt specific secondary structures but lack stable tertiary interactions. Distinctively, fold switchers can exist in two (or more) different secondary structure configurations with stabilizing tertiary interactions. Furthermore, unlike the locally unfolded intermediates of single-fold proteins, the equilibrium intermediates of fold switchers can fold into two different forms of secondary structure (Dingfelder et al., 2018).

Fold switchers also differ somewhat from metamorphic proteins (Murzin, 2008), the term first used to describe proteins that can radically change their structures. The main difference is that the secondary structure remodeling events of fold switchers can be either reversible or irreversible. In contrast, metamorphic proteins are defined to "undergo reversible conformational changes" (Murzin, 2008). Thus, we use the terms "fold-switching proteins" and "fold switchers" to encompass both reversible and irreversible secondary structure remodeling. Instances of both are discussed in this minireview, and both are functionally relevant.

Fold-Switching Events Foster Functional Changes

Currently, there are nearly 100 literature-validated examples of fold-switching proteins that can remodel their secondary structures in six different ways (Figure 1 and Table S1) with numerous functional consequences. The following secondary-structure-based classification differs from the oligomerization-based classification presented previously (Porter and Looger, 2018).

- Over 40% of known fold switchers change their conformation by reversibly remodeling α helices to β strands or vice versa. One recent example is PimA (Giganti et al., 2015), a membrane-associated bacterial glycosyltransferase involved in the biosynthesis of phosphatidylinositol mannosides. PimA's N-terminal domain (NTD) is flexible and can adopt either a closed or an open state. Its active substrate-bound closed state assumes a secondary structure configuration different from its apo open state, which may foster membrane association.
- Nearly 25% of proteins switch folds by extending their existing secondary structures. This includes the β-pore-forming toxin lysenin (Podobnik et al., 2016). In its monomeric pre-pore state, lysenin's fold-switching region adopts a compact structure with four β-strands packed against its neighboring domain. Upon membrane binding, however, this region re-forms into a long β hairpin that extends away from the neighboring domain (Figure 1), and lysenin forms a nonameric pore. More than half of secondary structure extenders form membrane-inserted pores and include viral fusion proteins, eukaryotic α-helical toxins, and other β-pores.
- Approximately 16% of fold switchers form domain-swapped dimers, which have been characterized extensively (Liu and Eisenberg, 2002). Domain-swapped fold switchers are a subset of domain-swapped dimers that remodel their secondary structures. One recent example is the NTD of human mitochondrial Hsp90, $TRAP1_N$ (Sung et al., 2016). Unliganded $TRAP1_N$ forms a domain-swapped coiled coil that allows the dimer to close partially. A recent model suggests that adding ATP to this partially closed conformation allows TRAP1 to form a fully closed domain-swapped β-strand conformation, fostering chaperone activity (Sung et al., 2016). $TRAP1_N$ can also bind ATP in an open conformation with no domain swapping; the functional relevance of this open structure is still under investigation.
- Nearly 11% of fold switchers change conformation by exchanging loops for regular secondary structure and vice versa. For example, CLIC1 is a human protein that adopts different folds and functions under oxidizing and reducing conditions (Littler et al., 2004). Oxidized CLIC1 forms a membrane-associated helical dimer believed to function as a chloride channel. In contrast, reduced CLIC1 folds into a soluble monomer with glutathione reductase activity (Al Khamici et al., 2015).
- Three percent of fold switchers have different β-sheet registers. Human calcineurin (CNA) appears to switch register through a *trans* ↔ *cis* proline isomerization (Guasch et al., 2015). Its two conformations have different active-site organizations that may explain their broad substrate specificities. Furthermore, *cis*-CNA appears unable to bind and activate the transcription factors activated by *trans*-CNA.
- The remaining 6% of fold switchers form fibrils. Despite their irreversibility, we include fibrils whose secondary structures differ from those of their unaggregated precursors. Often—but not always (Morris et al., 2019)—fibrils are associated with disease. For example, over 90% of type 2 diabetes patients have pathogenic deposits of amylin (Cao et al., 2020), a pancreatic peptide hormone that slows digestion and promotes satiety.

Fold-switching proteins occur in diverse life forms: archaea, bacteria, viruses, and lower and higher eukaryotes. To date, they are reported to perform 30 different functions. The majority (58%) function intracellularly, 10% function extracellularly, and 30% have at least one membrane-associated conformation, many of which are viral proteins associated with membrane fusion, a field with a rich history (White et al., 2008). The remaining 2% comprise a viral capsid protein and a protein of unknown function.

We now highlight two proteins that regulate biological processes through fold switching. To our knowledge, these are the only fold switchers whose regulatory roles have been

lysenin, 5EC5 (nonamer and monomer—shown for comparison with pre-pore monomer) and 3ZXG; $TRAP1_N$, 5F3K and 5F5R; CLIC1, 1RK4 and 1K0O; calcineurin, 5C1V_A and 5C1V_B; Amylin, 6VW2 and 2KB8. Since there is some overlap between categories, the structural transitions were classified in the following order: (1) domain-swapped dimers, (2) fibril formers, (3) α ↔ β remodelers, (4) secondary-structure extenders, (5) β-sheet shifters, (6) loop ↔ secondary structure exchanges. Protein structures in all figures were made with PyMOL.

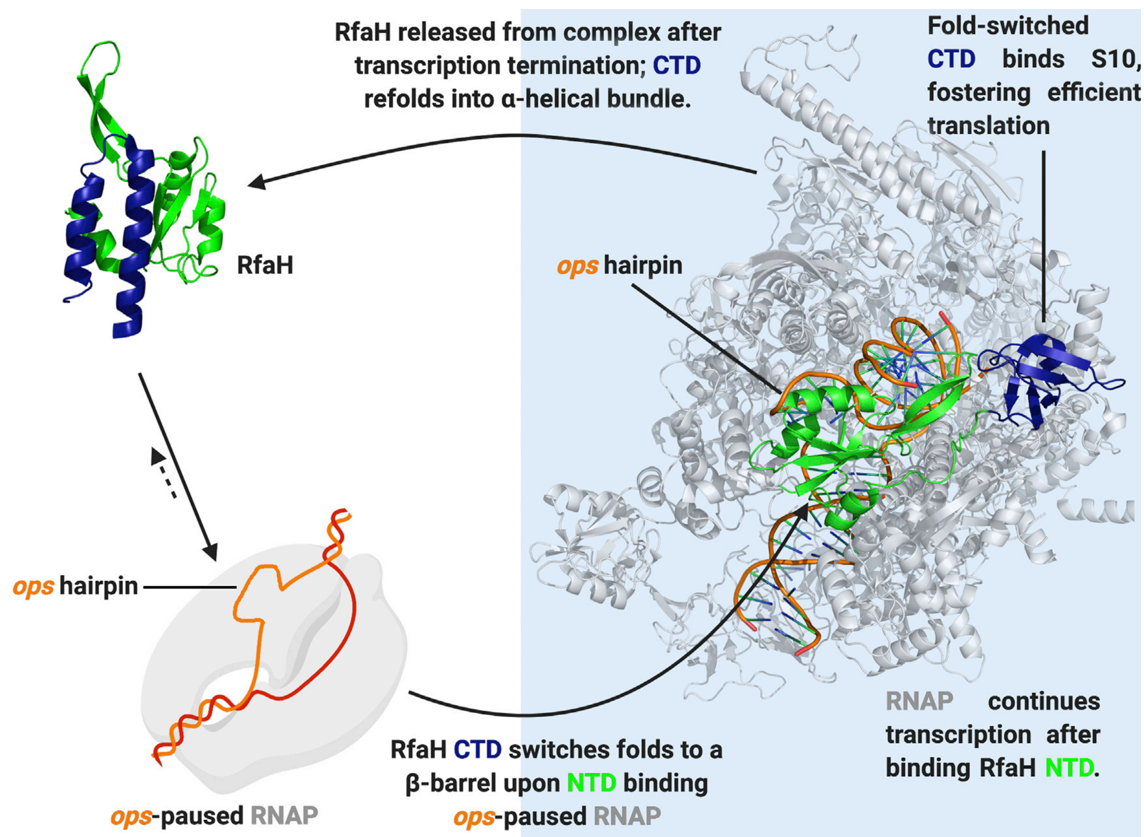

Figure 2. An Abbreviated Functional Cycle for RfaH

Autoinhibited RfaH (top left, PDB: 2OUG) NTD is colored green; its fold-switching CTD is dark blue. RfaH binds RNAP + *ops* (bottom left) with high affinity. Upon binding RNAP + *ops*, RfaH's CTD switches folds from an α-helical bundle to a β-barrel able to bind S10, an integral component of the 30S ribosomal subunit, fostering efficient translation. RfaH's NTD (green) binds both RNAP and *ops* DNA (PDB: 6C6S). White/blue backgrounds separate activities associated with autoinhibited/fold-switched RfaH, respectively. Created with BioRender.com.

characterized extensively over the past 5 years. In both cases, this characterization has combined major structural advances with diverse functional and cellular assays, placing the phenomenon of fold switching into a rich biological context. The first regulatory fold switcher is RfaH, whose C-terminal domain (CTD) switches from an α-helical turn to a β- barrel, regulating transcription and translation of specific genes in *E. coli*. The second is KaiB, whose fold-switched form fosters the 24-h periodicity of a cyanobacterial circadian clock.

Regulation through Fold Switching I: RfaH in Bacterial Gene Expression

NusG/Spt5 is the only family of transcription factors known to be conserved in all domains of life (Werner, 2012). These proteins colocalize with elongating RNA polymerases (RNAPs) and generally reduce RNAP pausing. All known NusG/Spt5 orthologs contain a NusG N-terminal (NGN) domain that binds RNAP (Kang et al., 2018). These NGN domains are covalently linked to one or more Kyprides-Ouzounis-Woese (KOW) domains that couple transcription elongation to diverse cellular processes such as translation, transcription termination, and chromatin remodeling (Werner, 2012). Furthermore, specialized NusG/Spt5 paralogs have been identified in both bacteria and eukaryotes (Kang et al., 2018).

RfaH is a NusG paralog found in *E. coli* and other bacterial strains. Discovered genetically in the 1970s (Beutin and Achtman, 1979), RfaH is required for the expression of secreted macromolecules, including the toxin hemolysin (Leeds and Welch, 1996). RfaH functions only in the presence of a short DNA consensus sequence called *ops* (operon polarity suppressor) (Artsimovitch and Landick, 2002). The non-template DNA *ops* sequence forms a hairpin that RfaH binds specifically (Zuber et al., 2018). This *ops*-site specificity distinguishes RfaH from NusG, which colocalizes with RNAP at most genes (Mooney et al., 2009), not just those containing *ops*.

The first full-length atomic-resolution structures of NusG (Steiner et al., 2002) suggested that bacterial NusGs and their paralogs would fold into a two-domain architecture: an N-terminal NGN fold and a C-terminal β-barrel. Surprisingly, when the full-length structure of RfaH was first solved, its CTD folded into an α-helical hairpin instead of the expected β-barrel (Belogurov et al., 2007) (Figure 2). This structure suggested that RfaH's CTD masks the RNAP-binding site on its NTD, but the structural details of how this RNAP-binding site becomes exposed remained an open question.

A breakthrough came 4 years later, when Burmann et al. showed that RfaH's CTD dramatically switches folds from an NTD-bound α-helical hairpin to a β-barrel that binds the S10

ribosomal subunit instead of the NTD (Burmann et al., 2012) (Figure 2). Thus, this dramatic, reversible (Zuber et al., 2019) fold-switching event couples *ops*-site-specific transcription elongation with translation. Key evidence corroborating this finding originated from an earlier study where a tobacco etch virus (TEV) protease cleavage site was inserted in the linker between RfaH's NTD and CTD (Belogurov et al., 2007). Adding trace amounts of TEV protease allowed the chemical shifts of the same RfaH sample to be monitored by solution NMR both before and after cleavage. Strikingly, after a 42-h TEV incubation, RfaH's chemical shifts changed from their initial α-helical pattern into a β-barrel pattern, demonstrating that RfaH's CTD switches folds upon NTD dissociation (Burmann et al., 2012).

CTD fold switching upon NTD dissociation raised the possibility that a trigger molecule could out-compete NTD:CTD interactions. RfaH's two known binding partners, RNAP and *ops* DNA, were likely suspects. Cellular assays showed that RfaH:RNAP binding is weak, however, and requires supplementation of cellular lysate to activate RfaH (Artsimovitch and Landick, 2002). Furthermore, autoinhibited RfaH binds *ops* DNA (Zuber et al., 2018). Together, these results show that neither RNAP nor *ops* DNA alone can activate RfaH *in vitro*. The resolution came from a recent NMR study showing that RfaH switches folds by binding both RNAP and *ops* DNA together (Zuber et al., 2019).

RfaH provides insight into how fold switching can directly regulate important cellular processes. In its autoinhibited conformation, the RNAP binding site on RfaH's NTD is masked by its C-terminal helical bundle domain (Figure 2). Upon encountering *ops*-paused RNAP, RfaH likely forms a transient complex with both RNAP and the non-template *ops* hairpin (Zuber et al., 2019), weakening its NTD:CTD interface. Thus, while RfaH's NTD maintains interactions with both RNAP and *ops*, its CTD refolds into a β-barrel (Figure 2) that recruits S10 and enhances translation. The requirement for *ops* binding is important as it ensures that RfaH enhances transcriptional processivity only at specific operons. This site-specific transcription is coupled with robust translation through the binding of S10 to the fold-switched CTD.

Regulation through Fold Switching II: KaiB in a Cyanobacterial Circadian Clock

Circadian clocks coordinate biological processes with the earth's 24-h cycle in all domains of life (Edgar et al., 2012). The circadian clock of the cyanobacterium *Synechococcus elongatus* orchestrates the expression of most of its genes (Vijayan et al., 2009), enhancing its fitness (Woelfle et al., 2004). This clock was initially believed to be regulated by a transcription-translation feedback loop, the most common molecular mechanism that drives circadian rhythm (Partch, 2020). In 2005, however, Kondo and colleagues demonstrated that this cyanobacterial circadian clock could be reconstituted *in vitro* with just three proteins, Mg^{2+}, and ATP (Nakajima et al., 2005). This elegant finding demonstrated that the clock was essentially regulated by a post-translational oscillator (PTO).

This cyanobacterial PTO is driven by a phosphorylation cycle involving three proteins: KaiA, KaiB, and KaiC. Of the three proteins, KaiC is the only enzyme, and its phosphorylation/dephosphorylation cycle has a ~24-h period. KaiC is composed of two ATPase domains: CI, which has demonstrated ATPase activity, and CII, which has autokinase and phosphotransferase activities (Tseng et al., 2017). These domains, connected by a short linker, form a hexameric ring (Figure 3). Compared with well-known motor proteins, which consume 10^3–10^7 ATP/day, CI activity is extremely slow, consuming only 10–12 ATP/day (Abe et al., 2015). Thus, CI's sluggish catalysis likely fosters the 24-h periodicity of KaiC's phosphorylation cycle.

Nevertheless, the slow ATPase activity of KaiC's CI domain does not fully explain the oscillatory behavior of this system. Sequential autophosphorylation of serine-431 (S431) and threonine-432 (T432) in KaiC's CII domain is a critical component of the oscillation circuit (Rust et al., 2007) and requires both KaiA and KaiB interactions. Specifically, KaiA activates the autophosphorylation activity of KaiC by binding the C-terminal extension of its CII domain, known as the A loop (Pattanayek and Egli, 2015). This binding triggers kinase activity that sequentially phosphorylates T432 and then S431. After S431 is phosphorylated, KaiB represses KaiA activity by sequestering it from the A loop (Phong et al., 2013), but the structural details of this process took years to elucidate.

The structural barrier was broken 10 years later, when LiWang and colleagues found that KaiB switches folds (Chang et al., 2015). In further detail, KaiB can fold into both an inactive tetrameric form ($KaiB^4$) and an active fold-switched monomeric form ($KaiB^{fs}$, Figure 3). Endogenous $KaiB^{fs}$ binds the CI domain of KaiC (Tseng et al., 2017) and exists in such low abundance that it cannot be detected by solution NMR. Thus, LiWang and colleagues engineered KaiB variants with NMR-detectable populations of the fold-switched conformation. Importantly, none of the genetically engineered strains of *S. elongatus* expressing these fold-switch-enhanced variants maintained their circadian rhythms (Chang et al., 2015). Since only $KaiB^{fs}$ is active (Tseng et al., 2017), this phenotypic readout suggests that the low abundance of $KaiB^{fs}$ probably plays an important role in maintaining the cyanobacterial circadian clock. The mechanistic details of this role are still up for debate, however (Chang et al., 2015; Koda and Saito, 2020).

Interestingly, $KaiB^{fs}$ functions beyond KaiC binding and KaiA sequestration (Figure 3). One of the signal transduction components downstream of the KaiA-KaiB-KaiC system is SasA, a histidine kinase that phosphorylates RpaA, a response regulator that activates transcription of the *kaiBC* genes (Takai et al., 2006). $KaiB^{fs}$ competes directly with SasA to bind CI (Tseng et al., 2014), likely downregulating SasA's activity at night (Tseng et al., 2017). Furthermore, $KaiB^{fs}$, when bound to KaiC, recruits not only KaiA but also CikA, a phosphatase that dephosphorylates RpaA, reducing *kaiBC* expression (Figure 3) (Gutu and O'shea, 2013). A structure-preserving CikA mutation that disrupted $KaiB^{fs}$:CikA binding showed phenotypic defects similar to *cikA* knockout strains *in vivo*, emphasizing that $KaiB^{fs}$ plays an important role in CikA activation (Tseng et al., 2017).

Similar to RfaH, KaiB fold switching regulates the cyanobacterial circadian clock in a few ways (Figure 3). First, $KaiB^{fs}$ sequesters KaiA only when KaiC's S431 is phosphorylated, allowing KaiC to be dephosphorylated and helping to maintain the phosphorylation cycle that underlies the cyanobacterial circadian clock. Second, $KaiB^{fs}$ downregulates SasA's phosphorylation activity by competing directly for CI binding. Third, $KaiB^{fs}$ antagonistically decreases the effect of SasA by recruiting CikA, which dephosphorylates SasA-activated RpaA. Finally, the relative low

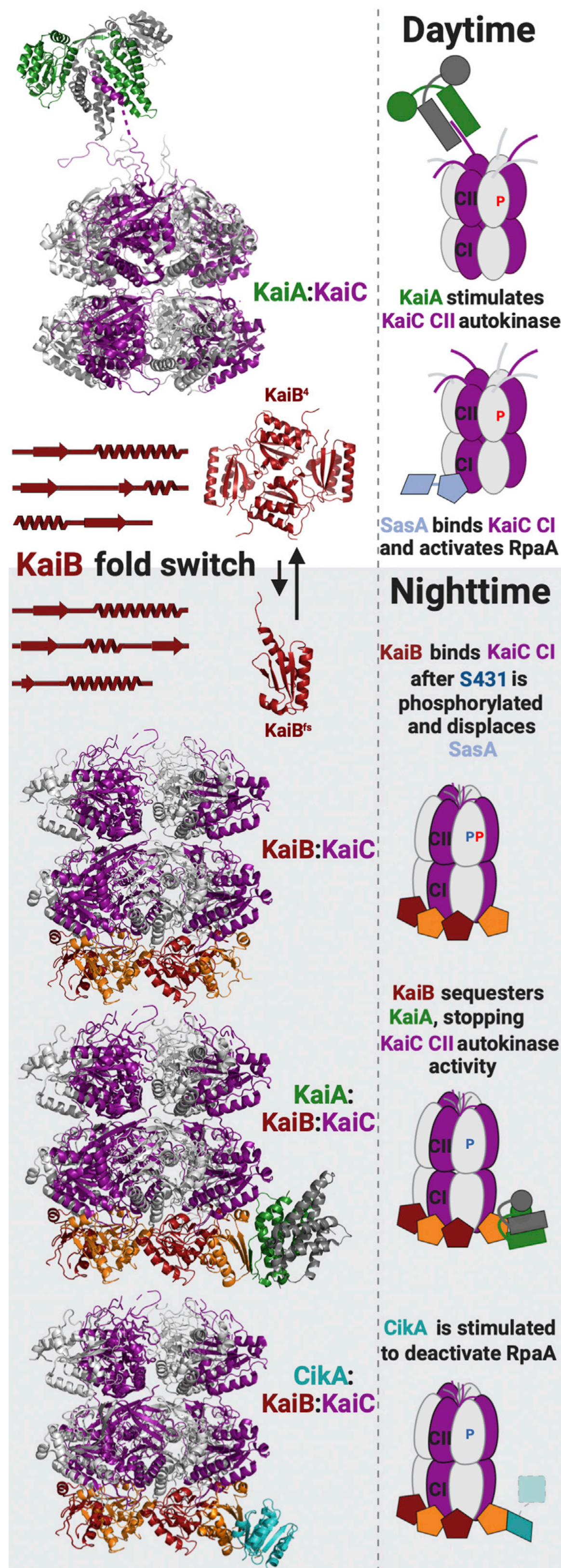

Figure 3. Complexes Associated with the Cyanobacterial Circadian Clock

Cartoons of each complex are shown to the right of the dotted line, and their corresponding molecular structures are shown to its left, except SasA:KaiC (not available). White/gray backgrounds separate daytime/nighttime complexes. Secondary structure diagrams of KaiB's two forms correspond to their experimentally determined three-dimensional structures. $KaiB^4$ refers to its inactive tetrameric form, while $KaiB^{fs}$ refers to its active fold-switched form. KaiA and KaiC form homo-oligomers, colored green/dark gray, purple/light gray to show different subunits. PDB IDs: KaiA:KaiC, 5C5E and 3DVL, respectively; $KaiB^4$, 2QKE; $KaiB^{fs}$, 5JYT; $KaiB^{fs}$:KaiC, 5JWQ. KaiA:-$KaiB^{fs}$:KaiC/CikA:$KaiB^{fs}$:KaiC complexes were modeled in PyMOL by superimposing a $KaiB^{fs}$ subunit from 5JWR/5JYV with a 5JWQ chain. Colors of $KaiB^{fs}$ subunits bound to KaiC alternate red/orange. Only the pseudo-receiver domain on CikA is shown in the molecular structure; the missing cartoon CikA domain is outlined with a dotted line and shown at 70% transparency to illustrate this. Phosphorylated T432/S431 are represented by red/blue P's. CI and CII domains are labeled on each cartoon. Cartoons are based on Figure 1 from Tseng et al., 2017. Created with Biorender.com.

abundance of $KaiB^{fs}$ appears to foster the periodicity of the cyanobacterial circadian clock, linking dramatic structural changes to a specific phenotypic readout.

Conclusions and Future Directions

Fold-switching proteins remodel their secondary structures in response to cellular stimuli, modulating their functions and regulating biological processes. So far, we have discussed what fold switching is, what some of its functional roles are, and how it can regulate two diverse biological processes. To conclude, we will discuss how the field of fold switching might be advanced.

One major barrier to this field is the relatively low number of documented fold switchers. Indeed, to our knowledge, all known fold switchers were stumbled upon by chance. Furthermore, the only high-throughput predictive method requires a solved protein structure (Mishra et al., 2019), greatly constraining its use.

Nevertheless, several computational and experimental approaches hold promise for the prediction and identification of new fold switchers. First, coevolutionary methods have been highly successful at predicting the structures of single-fold proteins (Marks et al., 2012; Tian et al., 2018), and two recent successes suggest that it may be possible to predict fold switchers from their genomic sequences by combining coevolutionary analysis with physically based modeling such as Monte Carlo or molecular dynamics simulations. The most recent success was achieved by Tian and Best, who combined residue coevolution with Monte Carlo simulations to accurately predict the structures of protein G variants with up to 98% sequence identity but very different folds (Tian and Best, 2020). In addition, more subtle conformational changes in proteins, such as rotational motions, were predicted with reasonable success by combining coevolutionary methods with discrete molecular dynamics simulations (Sfriso et al., 2016).

New fold switchers could also be identified through data mining. We recently identified nearly 100 fold-switching proteins by searching the Protein DataBank (PDB) for protein structures with identical (or nearly identical) sequences but different folds (Porter and Looger, 2018). We also used text searches of published abstracts to find 15 additional proteins with one solved structure and literature evidence for an alternative conformation (Porter and Looger, 2018). Periodically repeating similar efforts may uncover new fold switchers. In addition, discrepancies between experimentally determined protein structures and homology-based secondary structure predictions are a salient feature of fold switchers (Mishra et al., 2019). These could potentially be

used to identify fold-switching proteins with one experimentally determined conformation whose secondary structure differs from its homology-based secondary structure prediction.

Furthermore, it may be possible to discover new fold switchers by using the computational methods discussed previously to search protein families with members already known to switch folds. For example, eukaryotic RfaH/NusG orthologs couple transcription with various biological processes (Kang et al., 2018). It has been hypothesized that some of these orthologs might also switch folds (Burmann et al., 2012). This hypothesis (fold switching in eukaryotic transcriptional regulators) has precedent since the yeast transcriptional regulator *MATα2* switches folds (Tan and Richmond, 1998). A couple of other unrelated transcriptional regulators have also been found to switch folds. PrgX, which regulates transcription of virulent bacterial genes, is activated by binding an intracellular peptide pheromone, cCF10. Binding this pheromone appears to trigger the remodeling of PrgX's C-terminal α-helix into a β-hairpin, breaking the tetrameric interface of its apo form and allowing it to form active dimers (Kozlowicz et al., 2006; Shi et al., 2005). In addition, transcriptional repressors from the Cro and cI families share a common ancestor and have nevertheless evolved to assume different folds (Kumirov et al., 2018; Roessler et al., 2008). In sum, these results suggest that there are likely to be other undiscovered fold-switching transcriptional regulators, making them a promising target for efforts to identify more fold switchers.

NMR and cryoelectron microscopy (cryo-EM) are examples of experimental methods that hold promise for the discovery and characterization of fold-switching proteins. NMR played a critical role in the discovery of lymphotactin's ability to switch folds by revealing that it assumes two conformations at physiological temperatures (Tuinstra et al., 2008). NMR has also been used to structurally and functionally characterize RfaH, KaiB, and MinE, another fold switcher (Cai et al., 2019; Chang et al., 2015; Zuber et al., 2019). In addition, cryo-EM is likely to elucidate other fold switchers whose alternative conformations are accessed by forming large complexes and assemblies. Indeed, it has already played a critical role in uncovering new fold switchers, including β toxins (Iacovache et al., 2016), viral fusion proteins (Gui et al., 2017), and protein fibrils (Cao et al., 2020), as well as the structures of the large complexes in Figures 2 and 3 (Kang et al., 2018; Tseng et al., 2017). Both NMR and cryo-EM, as well as X-ray crystallography (Wei et al., 2020), could also be used to validate computational predictions of fold switchers.

As more fold switchers are discovered, we anticipate that their biological scope will expand. For example, a single cancer-associated mutation changes a native α-helix in MEF2B to a β-strand (Lei et al., 2018), indicating that single-nucleotide polymorphisms could induce fold switching and lead to disease. Furthermore, post-translational modifications can lead to dramatic functionally relevant structural changes in IDPs (Bah et al., 2015), suggesting that the same might be true for fold switchers.

As demonstrated above, multidisciplinary approaches are necessary to elucidate the biological relevance of fold switching. Indeed, what makes the RfaH and KaiB systems especially interesting is how they correlate fold switching with cellular assays and phenotypic readouts. These integrative approaches put the striking conformational transitions of fold switchers into a meaningful biological context. With the increasing accessibility of new biological methods and cross-disciplinary collaborations on the rise, this is an exciting time for the field.

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.str.2020.10.006.

ACKNOWLEDGMENTS

We thank Carolyn Ott, Loren Looger, David Landsman, and George Rose for critically reading the manuscript and Juliette Lecomte and Julie Forman-Kay for helpful discussion. This work was supported by the Intramural Research Program of the National Library of Medicine, National Institutes of Health.

AUTHOR CONTRIBUTIONS

Conceptualization, L.L.P.; Formal Analysis, A.K.K. and L.L.P.; Writing – Original Draft, A.K.K. and L.L.P.; Writing – Review & Editing, L.L.P. and A.K.K.; Visualization, L.L.P. and A.K.K.; Funding Acquisition, L.L.P.; Supervision, L.L.P.

DECLARATION OF INTERESTS

The authors declare no competing interests.

SUPPORTING CITATIONS

The following references appear in the supplemental information: Bryan and Orban (2010); Dishman and Volkman (2018); Goodchild et al. (2011); Kulkarni et al. (2018); Murzin (2008).

REFERENCES

Abe, J., Hiyama, T.B., Mukaiyama, A., Son, S., Mori, T., Saito, S., Osako, M., Wolanin, J., Yamashita, E., Kondo, T., and Akiyama, S. (2015). Circadian rhythms. Atomic-scale origins of slowness in the cyanobacterial circadian clock. Science *349*, 312–316.

Al Khamici, H., Brown, L.J., Hossain, K.R., Hudson, A.L., Sinclair-Burton, A.A., Ng, J.P., Daniel, E.L., Hare, J.E., Cornell, B.A., Curmi, P.M., et al. (2015). Members of the chloride intracellular ion channel protein family demonstrate glutaredoxin-like enzymatic activity. PLoS One *10*, e115699.

Artsimovitch, I., and Landick, R. (2002). The transcriptional regulator RfaH stimulates RNA chain synthesis after recruitment to elongation complexes by the exposed nontemplate DNA strand. Cell *109*, 193–203.

Bah, A., Vernon, R.M., Siddiqui, Z., Krzeminski, M., Muhandiram, R., Zhao, C., Sonenberg, N., Kay, L.E., and Forman-Kay, J.D. (2015). Folding of an intrinsically disordered protein by phosphorylation as a regulatory switch. Nature *519*, 106–109.

Belogurov, G.A., Vassylyeva, M.N., Svetlov, V., Klyuyev, S., Grishin, N.V., Vassylyev, D.G., and Artsimovitch, I. (2007). Structural basis for converting a general transcription factor into an operon-specific virulence regulator. Mol. Cell *26*, 117–129.

Beutin, L., and Achtman, M. (1979). Two Escherichia coli chromosomal cistrons, sfrA and sfrB, which are needed for expression of F factor tra functions. J. Bacteriol. *139*, 730–737.

Bryan, P.N., and Orban, J. (2010). Proteins that switch folds. Curr. Opin. Struct. Biol. *20*, 482–488.

Burmann, B.M., Knauer, S.H., Sevostyanova, A., Schweimer, K., Mooney, R.A., Landick, R., Artsimovitch, I., and Rosch, P. (2012). An alpha helix to beta barrel domain switch transforms the transcription factor RfaH into a translation factor. Cell *150*, 291–303.

Cai, M., Huang, Y., Shen, Y., Li, M., Mizuuchi, M., Ghirlando, R., Mizuuchi, K., and Clore, G.M. (2019). Probing transient excited states of the bacterial cell division regulator MinE by relaxation dispersion NMR spectroscopy. Proc. Natl. Acad. Sci. U S A *116*, 25446–25455.

Cao, Q., Boyer, D.R., Sawaya, M.R., Ge, P., and Eisenberg, D.S. (2020). Cryo-EM structure and inhibitor design of human IAPP (amylin) fibrils. Nat. Struct. Mol. Biol. *27*, 653–659.

Chang, Y.G., Cohen, S.E., Phong, C., Myers, W.K., Kim, Y.I., Tseng, R., Lin, J., Zhang, L., Boyd, J.S., Lee, Y., et al. (2015). Circadian rhythms. A protein fold switch joins the circadian oscillator to clock output in cyanobacteria. Science *349*, 324–328.

Dingfelder, F., Benke, S., Nettels, D., and Schuler, B. (2018). Mapping an equilibrium folding intermediate of the cytolytic pore toxin ClyA with single-molecule FRET. J. Phys. Chem. B *122*, 11251–11261.

Dishman, A.F., and Volkman, B.F. (2018). Unfolding the mysteries of protein metamorphosis. ACS Chem. Biol. *13*, 1438–1446.

Edgar, R.S., Green, E.W., Zhao, Y., van Ooijen, G., Olmedo, M., Qin, X., Xu, Y., Pan, M., Valekunja, U.K., Feeney, K.A., et al. (2012). Peroxiredoxins are conserved markers of circadian rhythms. Nature *485*, 459–464.

Giganti, D., Albesa-Jove, D., Urresti, S., Rodrigo-Unzueta, A., Martinez, M.A., Comino, N., Barilone, N., Bellinzoni, M., Chenal, A., Guerin, M.E., and Alzari, P.M. (2015). Secondary structure reshuffling modulates glycosyltransferase function at the membrane. Nat. Chem. Biol. *11*, 16–18.

Goodchild, S.C., Curmi, P.M.G., and Brown, L.J. (2011). Structural gymnastics of multifunctional metamorphic proteins. Biophys. Rev. *3*, 143.

Guasch, A., Aranguren-Ibanez, A., Perez-Luque, R., Aparicio, D., Martinez-Hoyer, S., Mulero, M.C., Serrano-Candelas, E., Perez-Riba, M., and Fita, I. (2015). Calcineurin undergoes a conformational switch evoked via peptidyl-prolyl isomerization. PLoS One *10*, e0134569.

Gui, M., Song, W., Zhou, H., Xu, J., Chen, S., Xiang, Y., and Wang, X. (2017). Cryo-electron microscopy structures of the SARS-CoV spike glycoprotein reveal a prerequisite conformational state for receptor binding. Cell Res. *27*, 119–129.

Gutu, A., and O'shea, E.K. (2013). Two antagonistic clock-regulated histidine kinases time the activation of circadian gene expression. Mol. Cell *50*, 288–294.

Henzler-Wildman, K.A., Lei, M., Thai, V., Kerns, S.J., Karplus, M., and Kern, D. (2007). A hierarchy of timescales in protein dynamics is linked to enzyme catalysis. Nature *450*, 913–916.

Iacovache, I., de Carlo, S., Cirauqui, N., dal Peraro, M., van der Goot, F.G., and Zuber, B. (2016). Cryo-EM structure of aerolysin variants reveals a novel protein fold and the pore-formation process. Nat. Commun. *7*, 12062.

Jaremko, M., Jaremko, L., Kim, H.Y., Cho, M.K., Schwieters, C.D., Giller, K., Becker, S., and Zweckstetter, M. (2013). Cold denaturation of a protein dimer monitored at atomic resolution. Nat. Chem. Biol. *9*, 264–270.

Kang, J.Y., Mooney, R.A., Nedialkov, Y., Saba, J., Mishanina, T.V., Artsimovitch, I., Landick, R., and Darst, S.A. (2018). Structural basis for transcript elongation control by NusG family universal regulators. Cell *173*, 1650–1662 e14.

Karamanos, T.K., Jackson, M.P., Calabrese, A.N., Goodchild, S.C., Cawood, E.E., Thompson, G.S., Kalverda, A.P., Hewitt, E.W., and Radford, S.E. (2019). Structural mapping of oligomeric intermediates in an amyloid assembly pathway. Elife *8*, e46574.

Kay, L.E. (1998). Protein dynamics from NMR. Biochem. Cell Biol. *76*, 145–152.

Koda, S.I., and Saito, S. (2020). An alternative interpretation of the slow KaiB-KaiC binding of the cyanobacterial clock proteins. Sci. Rep. *10*, 10439.

Kozlowicz, B.K., Shi, K., Gu, Z.Y., Ohlendorf, D.H., Earhart, C.A., and Dunny, G.M. (2006). Molecular basis for control of conjugation by bacterial pheromone and inhibitor peptides. Mol. Microbiol. *62*, 958–969.

Kulkarni, P., Solomon, T.L., He, Y., Chen, Y., Bryan, P.N., and Orban, J. (2018). Structural metamorphism and polymorphism in proteins on the brink of thermodynamic stability. Protein Sci. *27*, 1557–1567.

Kumirov, V.K., Dykstra, E.M., Hall, B.M., Anderson, W.J., Szyszka, T.N., and Cordes, M.H.J. (2018). Multistep mutational transformation of a protein fold through structural intermediates. Protein Sci. *27*, 1767–1779.

Leeds, J.A., and Welch, R.A. (1996). RfaH enhances elongation of Escherichia coli hlyCABD mRNA. J. Bacteriol. *178*, 1850–1857.

Lei, X., Kou, Y., Fu, Y., Rajashekar, N., Shi, H., Wu, F., Xu, J., Luo, Y., and Chen, L. (2018). The cancer mutation D83V induces an alpha-helix to beta-strand conformation switch in MEF2B. J. Mol. Biol. *430*, 1157–1172.

Li, R., and Woodward, C. (1999). The hydrogen exchange core and protein folding. Protein Sci. *8*, 1571–1590.

Littler, D.R., Harrop, S.J., Fairlie, W.D., Brown, L.J., Pankhurst, G.J., Pankhurst, S., Demaere, M.Z., CAMPBELL, T.J., Bauskin, A.R., Tonini, R., et al. (2004). The intracellular chloride ion channel protein CLIC1 undergoes a redox-controlled structural transition. J. Biol. Chem. *279*, 9298–9305.

Liu, Y., and Eisenberg, D. (2002). 3D domain swapping: as domains continue to swap. Protein Sci. *11*, 1285–1299.

Marks, D.S., Hopf, T.A., and Sander, C. (2012). Protein structure prediction from sequence variation. Nat. Biotechnol. *30*, 1072–1080.

Mishra, S., Looger, L.L., and Porter, L.L. (2019). Inaccurate secondary structure predictions often indicate protein fold switching. Protein Sci. *28*, 1487–1493.

Monteith, W.B., Cohen, R.D., Smith, A.E., Guzman-Cisneros, E., and Pielak, G.J. (2015). Quinary structure modulates protein stability in cells. Proc. Natl. Acad. Sci. U S A *112*, 1739–1742.

Mooney, R.A., Davis, S.E., Peters, J.M., Rowland, J.L., Ansari, A.Z., and Landick, R. (2009). Regulator trafficking on bacterial transcription units in vivo. Mol. Cell *33*, 97–108.

Morris, D.L., Kastner, D.W., Johnson, S., Strub, M.P., He, Y., Bleck, C.K.E., Lee, D.Y., and Tjandra, N. (2019). Humanin induces conformational changes in the apoptosis regulator BAX and sequesters it into fibers, preventing mitochondrial outer-membrane permeabilization. J. Biol. Chem. *294*, 19055–19065.

Motlagh, H.N., Wrabl, J.O., Li, J., and Hilser, V.J. (2014). The ensemble nature of allostery. Nature *508*, 331–339.

Murzin, A.G. (2008). Biochemistry. Metamorphic proteins. Science *320*, 1725–1726.

Mylona, A., Theillet, F.X., Foster, C., Cheng, T.M., Miralles, F., Bates, P.A., Selenko, P., and Treisman, R. (2016). Opposing effects of Elk-1 multisite phosphorylation shape its response to ERK activation. Science *354*, 233–237.

Nakajima, M., Imai, K., Ito, H., Nishiwaki, T., Murayama, Y., Iwasaki, H., Oyama, T., and Kondo, T. (2005). Reconstitution of circadian oscillation of cyanobacterial KaiC phosphorylation in vitro. Science *308*, 414–415.

Oldfield, C.J., Meng, J., Yang, J.Y., Yang, M.Q., Uversky, V.N., and Dunker, A.K. (2008). Flexible nets: disorder and induced fit in the associations of p53 and 14-3-3 with their partners. BMC Genomics *9* (*Suppl 1*), S1.

Partch, C.L. (2020). Orchestration of circadian timing by macromolecular protein assemblies. J. Mol. Biol. *432*, 3426–3448.

Pattanayek, R., and Egli, M. (2015). Protein-protein interactions in the cyanobacterial circadian clock: structure of KaiA dimer in complex with C-terminal KaiC peptides at 2.8 A resolution. Biochemistry *54*, 4575–4578.

Phong, C., Markson, J.S., Wilhoite, C.M., and Rust, M.J. (2013). Robust and tunable circadian rhythms from differentially sensitive catalytic domains. Proc. Natl. Acad. Sci. U S A *110*, 1124–1129.

Podobnik, M., Savory, P., Rojko, N., Kisovec, M., Wood, N., Hambley, R., Pugh, J., Wallace, E.J., Mcneill, L., Bruce, M., et al. (2016). Crystal structure of an invertebrate cytolysin pore reveals unique properties and mechanism of assembly. Nat. Commun. *7*, 11598.

Porter, L.L., and Looger, L.L. (2018). Extant fold-switching proteins are widespread. Proc. Natl. Acad. Sci. U S A *115*, 5968–5973.

Roessler, C.G., Hall, B.M., Anderson, W.J., Ingram, W.M., Roberts, S.A., Montfort, W.R., and Cordes, M.H. (2008). Transitive homology-guided structural studies lead to discovery of Cro proteins with 40% sequence identity but different folds. Proc. Natl. Acad. Sci. U S A *105*, 2343–2348.

Rust, M.J., Markson, J.S., Lane, W.S., Fisher, D.S., and O'shea, E.K. (2007). Ordered phosphorylation governs oscillation of a three-protein circadian clock. Science *318*, 809–812.

Sfriso, P., Duran-Frigola, M., Mosca, R., Emperador, A., Aloy, P., and Orozco, M. (2016). Residues coevolution guides the systematic identification of alternative functional conformations in proteins. Structure *24*, 116–126.

Shi, K., Brown, C.K., Gu, Z.Y., Kozlowicz, B.K., Dunny, G.M., Ohlendorf, D.H., and Earhart, C.A. (2005). Structure of peptide sex pheromone receptor PrgX and PrgX/pheromone complexes and regulation of conjugation in Enterococcus faecalis. Proc. Natl. Acad. Sci. U S A *102*, 18596–18601.

Smith, F.D., Reichow, S.L., Esseltine, J.L., Shi, D., Langeberg, L.K., Scott, J.D., and Gonen, T. (2013). Intrinsic Disorder within an AKAP-protein kinase A complex guides local substrate phosphorylation. Elife *2*, e01319.

Steiner, T., Kaiser, J.T., Marinkovic, S., Huber, R., and Wahl, M.C. (2002). Crystal structures of transcription factor NusG in light of its nucleic acid- and protein-binding activities. EMBO J. *21*, 4641–4653.

Sung, N., Lee, J., Kim, J.H., Chang, C., Joachimiak, A., Lee, S., and Tsai, F.T. (2016). Mitochondrial Hsp90 is a ligand-activated molecular chaperone coupling ATP binding to dimer closure through a coiled-coil intermediate. Proc. Natl. Acad. Sci. U S A *113*, 2952–2957.

Takai, N., Nakajima, M., Oyama, T., Kito, R., Sugita, C., Sugita, M., Kondo, T., and Iwasaki, H. (2006). A KaiC-associating SasA-RpaA two-component regulatory system as a major circadian timing mediator in cyanobacteria. Proc. Natl. Acad. Sci. U S A *103*, 12109–12114.

Tan, S., and Richmond, T.J. (1998). Crystal structure of the yeast MATalpha2/MCM1/DNA ternary complex. Nature *391*, 660–666.

Tian, P., and Best, R.B. (2020). Exploring the sequence fitness landscape of a bridge between two protein folds. PLoS Comput. Biol. *16*, e1008285.

Tian, P., Louis, J.M., Baber, J.L., Aniana, A., and Best, R.B. (2018). Co-evolutionary landscapes for sequence design. Angew. Chem. Int. Ed. *57*, 5674–5678.

Tseng, R., Chang, Y.G., Bravo, I., Latham, R., Chaudhary, A., Kuo, N.W., and Liwang, A. (2014). Cooperative KaiA-KaiB-KaiC interactions affect KaiB/SasA competition in the circadian clock of cyanobacteria. J. Mol. Biol. *426*, 389–402.

Tseng, R., Goularte, N.F., Chavan, A., Luu, J., Cohen, S.E., Chang, Y.G., Heisler, J., Li, S., Michael, A.K., Tripathi, S., et al. (2017). Structural basis of the day-night transition in a bacterial circadian clock. Science *355*, 1174–1180.

Tseng, T.S., Cheng, C.S., Chen, D.J., Shih, M.F., Liu, Y.N., Hsu, S.T., and Lyu, P.C. (2012). A molten globule-to-ordered structure transition of Drosophila melanogaster crammer is required for its ability to inhibit cathepsin. Biochem. J. *442*, 563–572.

Tuinstra, R.L., Peterson, F.C., Kutlesa, S., Elgin, E.S., Kron, M.A., and Volkman, B.F. (2008). Interconversion between two unrelated protein folds in the lymphotactin native state. Proc. Natl. Acad. Sci. U S A *105*, 5057–5062.

van der Lee, R., Buljan, M., Lang, B., Weatheritt, R.J., Daughdrill, G.W., Dunker, A.K., Fuxreiter, M., Gough, J., Gsponer, J., Jones, D.T., et al. (2014). Classification of intrinsically disordered regions and proteins. Chem. Rev. *114*, 6589–6631.

Vijayan, V., Zuzow, R., and O'shea, E.K. (2009). Oscillations in supercoiling drive circadian gene expression in cyanobacteria. Proc. Natl. Acad. Sci. U S A *106*, 22564–22568.

Wand, A.J., Roder, H., and Englander, S.W. (1986). Two-dimensional 1H NMR studies of cytochrome c: hydrogen exchange in the N-terminal helix. Biochemistry *25*, 1107–1114.

Wei, K.Y., Moschidi, D., Bick, M.J., Nerli, S., Mcshan, A.C., Carter, L.P., Huang, P.S., Fletcher, D.A., Sgourakis, N.G., Boyken, S.E., and Baker, D. (2020). Computational design of closely related proteins that adopt two well-defined but structurally divergent folds. Proc. Natl. Acad. Sci. U S A *117*, 7208–7215.

Werner, F. (2012). A nexus for gene expression-molecular mechanisms of Spt5 and NusG in the three domains of life. J. Mol. Biol. *417*, 13–27.

White, J.M., Delos, S.E., Brecher, M., and Schornberg, K. (2008). Structures and mechanisms of viral membrane fusion proteins: multiple variations on a common theme. Crit. Rev. Biochem. Mol. Biol. *43*, 189–219.

Woelfle, M.A., Ouyang, Y., Phanvijhitsiri, K., and Johnson, C.H. (2004). The adaptive value of circadian clocks: an experimental assessment in cyanobacteria. Curr. Biol. *14*, 1481–1486.

Wright, P.E., and Dyson, H.J. (2015). Intrinsically disordered proteins in cellular signalling and regulation. Nat. Rev. Mol. Cell Biol. *16*, 18–29.

Zuber, P.K., Artsimovitch, I., Nandymazumdar, M., Liu, Z., Nedialkov, Y., Schweimer, K., Rosch, P., and Knauer, S.H. (2018). The universally-conserved transcription factor RfaH is recruited to a hairpin structure of the non-template DNA strand. Elife *7*, e36349.

Zuber, P.K., Schweimer, K., Rosch, P., Artsimovitch, I., and Knauer, S.H. (2019). Reversible fold-switching controls the functional cycle of the antitermination factor RfaH. Nat. Commun. *10*, 702.