UNIVERSITY OF CALIFORNIA, MERCED

Structural Basis of Day/Night Transition in the Cyanobacterial Circadian Clock

A dissertation submitted in partial satisfaction of the
requirements for the degree of

Doctor of Philosophy

in

Quantitative and Systems Biology

by

Roger Tseng

Committee in charge:

Professor Michael E. Colvin, Chair
Professor Patricia LiWang
Professor Andy LiWang

2016

UNIVERSITY OF CALIFORNIA, MERCED
Graduate Division

The Dissertation of Roger Tseng is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Faculty Advisor:

_____
Andy LiWang

Committee Members:

_____
Chair: Michael E. Colvin

_____
Patricia LiWang

_____
Andy LiWang

_____
Date

DEDICATION

I never let my schooling interfere with my education.

-Mark Twain

To my advisor and my colleges, who have educated me every step of the way.

To my wife and family, who have supported my educational journey.

Thank you

ACKNOWEDGEMENT

I gratefully thank my mentors Andy LiWang, and Yonggang Chang for teaching me and walking me through this amazing journey. I am privileged to have met them and worked alongside with them. I would like to thank Dr. Carrie Partch and her great team at UCSC for this wonderful collaboration that led to these amazing discoveries. I would also like to thank Dr. Susan Cohen and Dr. Susan Golden for their great collaborative support. I thank Dr. Li Sheng and his team at UCSD. I thank Dr. Garrett at NIH for providing NMR software XIPP. I thank the materials provided by the NMRFAM NMR advanced workshop. I thank my committee members, Dr. Patricia LiWang and Dr. Michael Colvin. Last but not least, all the LiWang lab members for their support and company.

This research was supported by the U.S. Army (grant W911NF-10-1-0090), the U.S. AFOSR (grant FA 9550–13–1–0154), NIH (grant GM107521), and the National Science Foundation Graduate Research Fellowship.

I acknowledge the use of Beamline 8.3.1, which is supported by UC Office of the President, Multicampus Research Programs and Initiatives (grant MR-15-328599) and Program for Breakthrough Biomedical Research (partially funded by the Sandler Foundation).

VITA

Roger Tseng

DEGREE EARNED

Ph.D. 2010-2016 University of California, Merced, CA
Quantitative and Systems Biology

B.S. 2006-2009 University of California, Merced, CA
Biological Sciences
Emphasis*:* Molecular Biology & Biochemistry
Received with honors

PEER REVIEWED PUBLICATIONS

2015 Yong-Gang Chang, Susan E. Cohen, Connie Phong, William K. Myers, Yong-Ick Kim, **Roger Tseng**, Jenny Lin, Li Zhang, Joseph S. Boyd, Yvonne Lee, Shannon Kang, David Lee, Sheng Li, R. David Britt, Michael J. Rust, Susan S. Golden, Andy LiWang. A protein fold switch joins the circadian oscillator to clock output in cyanobacteria. *Science* 349(6245): 324-328.

2014 **Roger Tseng**, Yong-Gang Chang, Ian Bravo, Robert Latham, Abdullah Chaudhary, Nai-Wei Kuo and Andy LiWang. Cooperative KaiA–KaiB–KaiC Interactions Affect KaiB/SasA Competition in the Circadian Clock of Cyanobacteria. *Journal of Molecular Biology* 426(2):389-402.

2013 Yong-Gang Chang, **Roger Tseng**, Nai-Wei Kuo, and Andy LiWang. Nuclear magnetic resonance spectroscopy of the circadian clock of cyanobacteria. *Integrative and Comparative Biology* 53:93-102.

2012 Yong-Gang Chang, **Roger Tseng**, Nai-Wei Kuo and Andy LiWang. Rhythmic ring–ring stacking drives the circadian oscillator clockwise. *Proceedings of the National Academy of Sciences* 109(42):16847-51.

2011 Tracy R White, Michele M Conrad, **Roger Tseng**, Shaina Balayan, Rosemary Golding, António M de Frias Martins and Benoît Dayrat. Ten new complete mitochondrial genomes of pulmonates (Mollusca: Gastropoda) and their impact on phylogenetic relationships. *BMC Evolutionary Biology* 11:295.

2011 Yong-Gang Chang*, Nai-Wei Kuo*, **Roger Tseng** and Andy LiWang. Flexibility of the C-terminal, or CII, ring of KaiC governs the rhythm of the circadian clock of cyanobacteria**.** *Proceedings of the National Academy of Sciences* 108(35):14431-6.

HONORS AND AWARDS

2015 National Science Foundation Postdoctoral Research Fellowship in Biology

2013 UC Merced Outstanding Graduate Student Award

2013 Winner of the UC Merced Research Poster Competition

2013 UC Merced Quantitative and System Biology Graduate Group Travel Award

2012 UC Merced Quantitative and System Biology Graduate Group Research Award

2011 UC Merced Graduate and Research Council Summer Fellowship

2011 National Science Foundation Graduate Research Fellowship

2011 UC Merced Quantitative and System Biology Graduate Group Travel Award

2010 UC Merced Center of Excellence on Health Disparities Graduate Training Program Scholarship (funded by the National Institutes of Health)

2009 Minority Health and Health Disparities International Research Training (MHIRT) Program Scholarship (funded by the National Institutes of Health)

ABSTRACT

Structural Basis of Day/Night Transition in the Cyanobacterial Circadian Clock

Roger D. Tseng, B.S., University of California, Merced, 2009

Advisor: Dr. Andy C. LiWang

As an adaption to the solar day-night cycles, organisms have evolved endogenous circadian clocks in order to generate ~24-h rhythms in regulating their physiology and behavior. The goal of this study was to uncover mechanisms underlying the rhythm generation in the cyanobacterial clock, which, consisting of only three proteins, KaiA, KaiB, and KaiC, can be reconstituted in vitro. A wide array of biochemical and biophysical techniques including fluorescence spectroscopy, NMR, HDX-MS, and X-ray crystallography have been utilized for studying the transition of daytime (KaiA-KaiC) to nighttime (KaiA-KaiB-KaiC) complexes. In addition, genetic tools such as bioluminescence have been applied for functional validation in vivo. Specific findings to date include: (1) the day-to-night signal, KaiC phosphorylation at Ser431, is transmitted allosterically to a site over 60 Å away for initiating nighttime complex assembly; (2) KaiA cooperatively promotes the formation of the nighttime complex, in contrast to the view that KaiA is sequestered and inert at night; (3) KaiB is a metamorphic protein, forming a KaiA-KaiB-KaiC complex at night using its alternative thioredoxin-like fold; (4) KaiB inhibits KaiA through an unexpected mechanism in which KaiB disrupts KaiA-KaiC interactions by inducing a large conformational change on an α-helix of KaiA to block the binding site of KaiA for KaiC; (5) KaiB metamorphosis couples the circadian oscillator to clock output signaling by displacing the daytime output protein SasA, and engages the nighttime output protein CikA. In summary, this work reveals the role of allostery and large conformational changes in ~24-h rhythm generation in the cyanobacterial clock, creating a conceptual framework for understanding other clock systems.

TABLE OF CONTENTS

LIST OF FIGURES

Page

LIST OF TABLES

CHAPTER I

INTRODUCTION

Circadian rhythm

Organisms on Earth, ranging from bacteria, plants, fungi, insects, to animals, display circadian (~24 h) rhythms (1), as an adaption to the predictable environmental changes due to Earth's rotation. Such a rhythm was first reported in plants in 1729 by a French astronomer, de Mairan (2). The adaptive advantages of possessing such rhythms have been shown to enhance the fitness of organisms. For example, competitive growth experiments in plants and bacteria have shown that that the group whose circadian rhythm period closely matches a given period of light/dark cycles has competitive advantages over the desynchronized group (3, 4). Circadian rhythms have been implicated in aging (5), metabolism (6), cell division (7), sleep (8), neurological function (9), and cancer development (10). Understanding the basis of their generation would help tackle circadian related diseases.

Circadian clock

The molecular machinery underlying the generation of circadian rhythms are circadian clocks. Circadian clocks, have been identified through comparative studies to share three defining properties (11):

1. A free running oscillation—the clock is able to maintain a stable (~24 h) period of oscillation over a few days to weeks under constant conditions, i.e. without any environmental cues or zeitgeber (time giver).

2. Entrainment—the clock can adjust its phase and period, according to input(s) of environmental cue(s). Light and food intake are examples of strong environmental cues.

3. Temperature compensation—the period of a circadian clock is relatively insensitive to a wide range of temperature as compared to other biological systems. For example, whereas Q_{10} (the change in the rate of a reaction upon a temperature change ± 10 °C above ambient temperature) for most biological reactions is ~2-3 (12, 13), but for a circadian oscillator is between ~0.9-1.3 (14, 15). Q_{10} of 1 indicates no change due to temperature.

Components of circadian clock – input, oscillator, and output

A general scheme for describing known circadian clocks has been developed which includes input pathway(s), oscillator, and output pathway(s) (Figure 1.1) (1). Circadian oscillators lie at the core of 24-h rhythm generation. The input pathway and its components receive environmental cues to entrain the oscillator, i.e. influencing its phase and period. The output pathway and its components transduce timing signals from the oscillator to the rest of the cell, resulting in rhythmic gene expression. The circadian pattern of gene expression transduced from the output pathway probably give rise to most of the functional implications attributed to circadian clock, such as the regulation of physiology and behavior.

Circadian Oscillator Mechanism – TTFL

It is generally theorized that the oscillator functions through what is called a transcription-translation feedback loop (TTFL), which consists of positive and negative elements of the clock (1). In addition, posttranslational modifications (PTMs) of the clock proteins, such as phosphorylation, acetylation, ubiquitination and sumoylation, also have been shown to play important roles in the mechanism of circadian clocks, adding additional layers of complexity on top of the TTFL model. For example, Clock, Bmal1 and KaiC are core oscillator proteins that have circadian phosphorylation patterns that critically control circadian timing (16-18). Clock can also acetylate Bmal1 and enhance the negative element feedback of the clock (19). Phosphorylation of Per, which affects its ubiquitinated degradation, has been shown to change the length of oscillator period from 31 to 22 hours (20).

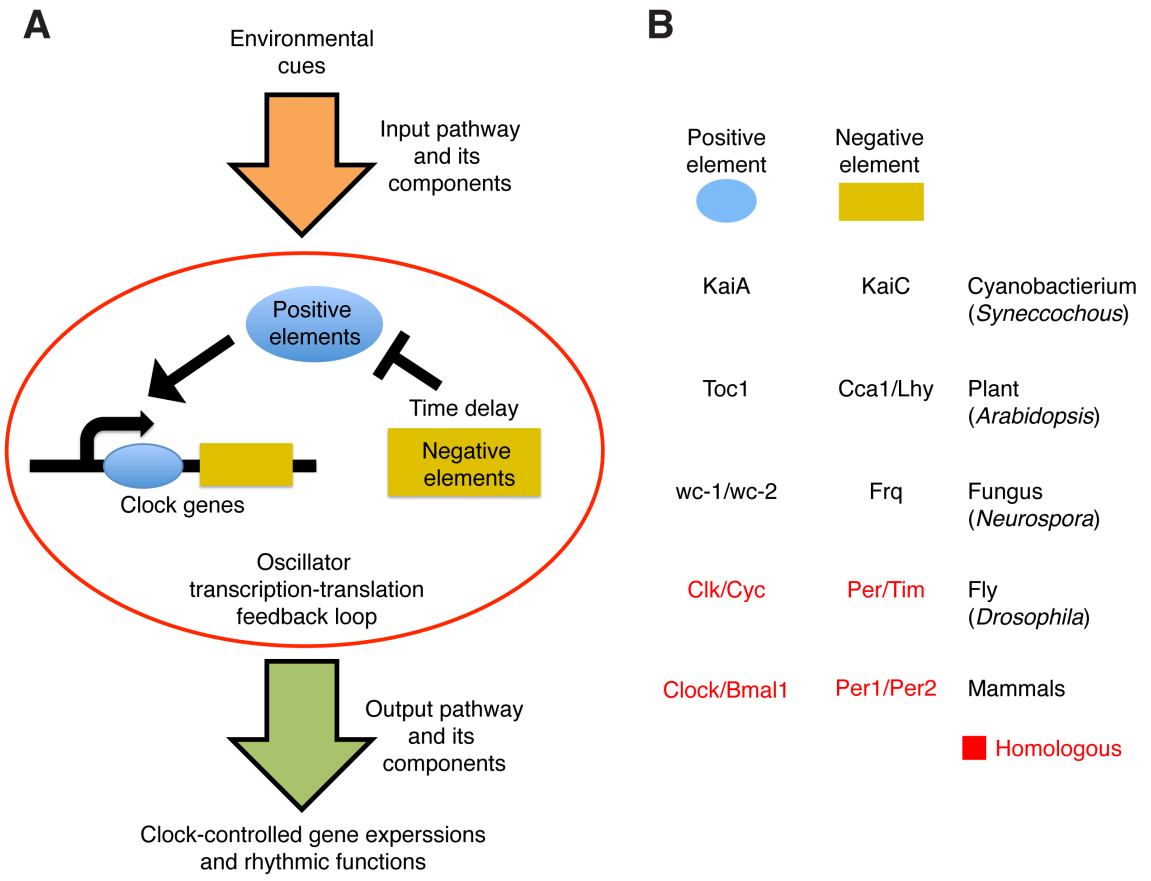


Figure 1.1. General scheme of a circadian circuit derived from comparative studies of multiple circadian models.

(A) Basic scheme of a general circadian circuit – input, oscillator, and output. Oscillator generates 24-h through a mechanism of delayed negative feedback (21). Environmental cues such as light can adjust the oscillator through the input pathway and its components. The output pathway is controlled by the oscillator genes, and drive clock-controlled gene expressions to ultimately produce the circadian phenotypes. (B) Examples of positive elements and negative elements in-cyanobacterium (*Synechococcus elongatus),* plant (*Arabidopsis thaliana*), fungus (*Neurospora crassa*), fly (*Drosophila melanogaster*), and mammals are shown: Toc1 (timing of cab expression 1), Cca1(circadian clock associated 1), Lhy (late elongated hypotcotyl), wc-1/wc-2 (*white collar 1/2*), frq (f*requency*), clk (*clock*), cyc (*cycle*), per (*period*), tim (*timeless*), bmal1 (*brain and muscle ARNT-Like 1*), and per1/per2 (*period 1/2*). Cyc is a homolog of bmal1. Per in fly and mammals are homologous as well.

The posttranslational oscillator of cyanobacteria

Although the TTFL model is prevalent in all organisms, the work by Kondo suggested that cyanobacteria harbor a clock, which can drive rhythmic gene expression in the absence of transcription-translation feedback (22). The discovery of its three components KaiA, KaiB, and KaiC, and the in vitro reconstitution of the clock by simply mixing the three components with ATP – the energy source (Fig. 1.2) – led to the establishment that cyanobacteria harbor a posttranslational oscillator. This "in vitro reconstitutability" makes the cyanobacterial clock an attractive system for biochemical, biophysical and quantitative investigation of the oscillator mechanism (23). In addition, genetic tractability of cyanobacteria allows functional validation in vivo (24). Thus, the cyanobacterial clock has become an important model system that helps provide mechanistic understanding of circadian rhythm generation.

The knowledge gained by studying the cyanobacterial clock could provide a framework for understanding eukaryotic clocks. Although sequence comparisons of the components of circadian clock models show that mammals do not possess homologs of the Kai proteins of cyanobacteria, they do share the same critical modification of component(s) – phosphorylation, and a hallmark of temporal assembly and disassembly of complexes. Interestingly, cyanobacteria and plant clock proteins also seem to utilize common protein module such as the pseudo-receiver domain in KaiA and Toc1 (25, 26). This suggests that the mechanisms of cyanobacterial clock may be applied to understand their similar modules in plant clocks.

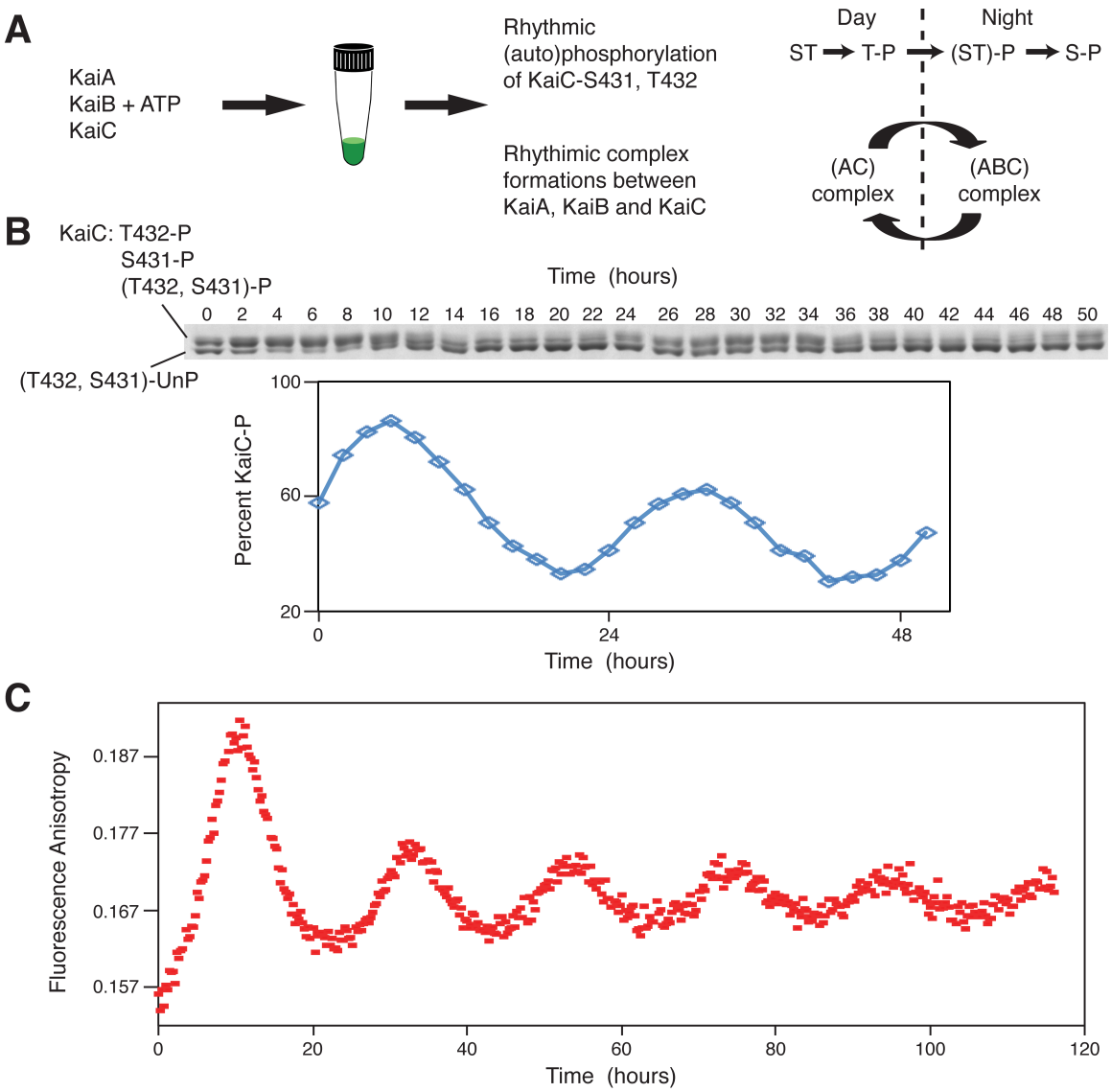


Figure 1.2. In vitro reconstitution of the cyanobacterial oscillator.

(A) Mixing purified KaiA, KaiB and KaiC proteins plus ATP as energy source generates in vitro oscillation of ordered phosphorylation of residues S431 and T432 on KaiC, as well as rhythmic complex formations between KaiA, KaiB and KaiC proteins. (B) A two-day oscillation of KaiC phosphorylation in the reconstituted oscillator. (Top panel) SDS-PAGE gel of KaiC showing phosphorylated KaiC bands and unphosphorylated KaiC band. (Bottom panel) Plot of the percentage of phosphorylated KaiC derived by densitometric analysis of the corresponding band(s) in gel. (C) An approximately five-day oscillation of fluorescent labeled KaiB in the reconstituted oscillator, measured by fluorescence anisotropy. The anisotropic changes indicate that the molecular size of KaiB is fluctuating due to rhythmic complex assembly and disassembly.

Mechanisms of the cyanobacterial clock

A bacterial circadian rhythm was first reported in 1986 in *Synechococcus* RF-1, a nitrogen fixing cyanobacteria strain (27). This finding revised the view at the time that bacterial cells lack the regulatory complexity exhibited by eukaryotic cells needed to possess a circadian clock (28). Further investigations in this regard indicate that the circadian clock facilitates growth of nitrogen fixing bacteria by temporally separating two incompatible reactions during day and night time, namely, oxygenic photosynthesis and nitrogen fixation. This temporal separation is important because oxygen is an inhibitor of the nitrogen fixing enzymes (29). Thus, this exemplifies a clear evolutionary advantage that possessing a circadian clock enhances fitness by regulating day and night time events.

Since the discovery of circadian rhythm in cyanobacteria, KaiA, KaiB and KaiC have been identified as the core oscillator proteins that regulate day and night time gene expression (30, 31). From a suite of genetic and biochemical studies, many output pathway components controlled by the oscillator have also been identified (32-40). Figure 1.3. Provides a general scheme of the cyanobacterial clock.

KaiC belongs to the AAA+ class of ATPases, and possesses autokinase, autophosphatase, and ATPase activities. KaiA and KaiB are positive and negative regulators, respectively, of KaiC phosphorylation. During subjective day, KaiA stimulates KaiC (auto)phosphorylation, whereas during subjective night, KaiB inhibits KaiA to allow KaiC to (auto)dephosphorylate. Once the (auto)dephosphorylation of KaiC returns to basal level, KaiA is released from the KaiA-KaiB-KaiC inhibitory complex and another cycle of KaiC phosphorylation begins anew (41-43).

In responding to the timing signals of the oscillator, SasA and CikA, two antagonistic enzymes, control RpaA phosphorylation levels, thereby driving distinct day and night gene expression (31, 34, 40, 44, 45). During the subjective day, SasA is stimulated and phosphorylates RpaA (33, 40, 44). At the subjective night, CikA is activated and dephosphorylates RpaA to reverse its effect (40). RpaA is a transcription factor that belongs to part of the bacterial two-component signal transduction system involving a sensor histidine kinase (SasA/CikA) and a response regulator (RpaA) (31, 40, 44). The antagonistic features of KaiA/KaiB and SasA/CikA are the bases of one of the more established pathways of day to night transition in the cyanobacterial clock.

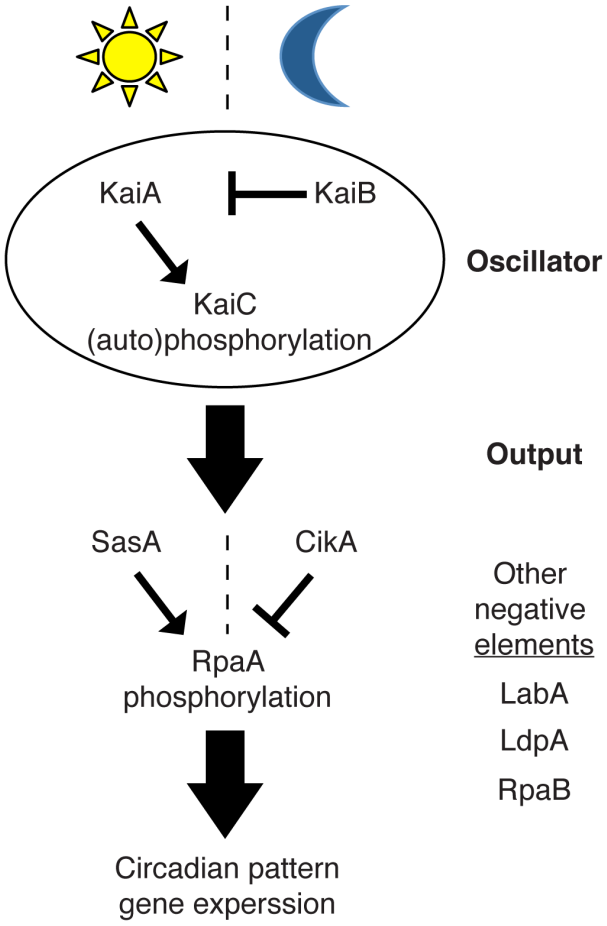


Figure 1.3. Schematic overview of the cyanobacterial clock.

In the oscillator, KaiA stimulates KaiC (auto)phosphorylation during the day; at night, KaiB promotes KaiC (auto)dephosphorylation by inhibiting KaiA. The two major output pathways involve the regulation of the phosphorylation level of RpaA, a major transcription factor that regulates circadian gene expression in cyanobacteria, by two antagonistic enzymes SasA and CikA; SasA stimulates RpaA phosphorylation during the day; at night CikA dephosphorylates RpaA. The activities of SasA and CikA are controlled by the status of the oscillator proteins. LabA, LdpA and RpaB have also been identified as output components that act as negative elements to repress gene activities. RpaB acts as a competitor to RpaA (32, 46). LdpA enhances CikA's function (37, 38). LabA is also a negative regulator of RpaA (36). However, the mechanisms of action for RpaB, LdpA and LabA are largely unclear.

Structure and dynamics of KaiC

From a series of genetic, biochemical and structural studies, KaiC has been found to be the central pacemaker of the oscillator that separates events of day from night. KaiC is a double doughnut shaped hexamer consisting of two homologous domains termed CI (N-terminus) and CII (C-terminus) forming two rings that stack on top of each other (47, 48) (Figure 1.4). KaiA stimulates KaiC (auto)phosphorylation on CII domain residues S431 and T432, from hypophosphorylated (ST) to T432 phosphorylated (T-P) during the subjective day by binding to the so-called A-loop region (residues 488-518) of KaiC, which can sample buried and exposed positions, with the former promoting KaiC (auto)dephosphorylation and the latter KaiC (auto)phosphorylation (49, 50). Structures of KaiA in complex with A-loop peptides have provided an explanation for the mechanism underlying KaiA-stimulated (auto)phosphorylation of KaiC. Several hydrophobic A-loop residues are initially buried in KaiC structure, but exit the CII pore of KaiC upon binding to KaiA (49, 51). This conformational change destabilizes a network of hydrogen bonds in the CII ring, thereby increasing the flexibility of the CII ring, which likely enhances the kinase activity of KaiC (50, 52). Indeed, KaiA enhances ADP/ATP nucleotide exchange of KaiC (53), probably by this mechanism. As KaiA continues to stimulate KaiC (auto)phosphorylation from T-P to both S431 and T432 phosphorylated ((ST)-P)[*], the phosphorylation on S431 signals the onset of the subjective night. Nuclear magnetic resonance (NMR) spectroscopy studies of KaiC have revealed that upon S431 phosphorylation, the CII ring of KaiC rigidifies and forms a more stable hexamer (52). This CII ring rigidification even enhances stacking interactions with CI ring, which induces long-range communication between CII and CI depending on the phosphorylation status of KaiC (i.e. S431 phosphorylation) (48, 52). This ring-on-ring interaction and communication between CI and CII domains is thought to be important for KaiB binding to KaiC and sequestering KaiA at night.

KaiB-KaiC complex formation initiates day/night transition in oscillator

KaiB inhibits KaiA at night, not during the day. This temporal selectivity is facilitated by KaiB preferentially binding to only two of the four phosphorylated states of KaiC (54, 55) (Figure 1.4). The mechanism of how KaiB senses S431 phosphorylation has been elusive. Several studies based on electron microscopy (EM) and hydrogen deuterium exchange mass spectroscopy (HDX-MS) have led to a model where KaiB binds to CII domain of KaiC to inhibit KaiA (56-58). However, recent biochemical and structural studies by our lab provide evidence

[*] The ordered phosphorylation of KaiC can be written as
ST → T-P → (ST)-P → S-P. It can also be written as ST → SpT → pSpT → pST. (S is Ser431, T is Thr432, P or p is phosphate).

supporting direct binding of KaiB to CI, not CII (48, 59, 60). KaiB binds to KaiC at the B-loop region (residues 116-123) on the CI domain (60). The KaiB-KaiC complex is only detected when KaiC is either in (ST)-P state or just S431 phosphorylated state (S-P) (54, 55). The phosphorylation dependent ring-on-ring interaction between CII and CI domains provides an explanation for how KaiB can sense S431 phosphorylation on CII domain that is over 60 Å away from its CI binding site. The details for how CII-CI ring stacking influence KaiB binding remain unclear. However, a possible mechanism could be that the ring-ring stacking regulates the CI ATPase activity as it has been shown that disruption of CI ATPase activity dramatically abolished KaiB binding to KaiC (53, 61). Intriguingly, KaiB has been shown to bind to isolated CI monomers much tighter than isolated CI rings (48), suggesting that the role of CI ATPase in KaiB binding may be to use the energy of ATP hydrolysis to loosen the CI hexameric interfaces (48). The B-loop region where KaiB binds to is also near the CI subunit interfaces (60). The nature of KaiB-KaiC interactions is still largely unclear and even debated due to the lack of high-resolution structure of the complex.

KaiB binding to KaiC is the first critical step that initiates the day/night transition in the cyanobacteria clock. The essential KaiB-KaiC complex strongly interacts with KaiA. Several studies have found that the newly formed KaiA-KaiB-KaiC complex is not through interactions with A loop on KaiC, but through an unidentified binding site on KaiB (43, 60). An A loop truncated variant of KaiC and isolated CI domain of KaiC can both form this stable ternary KaiA-KaiB-KaiC complex that seems to be the key state for understanding KaiA inhibition at night. However, without KaiC, KaiA-KaiB direct interactions are virtually not detectable as shown by many studies (42, 43, 62). Even though EM studies have captured this KaiA-KaiB-KaiC inhibitory complex (63, 64), the low resolution limits understanding of how KaiA is inhibited during nighttime. Thus, the mechanism of KaiA inhibition is challenging and remains a mystery.

Thus, many research groups have focused on understanding the KaiB-KaiC interactions first. Several investigations have led to the conclusion that KaiB monomerizes from a free homotetramer upon binding to KaiC and possibly undergoes large structural changes (48, 57, 65). A recent discovery has further found that KaiB switches folds upon binding to the isolated CI domain of KaiC (59). By engineering a KaiB mutant that adopts this fold-switched state (fsKaiB), it binds KaiC tightly, forms a stable ternary KaiA-KaiB-KaiC inhibitory complex and even binds to KaiA in the absence of KaiC. In vivo circadian gene expression experiments have shown that fsKaiB severely disrupts circadian rhythms by inhibiting KaiA prematurely prior to the onset of the subjective night (59). Thus, the fold switching behavior of KaiB provides a strong and rational explanation of why the KaiB homotetramer does not inhibit KaiA during the subjective day, but only during the subjective night when fsKaiB can be stabilized by KaiC. The

fsKaiB fold resembles thioredoxin (59). And this newly found state of KaiB is not only important for understanding the KaiA inhibitory mechanism but also helps explain the output mechanism involving SasA, CikA and RpaA.

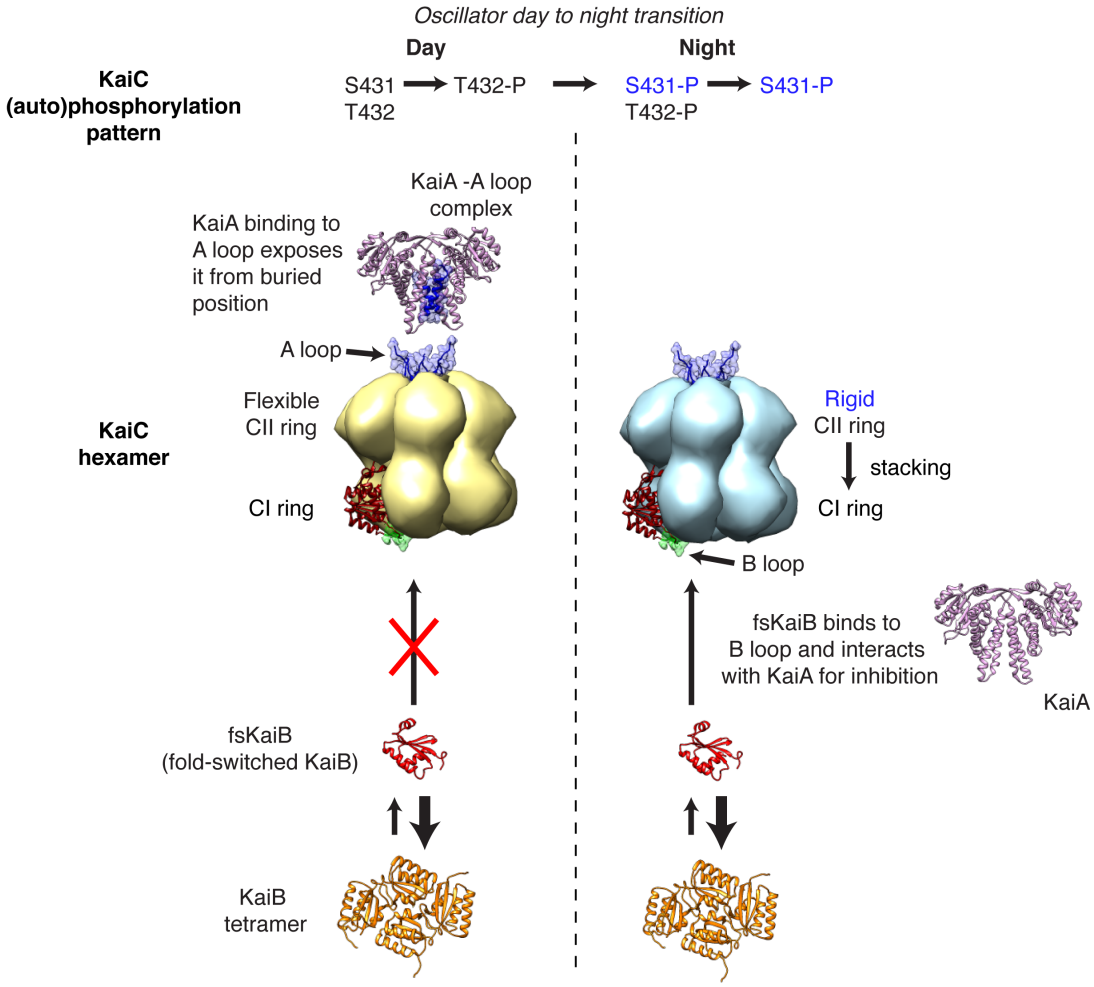


Figure 1.4. Mechanisms of day/night transition in the oscillator.

From day to night, KaiC phosphorylation occurs in an ordered pattern: ST→ T-P → (ST)-P → S-P (S indicates residue Ser431 and T is Thr432) (top). The A loop (residues 488-518 on CII domain of KaiC) is the switch for controlling KaiC phosphorylation. It can sample buried and exposed positions. In the two states, ST, T-P, where Ser431 is not phosphorylated, the CII ring is flexible, allowing KaiA to stabilize the A loop in the exposed state, which promotes KaiC (auto)phosphorylation during the day (left). Phosphorylation at S431 rigidifies KaiC CII ring, resulting in CII-CI ring stacking. Stacking is critical in promoting KaiB binding. KaiB undergoes fold-switching, with the fold-switched state (fsKaiB) binding to the B loop of CI (residues 116-123 of KaiC). The KaiC-fsKaiB complex (middle right) inhibits KaiA, initiating the night phase. PDB IDs for structures depicted in this figure: 5C5E (KaiA-A loop complex), 4O0M (KaiC hexamer), 2QKE (KaiB tetramer) and 1R8J (free KaiA).

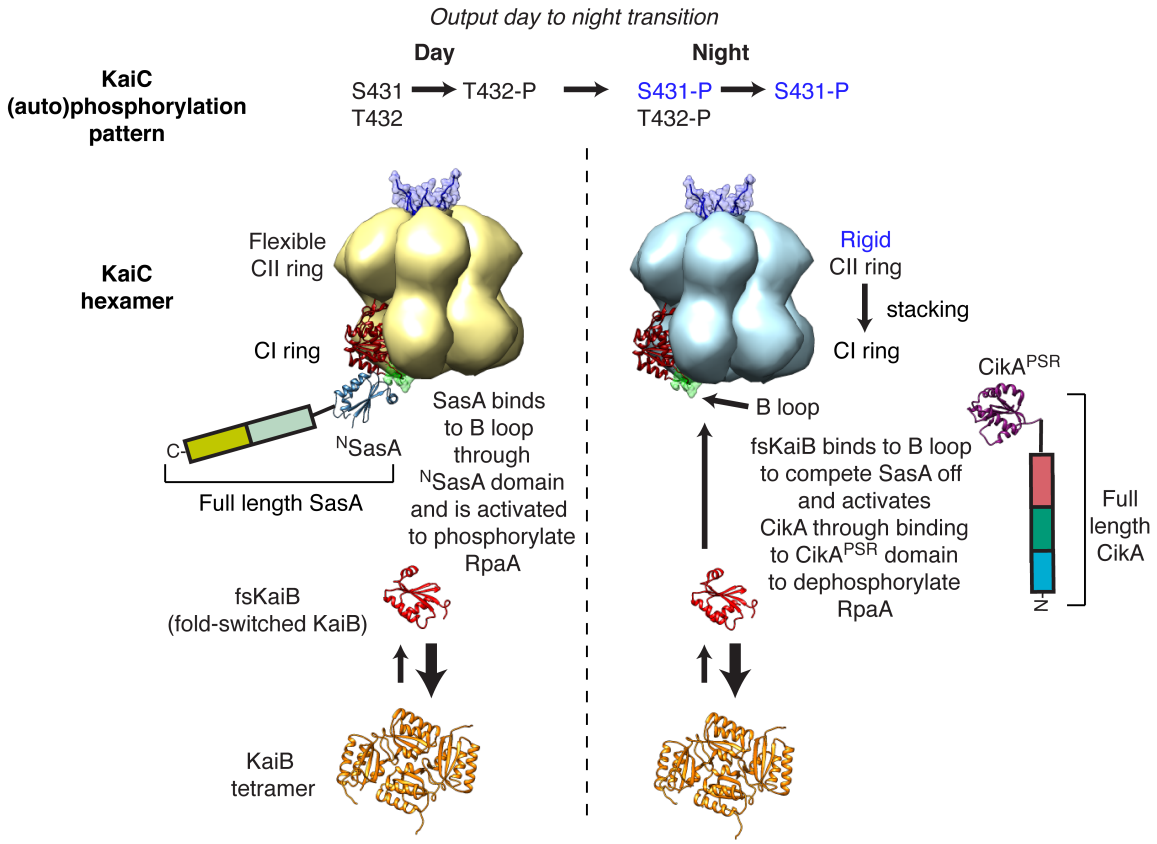


Figure 1.5. Mechanism of day/night transition and output signaling.

As KaiC executes ordered phosphorylation ST→ T-P → (ST)-P → S-P (S indicates residue S431 and T is T432) from day to night (top), the timing signals are transduced via the SasA and CikA output pathways. In the daytime, full-length SasA binds to the B loop (residues 116-123 on CI domain of KaiC) through N-terminal domain of SasA (^{N}SasA) (middle left), leading to activation of SasA-catalyzed RpaA phosphorylation. At night, KaiB is stabilized in the fold-switched state by binding to the B loop, which displaces ^{N}SasA and activates full length CikA through binding to the pseudo-receiver domain of CikA (CikA^{PSR}). CikA binding to fsKaiB at night dephosphorylates RpaA, reversing RpaA's daytime activity. PDB IDs for structures depicted in this figure: 1T4Y (^{N}SasA), 2J48 (CikA^{PSR}), 4O0M (KaiC hexamer) and 2QKE (KaiB tetramer).

KaiB-KaiC complex formation regulates day/night transition in output

During the subjective day, in the output mechanism involving SasA, CikA and RpaA, SasA is stimulated and actively phosphorylates RpaA (44). The mechanism of SasA-stimulated RpaA phosphorylation probably resembles a typical bacterial two-component system in which the sensor histidine kinase (SasA) transfers phosphorylation to the response regulator (RpaA) via a conserved histidine to conserved aspartyl acid residues (66). However, SasA is only stimulated upon KaiC binding (33, 67). Several studies have found that KaiB competes with SasA for binding to KaiC (60, 62, 64). Since the KaiB-KaiC complex occurs only at night, the output mechanism of the day/night transition proposed by many groups is that SasA binds to KaiC during subjective day and KaiB competes SasA off during subjective night to turn off the kinase activity of SasA (31, 40, 62). Full length SasA is ~389 residues with an N-terminal domain (^{N}SasA), a histidine kinase domain, and a histidine kinase-like ATPase domain. The latter two C-terminal domains are typical of sensor histidine kinases in two-component systems. ^{N}SasA has been identified to be essential in binding to KaiC (33). Several studies have shown that ^{N}SasA binds to the CI domain of KaiC (52, 67), and near B loop region (60). In light of the finding that the fsKaiB structure also binds to the B-loop region on the CI domain, ^{N}SasA/fsKaiB competition provides convincing evidence for the proposed day/night transition mechanism of SasA. In fact, ^{N}SasA shares ~60% sequence homology with KaiB and adopts a thioredoxin fold that is similar to that of fsKaiB (33, 59, 68). Currently, there is a HADDOCK model using mutagenesis and double electron-electron resonance (DEER) distant restraints showing the possible ^{N}SasA and CI complex interactions near the B loop (59). The fsKaiB findings of sharing sequence homology, similar folds to ^{N}SasA and competing for B-loop binding collectively provide strong support for KaiB binding to CI rather than to CII.

The discovery of fsKaiB also shed light on the activation mechanism of CikA during the subjective night (Figure 1.5). Full-length CikA is a 754-residue protein with four distinct domains: GAF, histidine kinase, histidine kinase-like ATPase and pseudo receiver (CikA^{PSR}) domains. Recent findings have shown that CikA antagonizes SasA's kinase activity toward RpaA (31, 40). However, CikA only antagonizes SasA when it interacts with the KaiB-KaiC complex (40). This suggests that the output mechanism of CikA is during the subjective night. Indeed, fsKaiB has been shown to bind to CikA^{PSR} directly and addition of CikA^{PSR} perturbs the in vitro KaiC phosphorylation rhythm, matching previous in vivo results (59, 69). Furthermore, CikA has been found to co-localize with KaiC at cellular poles during the subjective night (70), but not in a construct where the PSR domain is missing. Collectively, these results strongly suggest that CikA^{PSR} is the domain that binds to the KaiB-KaiC complex and is responsible for activating CikA's phosphatase activity during subjective night to antagonize

SasA. How $CikA^{PSR}$ binding to KaiB-KaiC complex is able to stimulate CikA's phosphatase is unclear. Removing $CikA^{PSR}$ has been shown to elevate CikA's (auto)kinase activity at a conserved histidine residue by 10 fold (71). Thus, $CikA^{PSR}$ has been proposed to block CikA's (auto)kinase activity (71). Perhaps upon binding to KaiB-KaiC complex, this suppression by $CikA^{PSR}$ is removed. Although the NMR structure of $CikA^{PSR}$ alone is available (72), the interactions between $CikA^{PSR}$ and the oscillator KaiB-KaiC complex remain to be delineated.

Questions to be addressed in understanding day/night transition

So far, a wealth of genetic, biochemical and structural investigations has enabled understanding of individual components and daytime events in the oscillator and output of the cyanobacterial clock. For example, structures of individual KaiA, KaiB, fsKaiB, KaiC, ^{N}SasA and $CikA^{PSR}$ proteins have been elucidated, including several key daytime complexes such as KaiA-KaiC and ^{N}SasA-KaiC. These key structures have provided great insights into clock function: for example, how KaiA stimulates KaiC (auto)phosphorylation by manipulating CII flexibility, and how KaiB separates day and nighttime events by adopting two district folds. However, how KaiB engages KaiC and inhibits KaiA remains unclear, which necessitates structural determination of the nighttime complexes, namely, KaiB-KaiC, KaiA-KaiB-KaiC, and KaiB-CikA. Thus my research has been focused on elucidating KaiA inhibition and CikA activation by solving the high-resolution structures of the nighttime complexes.

Chapter II of this dissertation investigates several molecular interactions important for the clock in the *Thermosynechococcus elongatus* system, such as KaiA-KaiC, KaiB-KaiC, KaiA-KaiB-KaiC and SasA-KaiC interactions. The biochemical studies included in this chapter provides accumulating support for a KaiB-CI binding model, shifting away from the previously proposed CII binding model. Chapter III elucidates three high-resolution nighttime complexes by X-ray crystallography and NMR, namely, the fsKaiB-KaiC, KaiA-fsKaiB-KaiC and fsKaiB-$CikA^{PSR}$ complexes. In the process, two additional NMR structures of individual $CikA^{PSR}$ and fsKaiB proteins are solved as well. The elucidation of these complexes provides first insights into nighttime KaiA inhibition and CikA activation by KaiB that is important for the day/night transition of the clock.

CHAPTER II

COOPERATIVE KAIA-KAIB-KAIC INTERACTIONS AFFECT KAIB/SASA COMPETITION IN THE CIRCADIAN CLOCK OF CYANOBACTERIA[†]

Synopsis

The circadian oscillator of cyanobacteria is composed of only three proteins, KaiA, KaiB, and KaiC. Together, they generate an autonomous ~ 24-h biochemical rhythm of phosphorylation of KaiC. KaiA stimulates KaiC phosphorylation by binding to the so-called A loops of KaiC, whereas KaiB sequesters KaiA in a KaiABC complex far away from the A loops, thereby inducing KaiC dephosphorylation. The switch from KaiC phosphorylation to dephosphorylation is initiated by the formation of the KaiB–KaiC complex, which occurs upon phosphorylation of the S431 residues of KaiC. We show here that formation of the KaiB–KaiC complex is promoted by KaiA, suggesting cooperativity in the initiation of the dephosphorylation complex. In the KaiA–KaiB interaction, one monomeric subunit of KaiB likely binds to one face of a KaiA dimer, leaving the other face unoccupied. We also show that the A loops of KaiC exist in a dynamic equilibrium between KaiA-accessible exposed and KaiA-inaccessible buried positions. Phosphorylation at the S431 residues of KaiC shift the A loops toward the buried position, thereby weakening the KaiA–KaiC interaction, which is expected to be an additional mechanism promoting formation of the KaiABC complex. We also show that KaiB and the clock-output protein SasA compete for overlapping binding sites, which include the B-loops on the CI ring of KaiC. KaiA strongly shifts the competition in KaiB's favor. Thus, in addition to stimulating KaiC phosphorylation, it is likely that KaiA plays roles in switching KaiC from phosphorylation to dephosphorylation, as well as regulating clock output.

[†] Reprinted from Journal of Molecular Biology, Vol. 426, by Tseng R., Chang YG., Bravo I., Latham R., Chaudhary A., Kuo NW. and LiWang A., titled "Cooperative KaiA-KaiB-KaiC interactions affect KaiB/SasA competition in the circadian clock of cyanobacteria", pp. 389-402, Copyright (2014), with permission from Elsevier

Introduction

Evolution of life under daily swings of ambient light and temperature produced biological timekeeping systems called circadian clocks (73). These cellular clocks prepare organisms for environmental oscillations by imposing circadian rhythms on gene expression, metabolism, physiology, and behavior in-phase with the rising and setting of the sun. These internal rhythms originate from the oscillator components of circadian clocks, and are transmitted downstream through output pathways, resulting in clock control over cellular processes with important consequences to health and reproductive fitness (7, 74-78). At the other end, environmental cues entrain the oscillators through sensory input pathways. The mechanisms of biological timekeeping at the molecular level remain largely unresolved, however.

With regard to elucidating the molecular mechanics of circadian oscillators, most of the inroads have been made in the cyanobacterial system, which is composed of only three proteins, KaiA, KaiB, and KaiC (30) (Figure 2.1). A mixture of these proteins from *Synechococcus elongatus* with ATP produces an autonomous, entrainable, and temperature-compensated circadian rhythm of KaiC (auto)phosphorylation and (auto)dephosphorylation (22, 79, 80). Recently, circadian rhythms of KaiC phosphorylation were also reconstituted in vitro using *Thermosynechococcus elongatus* proteins (62). KaiC is a homohexamer in which each subunit is composed of two RecA-like domains, called CI and CII (42, 81). The CI and CII domains self-associate into rings, and stack together like two donuts (47, 82). KaiA stimulates phosphorylation (26) by binding to the so-called A- loops of KaiC (49, 50, 83). The A-loops are located at the pore of the CII ring. Each CII domain contains two residues, S431 and T432, which in the presence of KaiA phosphorylate (S → pS and T → pT) in the following order: ST → SpT → pSpT (54, 55). Upon phosphorylation of S431, the CII ring transitions from loose to tight, and then stacks on the CI ring (52). This stacking interaction exposes part of the KaiB-binding site on CI (48). KaiB binds to KaiC and inactivates KaiA (54, 55). KaiC then passes sequentially through the phosphoforms pSpT → pST → ST, whereupon the rings unstack and a new cycle begins. This ring-ring communication explained how phosphorylation on the CII ring could induce new protein-protein interactions on the CI ring. Recently, it was demonstrated that dephosphorylation likely proceeds through phosphoryl transfer from KaiC to bound ADP molecules, followed by hydrolysis of the newly formed ATP intermediates (84, 85).

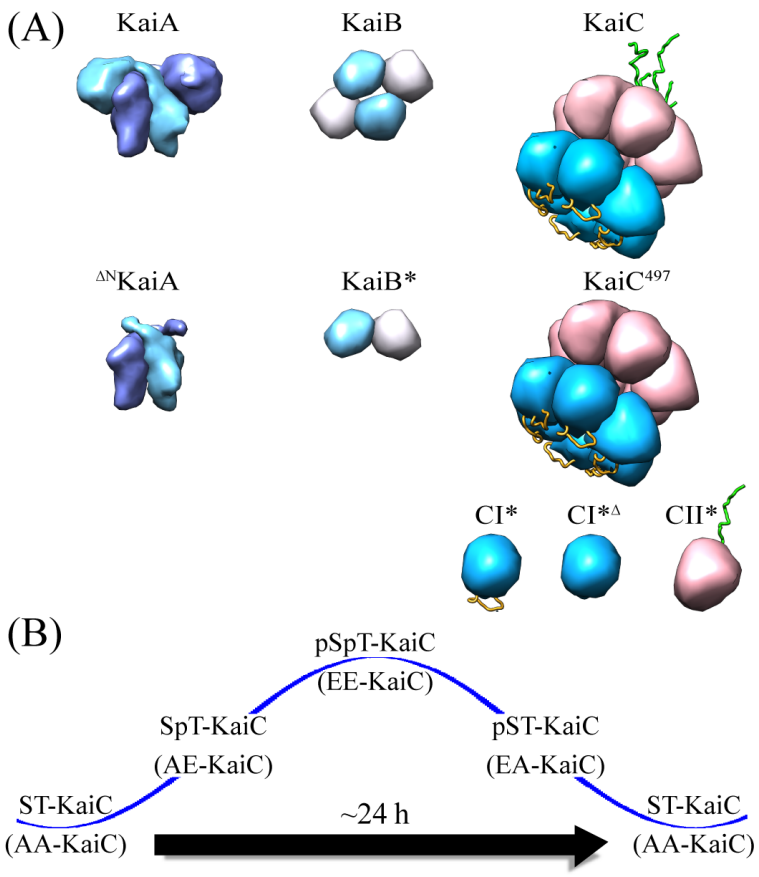


Figure 2.1 A summary of KaiABC oscillator proteins and their variants used in this study.

(A) Pictorial representations of wild-type cyanobacterial clock proteins: KaiA, KaiB and KaiC as, respectively, homodimer, homotetramer and homohexamer. Important variants of KaiA and KaiB and KaiC proteins used in this study are shown as well: $^{\Delta N}$KaiA (missing its N-terminal domain by truncating before residue S147), KaiB* (a dimeric variant by truncating after residue Y94, and with Y8A and Y94A substitutions) (48), KaiC^{497} (missing A-loops by truncating after I497) (50), CI* (monomeric CI domain: KaiC residues 1-247, with R41A and K173A substitutions) (48) , CI*$^{\Delta}$ (same as CI*, but with the deletion of B-loop from residues 116-123), and CII* (monomeric CII domain: KaiC residues 249-518, with S431E, T432E and E444D substitutions). Complete construct information is provided in Table 2.1. (B) A phosphorylation cycle of KaiC. ST-KaiC (unphosphorylated), SpT-KaiC (T432 phosphorylated), pSpT-KaiC (S431 and T432 phosphorylated), and pST-KaiC (S431 phosphorylated) can be mimicked by, respectively, AA-KaiC, AE-KaiC, EE-KaiC, and EA-KaiC. Alanyl/glutamyl substitutions mimic the unphosphorylated/phosphorylated states of S431 and T432.

With regard to entrainment, instead of having photoreceptor-mediated input pathways, the cyanobacterial clock relies on metabolic products of photosynthesis (86, 87). Oxidized quinones signal the onset of darkness by directly binding to and inactivating KaiA (88, 89) and the input pathway protein CikA (35, 39, 69, 72), whereas high ADP/ATP ratios, which act directly on KaiC, are a measure of the duration of darkness (90). Through output pathways, the cyanobacterial clock controls cellular processes such as genome-wide gene expression (91-94) and cell division (95, 96). The output pathway protein SasA (33), a sensor histidine kinase, receives timing signals from the oscillator through direct interactions of its N-terminal domain with KaiC. This interaction stimulates SasA phosphorylation (97) and subsequent phosphoryl transfer to and activation of the transcription factor RpaA (44). Whereas SasA can bind to each 17hosphomimic of KaiC (67), KaiB only binds to KaiC when the S431 residues are phosphorylated (54, 55), and can displace SasA from those phosphoforms (62, 64). Interestingly, stimulation of CikA by the KaiB-KaiC complex during the subjective night induces dephosphorylation of RpaA (40). Thus, rhythms in KaiB-KaiC binding may regulate clock output by controlling sequential action of SasA and CikA on RpaA. Recently, we showed that both SasA (52) and KaiB (48) bind to the CI domain of KaiC, suggesting that they compete for overlapping binding sites.

Until now, it has been thought that KaiA only stimulates KaiC phosphorylation (26, 98) and signals the onset of darkness (86). In other words, KaiA has not been presumed to play active roles in the dephosphorylation of KaiC or clock output. Here, using proteins from *T. elongatus*, we investigated whether those two presumptions are warranted. We show that in fact, KaiA plays a cooperative role in the assembly of the KaiA-KaiB-KaiC dephosphorylation complex, and that decreasing KaiA-KaiC affinities during KaiC phosphorylation enhances this cooperativity. We also show that KaiA promotes the displacement of SasA by KaiB from KaiC, implicating KaiA as one of the regulators of clock output. Apparently, a dimer of KaiA binds only one KaiB monomer. Additionally, we show that KaiB and SasA bind to the B-loop of the CI domain of KaiC.

Materials and methods

Cloning and Constructs of kaiA, *kaiB*, and kaiC: All genes were cloned into pET-28b using the Nde I/Hind III sites by PCR. The gene encoding SUMO was spliced with the *kai* genes using PCR. Details of the cloning protocol have been described previously (52). Please see Table 2.1 for the list of protein constructs.

Protein expression and purification: Proteins were expressed in BL21(DE3) Escherichia coli (Novagen) and purified by Ni-NTA affinity chromatography and size-exclusion chromatography as described previously (48), except KaiC proteins were expressed in M9 (H_2O) minimal medium with $^{14}NH_4Cl$ for higher purity and yield. The pET-28b plasmids harboring *kaiA, kaiB,* or *kaiC* were used to transform *Escherichia coli* BL21(DE3) cells (Novagen). Cells were grown in LB (H_2O), M9 (H_2O), and M9 (D_2O) media for production of unlabeled, U-[^{15}N, ^{1}H]-labeled, and [^{15}N, ^{2}H, $^{1}H^{N}$]-labeled or U-[^{15}N, ^{2}H]-Ile-δ1-[^{13}C, ^{1}H]-labeled proteins, respectively. IPTG (isopropyl β-D-1-thiogalactopyranoside) (Research Products International) at a final concentration of 0.2 mM was added to induce protein expression when OD_{600} reached 0.4-0.6. The induction time was ~12 h at 25°C for over-expression of unlabeled proteins, or ~12 h for production of ^{15}N-labeled, or [^{15}N, ^{2}H, $^{1}H^{N}$]-labeled or U-[^{15}N, ^{2}H]-Ile-δ1-[^{13}C, ^{1}H]-labeled proteins. Cells were harvested by centrifugation at 5,000 rpm for 15 min, followed by cell lysis and protein purification. Proteins were purified by Ni-NTA affinity chromatography and gel filtration chromatography. The disposable 10 ml Polypropylene Columns (Thermo Scientific) packed with Ni-NTA agarose (VWR) were used for Ni-NTA gravity chromatography, whereas pre-packed columns including HiLoad 16/60 Superdex 75 Column (GE Healthcare), and HiLoad 16/60 Superdex 200 Column (GE Healthcare) were used for gel-filtration chromatography.

Labeling proteins with 6-IAF (Invitrogen) was achieved by adding the dye to protein samples at a 5:1 dye:protein molar ratio followed by overnight incubation of the mixture at 4°C in darkness. A subsequent desalting step was applied to remove residual dye. Note that the protein samples were supplemented with 1 mM TCEP before addition of the dye to keep the Cys residues reduced for efficient fluorescence labeling. All protein samples were concentrated with Amicon Stirred Ultrafiltration Cell (Millipore) by applying nitrogen gas at 45 psi and using a YM-10 ultrafiltration membrane (Millipore). Protein concentrations were determined using Coomassie Plus Assay (Pierce) by following the manufacturer's protocol. The absorbance for BSA at a series of concentrations (0, 0.2, 0.4, 0.6, and 0.8 mg/ml) was obtained to generate the standard curve. The protein (or BSA) and dye volumes were 5 µL and 150 µL,

respectively. The volume for OD_{595} measurement was 100 µL. The OD_{595} reading was obtained on a DU 640 UV/VIS spectrophotometer (Beckman).

Analytical gel-filtration chromatography: All gel-filtration chromatography experiments were performed with a Superdex 200 10/300 GL column (GE Healthcare) as described previously (48), except that the sample injection volume was 250 µL. Briefly, desired protein sample(s) and buffer were mixed to total of 250 µL, with or without incubation time in room temperature depending on experimental needs. The reaction sample(s) then is applied to a Superdex 200 10/300 GL column with a flow rate of 0.5 mL/min, at room temperature. The volume of the sample loop was 100 µL. Peak fractions of the run were collected and subjected to lyophilization and SDS/PAGE. Details of experimental setup and proteins used in each experiment are provided in Table 2.1.

Molecular weight markers used to calibrate the size-exclusion columns: Ferritin (440 kDa), albumin (67 kDa), ovalbumin (43 kDa), and ribonuclease A (13.7 kDa) from the gel-filtration HMW and LMW calibration kits (GE Healthcare) were used as molecular weight markers.

Fluorescence spectroscopy: All data were measured with an ISS PC1 spectrofluorometer. The thiol-reactive fluorophores used for all experiments were 6-iodoacetamidofluorescein (6-IAF) (Invitrogen). Data were fit to a single exponential equations [1-4] depending on the experimental setup (99), with Mathematica software (Wolfram).

$$F = Fmin + (Fmax - Fmin)\ \frac{[L]+[P]+KDapp-\sqrt{([L]+[P]+KDapp)^2-4[L][P]}}{2[P]} \qquad [1]$$

$$F = Fmin + (Fmax - Fmin)\ \frac{[L]}{[L]+KDapp} \qquad [2]$$

where F is fluorescence anisotropy, $[P]$ is total concentration of fluorophore-labeled protein, and $[L]$ is total concentration of ligand added. The standard errors were estimated from duplicate runs.

$$F = (Fmax - Fmin)(1 - e^{-kobst}) + Fmin \qquad [3]$$

$$F = (Fmax - Fmin)(e^{-kobst}) + Fmin \qquad [4]$$

where F is fluorescence anisotropy, and t is in hours.

Details of experimental setup and proteins used in each experiment are provided in Table 2.1. An ISS PC1 spectrofluorimeter was operated by using the Vinci software provided. A water bath (VWR International) was connected to the

spectrofluorimeter to control the temperature of the cuvette compartment. Excitation wavelength and slit width were 492 nm and 2 mm, respectively, with a 497/16 nm BrightLine® single-band bandpass filter installed. Emission wavelength and slit width were 530 nm and 2 mm, respectively, with a 524/24 nm BrightLine® single-band bandpass filter and monochromator installed. Detailed setup was followed as previously described (48).

A-loop proteolysis reactions: Proteolysis experiments were carried out using 10 μM of each of the five KaiC variants (AA-KaiC, AE-KaiC, EE-KaiC, EA-KaiC, and EA-KaiC^{E444D}) with a thrombin cut site (LVPRGS) replacing residues I497 to K502. KaiC samples were mixed with 2 units of thrombin (Sigma) and incubated at 30°C in a reaction buffer (20 mM Tris, 50 mM NaCl, 1 mM ATP, 1 mM $MgCl_2$, 0.02% NaN_3, pH 7.0). For control experiments, buffer was added instead of thrombin. 10 μL aliquots were taken at indicated time points in Figure 2.10 from the reaction mixtures for SDS-PAGE analysis. The reaction of each time point was stopped by adding of 10 μL of SDS-PAGE gel-loading dye (100 mM Tris, 4% SDS, 0.2% bromophenol blue, 20% glycerol, 400 mM β-mercaptoethanol, pH 6.8) and freezing at -20°C. The samples were loaded onto gels: 9 x 10 cm SDS polyacrylamide gels (4% of acrylamide/bisacrylamide for stacking gel, and 9% of acrylamide/bisacrylamide for running gel) with 15 wells (10 x 3 x 0.75 mm). The gels were subjected to electrophoresis at 60 volts in glycine buffer for 30 minutes and then at 140 volts for 120 minutes, with the buffer chamber surrounded by ice. Gels were stained with Coomassie brilliant blue R250. The percentage of thrombin proteolysis product in each lane was determined by densitometric analysis using Image J (National Institutes of Health) and PeakFit (SeaSolve Software, Inc.). The k_{obs} values were fit to equation [5] using Mathematica,

$$P(t) = (Pmax - Pmin)(1 - e^{-kobst}) + Pmin \qquad [5]$$

where $P(t)$ is percentage of cut product at time t (hours). *Pmin* and *Pmax* are respectively the percent cut at $t = 0$, and $t = \infty$. The standard error was estimated from duplicate runs. For details on the protein constructs and experimental setup, please see Table 2.1.

In vitro KaiC phosphorylation reactions: Experimental setup was as described previously (52), except that 5 μM of KaiC was incubated with 5 μM KaiB and 1.67 μM of KaiA or $^{\Delta N}$KaiA at 35 °C. Briefly, KaiC was preincubated at 37°C for 24 hours to allow samples to autodephosphorylate to low levels of phosphorylation before use in experiments. KaiC proteins were then incubated with KaiB and KaiA (or $^{\Delta N}$KaiA), in a reaction buffer (20 mM Tris, 150 mM NaCl, 0.5 mM EDTA, 1 mM ATP, 5 mM $MgCl_2$, 0.02% NaN_3, pH 7.0) at 35°C. The final concentrations of KaiA (or $^{\Delta N}$KaiA), KaiB, and KaiC were 1.67, 5, and 5 μM,

respectively. 10 μL aliquots were taken at indicated time points from the reaction mixtures for SDS-PAGE analysis. The reaction of each time point was stopped by adding 10 μL of SDS-PAGE gel-loading dye (100 mM Tris, 4% SDS, 0.2% bromophenol blue, 20% glycerol, 400 mM β-mercaptoethanol, pH 6.8) and freezing at -20°C. Approximately 100 μL of transparent mineral oil (CMP 25, Cambridge mill products. Inc.) was layered on top of the reaction solution to prevent evaporation. Aliquots were obtained at indicated time points for SDS/PAGE analysis. Following electrophoresis as described previously (52), gels were stained with Coomassie brilliant blue R250 (EMD Chemicals), and the percentage of KaiC phosphorylation was determined by densitometry using ImageJ (National Institutes of Health) and PeakFit (SeaSolve Software). Details of experimental setup and proteins used in each experiment are provided in Table 2.1.

Without λ phosphatase: 10 μM of KaiC and KaiC^{E444D} were incubated at 37°C for 24 hours in a reaction buffer (20 mM Tris, 150 mM NaCl, 0.5 mM EDTA, 1 mM ATP, 5 mM $MgCl_2$, 0.02% NaN_3, pH 7.0). 10 μL aliquots were taken at indicated time points from the reaction mixtures for SDS-PAGE analysis. The reaction of each time point was stopped by adding 10 μL of SDS-PAGE gel-loading dye (100 mM Tris, 4% SDS, 0.2% bromophenol blue, 20% glycerol, 400 mM β-mercaptoethanol, pH 6.8) and freezing at -20°C. Details of the running conditions for SDS-PAGE and densitometric analysis were the same as for in vitro KaiC phosphorylation reactions. Details on the protein constructs and experimental setup are provided in Table 2.1.

With λ phosphatase: 10 μM of KaiC and KaiC^{E444D} were incubated at 30°C for 2 hours in a reaction buffer (20 mM Tris, 50 mM NaCl, 1 mM ATP, 1 mM $MgCl_2$, 5 mM $MnCl_2$, 0.02% NaN_3, pH 7.0) with 4000 units of λ phosphatase (New England BioLabs Inc.) ± 50 mM EDTA (λ phosphatase inhibitor). As a control, buffer was added instead of λ phosphatase. 10 μL aliquots were taken at indicated time points from the reaction mixtures for SDS-PAGE analysis. The reaction of each time point was stopped by adding 10 μL of SDS-PAGE gel-loading dye (100 mM Tris, 4% SDS, 0.2% bromophenol blue, 20% glycerol, 400 mM β-mercaptoethanol, pH 6.8) and freezing at -20°C. Details of the running conditions for SDS-PAGE and densitometric analysis were the same as for in vitro KaiC phosphorylation reactions. Details on the protein constructs and experimental setup are provided in Table 2.1.

NMR spectroscopy: All NMR experiments were run on a Bruker 600MHz AVANCE III spectrometer equipped with a TCI cryoprobe. Chemical shifts were referenced to internal DSS. Data were processed using NMRPipe and visualized using NMRDraw (100). For details on the protein constructs and experimental setup, please see Table 2.1.

Table 2.1. Abbreviation and full name of *Thermosynechococcus elongatus* Kai proteins used in each experimental condition in Chapter Two.

Experiment Type	[Abbreviation]-Protein Full Name (N-terminal Tag_Protein_Mutation_Length of Protein_ C-terminal Tag) / Final Concentration (μM)	Experimental Condition (all samples include 0.02% w/v ratio of NaN_3)
Analytical Gel Filtration Chromatography (Figure 2.2, 2.12, 2.3A.)	1) [$^{\Delta N}$KaiA]-FLAG_KaiA_147-283 / 50, 100, 150 2) [KaiB*]-FLAG_KaiB_Y8A-Y94A_1-94_FLAG / 50 3) [CI*]-FLAG_CI-KaiC_R41A-K173A_1-247_FLAG / 50 4) [CI*$^{\Delta}$]-FLAG_CI-KaiC_R41A-K173A-Δ116-123_1-247_FLAG / 50 5) [^{N}SasA]-FLAG_SasA_P16A_16-107_FLAG / 50	• Volume: 250 μL • Pre-experiment incubation time of mixture: 4 hrs • Temperature: 30 °C • Buffer: 20 mM Tris, 50 mM NaCl, 5 mM DTT, 1 mM ATP, 1 mM $MgCl_2$, pH 7
Analytical Gel Filtration Chromatography (Figure 2.3B.)	1) [EE-KaiC^{497}]-FLAG_KaiC_S431E-T432E_1-497_ / 50 2) [KaiB]-AMA_KaiB / 50	• Volume: 250 μL • Pre-experiment incubation time of mixture: 24 hrs • Temperature: 30 °C • Buffer: 20 mM Tris, 50 mM NaCl, 5 mM DTT, 1 mM ATP, 1 mM $MgCl_2$, pH 7
Fluorescence Spectroscopy (Figure 2.4B)	1) [KaiB]-AMA_KaiB_A54C (6-IAF-labeled) / 0.02 2) [CI*]-FLAG_CI-KaiC_R41A-K173A_1-247_FLAG / 1, 5, 10, 20, 40, 150 3) KaiA / 5	• Volume: 300 μL • Pre-experiment incubation time of mixture: 24 hrs • Temperature: 25 °C • Buffer: 20 mM Tris, 50 mM NaCl, 5 mM DTT, 1 mM ATP, 1 mM $MgCl_2$, pH 7
Fluorescence Spectroscopy (Figure 2.7A)	1) [KaiB]-AMA_KaiB_A54C (6-IAF-labeled) / 0.02 2) [CI*]-FLAG_CI-KaiC_R41A-K173A_1-247_FLAG / 5 3) KaiA / 5	• Volume: 400 μL • Pre-experiment incubation time: 20 mins • Experiment time: 24 hrs • Temperature: 25 °C • Buffer: 20 mM Tris, 50 mM NaCl, 5 mM DTT, 1 mM ATP, 1 mM $MgCl_2$, pH 7 • Measurement time interval: every 30 mins

Fluorescence Spectroscopy (Figure 2.9, 2.10C)	1) [$^{\Delta N}$KaiA]-KaiA_147-283 (6-IAF-labeled) / 0.2 2) [AA-KaiC]-FLAG_KaiC_S431A-T432A / 0, 0.1, 0.25, 0.5, 1, 1.5, 2, 3, 6, 9, 12, 25 3) [AE-KaiC]-FLAG_KaiC_S431A-T432E / 0, 0.1, 0.25, 0.5, 1, 1.5, 2, 3, 6, 9, 12, 25 4) [EE-KaiC]-FLAG_KaiC_S431E-T432E / 0, 0.1, 0.25, 0.5, 1, 1.5, 2, 3, 6, 9, 12, 25 5) [EA-KaiC]-FLAG_KaiC_S431E-T432A / 0, 0.1, 0.25, 0.5, 1, 1.5, 2, 3, 6, 9, 12, 25 6) [EA-KaiC^{E444D}]-FLAG_KaiC_S431E-T432A-E444D / 0, 0.1, 0.25, 0.5, 1, 1.5, 2, 3, 6, 9, 12, 25	• Volume: 300 µL • Pre-experiment incubation time of mixture: 20 mins • Temperature: 25 °C • Buffer: 20 mM Tris, 50 mM NaCl, 5 mM DTT, 1 mM ATP, 1 mM $MgCl_2$, pH 7
Fluorescence Spectroscopy (Figure 2.11A, 2.7B.)	1) [^{N}SasA]-FLAG_SasA_P16A-G57C_16-107_FLAG (6-IAF-labeled) / 0.025 2) [EE-KaiC^{497}]-FLAG_KaiC_S431E-T432E_1-497 / 2.5 3) [KaiB*]-FLAG_KaiB_Y8A-Y94A_1-94 / 10 4) [KaiB]-AMA_KaiB / 10 5) KaiA / 10, 20 6) [$^{\Delta N}$KaiA]-FLAG_KaiA_147-283 / 10, 20	• Volume: 400 µL • Pre-experiment incubation time: 2 hrs • Experiment time: 8 hrs • Temperature: 30 °C • Buffer: 20 mM Tris, 50 mM NaCl, 5 mM DTT, 1 mM ATP, 1 mM $MgCl_2$, pH 7 • Measurement time interval: 1.67, 2.5, 5 and 10 mins, then every 20 mins hereafter. (additional 0.2 and 0.41 mins were added for KaiB* measurement)
Fluorescence Spectroscopy (Figure 2.11B, 2.7C.)	1) [^{N}SasA]-FLAG_SasA_P16A-G57C_16-107_FLAG (6-IAF-labeled) / 0.025 2) [CI*]-FLAG_CI-KaiC_R41A-K173A_1-247_FLAG / 2.5 3) [KaiB*]-FLAG_KaiB_Y8A-Y94A_1-94 / 10 4) [KaiB]-AMA_KaiB / 10 5) KaiA / 10, 20 6) [$^{\Delta N}$KaiA]-FLAG_KaiA_147-283 / 10, 20	• Volume: 400 µL • Pre-experiment incubation time: 20 mins • Experiment time: 24 hrs • Temperature: 30 °C • Buffer: 20 mM Tris, 50 mM NaCl, 5 mM DTT, 1 mM ATP, 1 mM $MgCl_2$, pH 7 • Measurement time interval: 2.5, 5, and 10 mins, then every 20 mins hereafter

In Vitro KaiC Phosphorylation Reactions (Figure 2.6A-B)	1) KaiA / 1.67 2) [$^{\Delta N}$KaiA]-FLAG_KaiA_147-283 / 1.67 3) [KaiB]-AMA_KaiB / 5 4) [KaiC]-FLAG_KaiC / 5	• Volume: 200 µL • Experiment time: 24 hrs • Temperature: 35 °C • Buffer: 20 mM Tris, 150 mM NaCl, 0.5 mM EDTA, 1 mM ATP, 5 mM MgCl2, pH 7.0
Autodephosphorylation Reactions of KaiC and KaiC^{E444D} (Figure 2.6C.)	1) [KaiC]-FLAG_KaiC / 10 2) [KaiC^{E444D}]-FLAG_KaiC_E444D / 10	• Volume: 50 µL • Experiment time: 24 hrs • Temperature: 37 °C • Buffer: 20 mM Tris, 150 mM NaCl, 0.5 mM EDTA, 1 mM ATP, 5 mM $MgCl_2$, pH 7.0
Autodephosphorylation Reactions of KaiC and KaiC^{E444D} with λ phosphatase (Figure 2.6D.)	1) [KaiC]-FLAG_KaiC / 10 2) [KaiC^{E444D}]-FLAG_KaiC_E444D / 10	• Volume: 50 µL • Experiment time: 2 hrs • Temperature: 30°C • Buffer: 20 mM Tris, 50 mM NaCl, 1 mM ATP, 1 mM $MgCl_2$, 5 mM $MnCl_2$, pH 7.0 • λ phosphatase added: 4000 units • λ phosphatase inhibitor added: 50 mM EDTA
A-loop Proteolysis (Figure 2.9, 2.10A-B.)	1) [AA-KaiC^{LVPRGS}]-FLAG_KaiC_S431A-T432A-I497L-S498V-V499P-D500R-E501G-K502S / 10 2) [AE-KaiC^{LVPRGS}]-FLAG_KaiC_S431A-T432E-I497L-S498V-V499P-D500R-E501G-K502S / 10 3) [EE-KaiC^{LVPRGS}]-FLAG_KaiC_S431E-T432E-I497L-S498V-V499P-D500R-E501G-K502S / 10 4) [EA-KaiC^{LVPRGS}]-FLAG_KaiC_S431E-T432A-I497L-S498V-V499P-D500R-E501G-K502S / 10 5) [EA-KaiC$^{E444D\text{-}LVPRGS}$]-FLAG_KaiC_S431E-T432A-E444D-I497L-S498V-V499P-D500R-E501G-K502S /10	• Volume: 100 µL • Experiment time: 24 hrs • Temperature: 30 °C • Buffer: 20 mM Tris, 50 mM NaCl, 1 mM ATP, 1 mM $MgCl_2$, pH 7 • Thrombin added: 2 units

NMR Methyl-TROSY (Figure 2.4A, 2.5.)	1) [$^{\Delta N}$KaiA]-FLAG_KaiA_147-283 (U-[^{15}N, ^{2}H]-Ile-1-[^{13}C, ^{1}H] labeled) / 20 2) [KaiB*]-FLAG_KaiB_Y8A-Y94A_1-94 / 50, 100, 200 3) [KaiB]-AMA_KaiB / 50, 200 4) [CI*]-FLAG_CI-KaiC_R41A-K173A_1-247_FLAG / 50 5) [CII*]-FLAG_CII-KaiC_S431E-T432E-E444D_249-518 / 50	• Volume: 350 µL • Pre-experiment incubation time of mixture: 8 hrs • Number of scan: 256 • ^{1}H / ^{13}C sweep width (ppm): 6.48 / 8.00 • ^{1}H / ^{13}C carrier (ppm): -0.657 / 13.397 • ^{1}H / ^{13}C acquisition time (ms): 64.0 / 82.7 • Temperature: 30 °C • Buffer in 99.96 % D_2O: 20 mM Tris, 75 mM NaCl, 5 mM DTT, 1 mM ATP, 1 mM $MgCl_2$, 10 µM DSS, pH 7 • Shaped tube
NMR Methyl-TROSY (Figure 2.13, 2.14.)	1) [CI*]-FLAG_CI-KaiC_R41A-K173A_1-247_FLAG (U-[^{15}N, ^{2}H]-Ile-1-[^{13}C, ^{1}H] labeled) / 20 2) [^{N}SasA]-FLAG_SasA_P16A_16-107_FLAG / 50 3) [KaiB*]-FLAG_KaiB_Y8A-Y94A_1-94 / 50 4) [$^{\Delta N}$KaiA]-FLAG_KaiA_147-283 / 50	• Volume: 350 µL • Number of scan: 256 • ^{1}H / ^{13}C sweep width (ppm): 6.48 / 8.00 • ^{1}H / ^{13}C carrier (ppm): -0.609 / 13.445 • ^{1}H / ^{13}C acquisition time (ms): 64.0 / 82.7 • Temperature: 30 °C • Buffer in 99.96 % D_2O: 20 mM Tris, 50 mM NaCl, 5 mM DTT, 1 mM ATP, 1 mM $MgCl_2$, 10 µM DSS, pH 7 • Shaped tube
NMR ^{15}N-TROSY (Figure 2.8A.)	1) [KaiB*]-FLAG_KaiB_Y8A-Y94A_1-94 (U-[^{15}N, ^{2}H] labeled) / 150 2) [$^{\Delta N}$KaiA]-FLAG_KaiA_147-283 / 675	• Volume: 350 µL • Pre-experiment incubation time of mixture: 8 hrs • Number of scan: 128 (free KaiB*), 256 (KaiB* + $^{\Delta N}$KaiA) • ^{1}H / ^{15}N sweep width (ppm): 16.00 / 26.48 • ^{1}H / ^{15}N carrier (ppm): 4.696 / 119.279 • ^{1}H / ^{15}N acquisition time (ms): 69.8 / 79.4 • Temperature: 30 °C • Buffer in 5% D_2O: 5 mM Tris, 50 mM NaCl, 10 µM DSS, pH 7 • Shaped tube

NMR ^{15}N-TROSY (Figure 2.8B.)	1) [CI*]-FLAG_CI-KaiC_R41A-K173A_1-247_FLAG (U-[^{15}N] labeled) / 50 2) [CI*$^{\Delta}$]-FLAG_CI-KaiC_R41A-K173A-Δ116-123_1-247_FLAG (U-[^{15}N] labeled) / 50	• Volume: 350 µL • Number of scan: 512 • ^{1}H / ^{15}N sweep width (ppm): 18.00 / 35.23 • ^{1}H / ^{15}N carrier (ppm): 4.756 / 119.628 • ^{1}H / ^{15}N acquisition time (ms): 64.9 / 29.8 • Temperature: 25 °C • Buffer in 5% D_2O: 20 mM Tris, 50 mM NaCl, 5 mM DTT, 1 mM ADP, 1 mM $MgCl_2$, 10 µM DSS, pH 7 • Shaped tube
NMR ^{15}N-TROSY (Figure 2.8C.)	1) [^{N}SasA]-FLAG_SasA_P16A_16-107 (U-[^{15}N] labeled) / 80 2) [CI*]-FLAG_CI-KaiC_R41A-K173A_1-247_FLAG / 156	• Volume: 350 µL • Number of scan: 264 • ^{1}H / ^{15}N sweep width (ppm): 16.00 / 26.48 • ^{1}H / ^{15}N carrier (ppm): 4.746 / 119.331 • ^{1}H / ^{15}N acquisition time (ms): 69.8 / 79.4 • Temperature: 25 °C • Buffer in 5% D_2O: 20 mM Tris, 50 mM NaCl, 10 µM DSS, pH 7 • Shaped tube

Results

Resolving whether the N-terminal domains of KaiA are necessary for binding KaiB

The N-terminal domains of KaiA from *Synechococcus elongatus* have recently been demonstrated to play a critical role in entrainment by sensing oxidized quinones (86). An X-ray crystal structure of KaiA verified that the N-terminal domains bind directly to the oxidized form of quinone analog 2,5-dibromo-3-methyl-6-isopropyl-*p*-benzoquinone (DBMIB) (89). Less clear is the role of the N-terminal domains in generating KaiC phosphorylation rhythms. Recent electron microscopy studies have suggested that, in order for KaiC to dephosphorylate, KaiB sequesters monomeric subunits of KaiA by binding exclusively to its N-terminal domains (64). If correct, the N-terminal domains of KaiA would be necessary for the generation of phosphorylation rhythms. However, electron spin resonance spectroscopy suggests that the C-terminal domains of KaiA bind KaiB (101). Additionally, *Anabaena* sp. Strain PCC 7120, which supposedly displays circadian rhythms (102), lacks N-terminal domains (103). Also, truncating the N-terminal domains of KaiA did not abolish in vivo rhythms of bioluminescence (104). We therefore wanted to clarify these contradicting models. In Figure 2.1, we provided a list of the different constructs of KaiA, KaiB, and KaiC used in this study.

Here, using size-exclusion chromatography of *T. elongatus* proteins, we found that a KaiA construct, $^{\Delta N}$KaiA, which lacked its N-terminal domains by truncating before residue S147 (Figure 2.1A), formed a stable ternary complex with KaiB* (a dimeric variant of KaiB) (48) and CI* (a monomeric variant of the CI domain) (48) (Figure 2.2A, left panel). Because KaiB* binds to CI as a monomer (48), it is likely that KaiB* is a monomer in this ternary complex as well. Next, we tested formation of the $^{\Delta N}$KaiA-KaiB*-CI* complex by NMR. Chemical shift perturbations in methyl-TROSY NMR (105) spectra of U-[^{15}N, ^{2}H]-Ile-δ1-[^{13}C, ^{1}H]-labeled $^{\Delta N}$KaiA in the presence of unlabeled KaiB* and CI* (Figure 2.4A, columns 1-2), also indicated complex formation, corroborating our chromatography results. Identical spectra were obtained when using KaiB instead of KaiB* to form the $^{\Delta N}$KaiA-KaiB-CI* complex (Figure 2.4A, column 3), suggesting that these interactions are not artifacts from using a mutant of KaiB. As further evidence that the N-terminal domains of KaiA are not necessary for KaiA-KaiB interactions, a full phosphorylation cycle was observed using $^{\Delta N}$KaiA instead of KaiA (Figure 2.6A-B). Thus, the recent model in which the N-terminal domains of KaiA are necessary for KaiA-KaiB binding (64), albeit in the *S. elongatus* system, may need to be revised.

Figure 2.2. Formation of $^{\Delta N}$KaiA-KaiB*-CI* and $^{\Delta N}$KaiA-KaiB* complexes.

(Figure is on the next page)

(A) Gel filtration profiles of $^{\Delta N}$KaiA alone (blue), KaiB* alone (green), $^{\Delta N}$KaiA + KaiB* (black), and KaiB* + CI* + $^{\Delta N}$KaiA (red) (*top left*). Control mixtures of KaiB* + CI* (purple) and CI* alone (orange) are shown for comparison (*bottom left*). Gel filtration profiles of KaiB* + CI* + $^{\Delta N}$KaiA with increasing amounts of $^{\Delta N}$KaiA added are shown on the *top right panel* (1×-red; 2×-cyan and 3×-black dash). Control mixtures of KaiB* + CI* (purple); CI* alone (orange); $^{\Delta N}$KaiA alone (blue) and KaiB* alone (green) are shown for comparison (*bottom right*). 1× corresponds to 50 μM monomer concentration. Chromatogram of $^{\Delta N}$KaiA + CI* is shown in Figure 2.3A. Elution positions of molecular weight markers are indicated by inverted triangles and molecular weights in kDa along the top of the gel filtration chromatograms. Peaks denoted by *a*, *b*, *c*, *d*, *e*, *f*, *g*, *h*, *I*, and *j* were checked by SDS/PAGE.

(B) SDS/PAGE gel of peaks *a*, *b*, *c*, *d*, *e*, *f*, *g*, *h*, *I*, and *j* in (A). All lanes shown were run on the same gel. The band marked "*#*" in lane *b* indicates the presence of $^{\Delta N}$KaiA in peak *b*. A comparison of peak *b* with peaks from other chromatograms (*f, h,* and *i*) suggests that the presence of $^{\Delta N}$KaiA in *b* was likely due to overlap of peaks *a* and *b* in the 1:1:1 $^{\Delta N}$KaiA:KaiB*:CI* chromatogram. The presence of a peak corresponding to free CI* in the 1:1:1, 2:1:1, and 3:1:1 $^{\Delta N}$KaiA:KaiB*:CI* chromatograms suggests that the true concentration of CI* was systematically underestimated. This underestimation however does not affect our analysis because CI* and $^{\Delta N}$KaiA do not interact with each other.

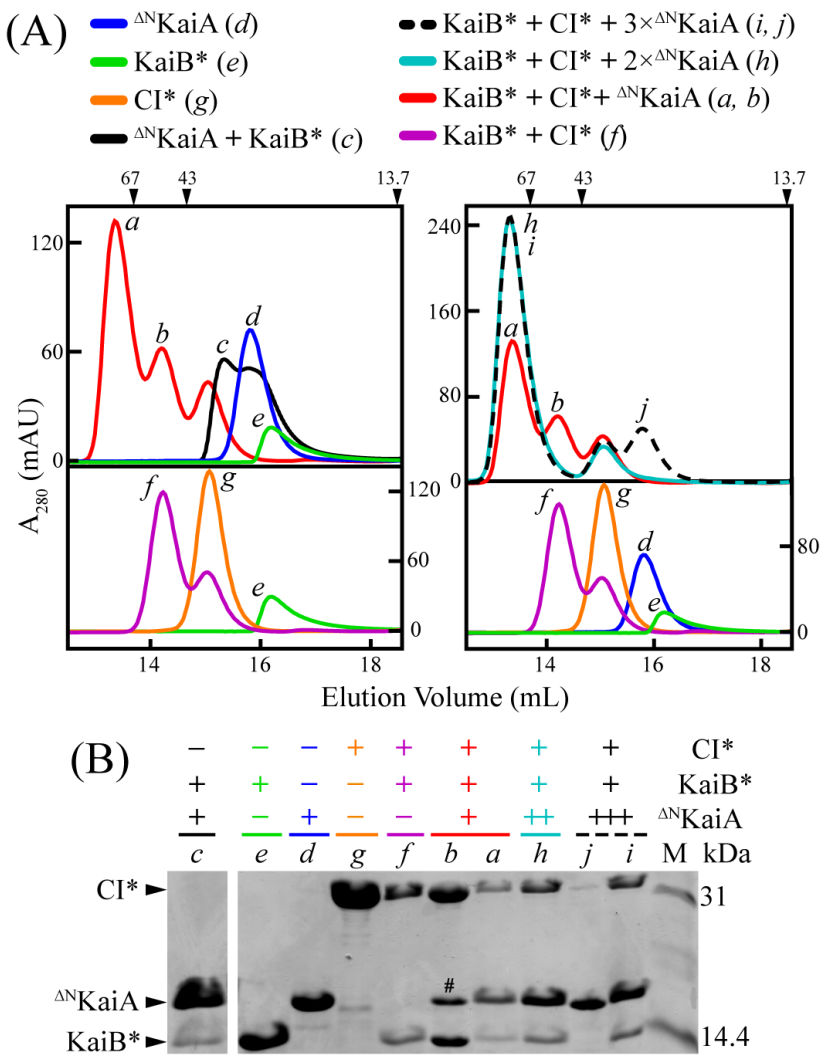


Figure 2.2. Formation of $^{\Delta N}$KaiA-KaiB*-CI* and $^{\Delta N}$KaiA-KaiB* complexes.

(Figure caption is on the previous page)

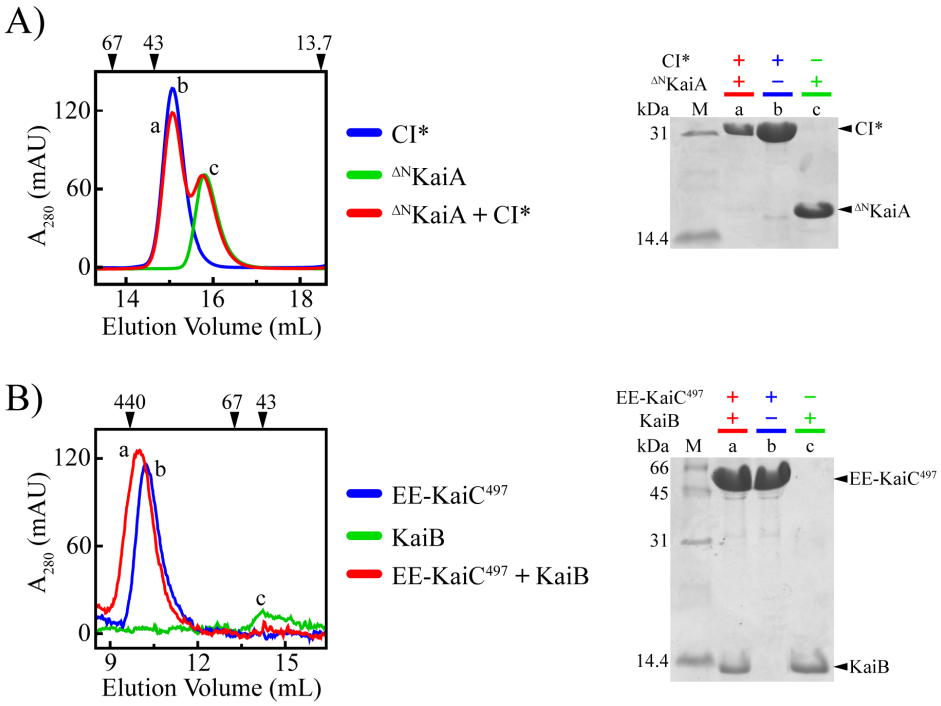


Figure 2.3. Gel filtration chromatograms for Figure 2.2.

(A) $^{\Delta N}$KaiA and CI* have no detectable interaction, and (B) wild-type KaiB interacts, as expected, with EE-KaiC497 (C-terminal truncation after residue I497). SDS/PAGE analysis of the chromatographic peaks in (A) and (B) are shown on the right.

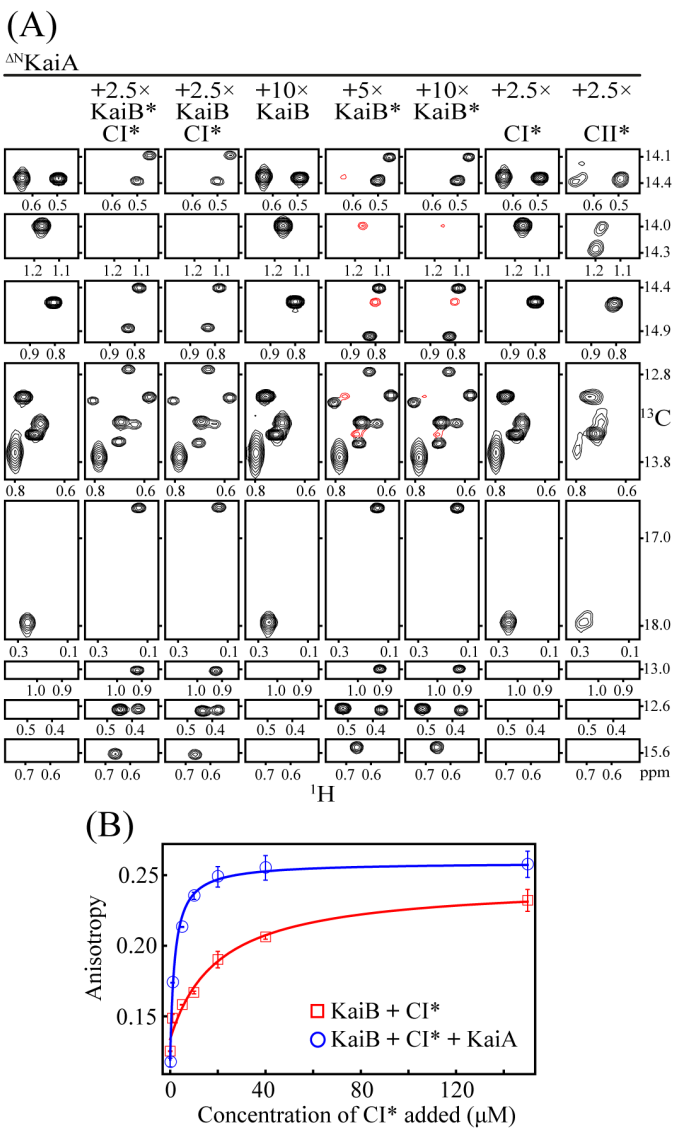


Figure 2.4. Formation of a $^{\Delta N}$KaiA-KaiB*-CI* complex is a cooperative process. (A) Selected regions from methyl-TROSY spectra of U-[^{15}N, ^{2}H]-Ile-δ1-[^{13}C, ^{1}H]–labeled 1× $^{\Delta N}$KaiA alone (column 1); or in the presence of 2.5× (KaiB* + CI*) (column 2); 2.5× (KaiB + CI*) (column 3); 10× KaiB (column 4); 5× KaiB* (column 5); 10× KaiB* (column 6); 2.5× CI* (column 7); or 2.5× CII* (column 8). 1× corresponds to 20 μM monomer concentration. NMR peaks marked in red are at positions that correspond to resonances in free $^{\Delta N}$KaiA (columns 4-5). Full spectra are shown in Figure 2.5. (B) Fluorescence anisotropy measurements of apparent dissociation constants, K_D^{app}, of 6-IAF labeled KaiB binding to CI* in the presence (blue) or absence (red) of 5 μM KaiA. Data were fit to equation [2]. The K_D^{app} values for CI* and CI*+KaiA were 2.0 ± 0.4 and 20.5 ± 1.0 μM, respectively. Standard error was estimated from duplicate measurements. Equilibrium prior to measurements was achieved by incubation of samples for 24 h (Figure 2.7A).

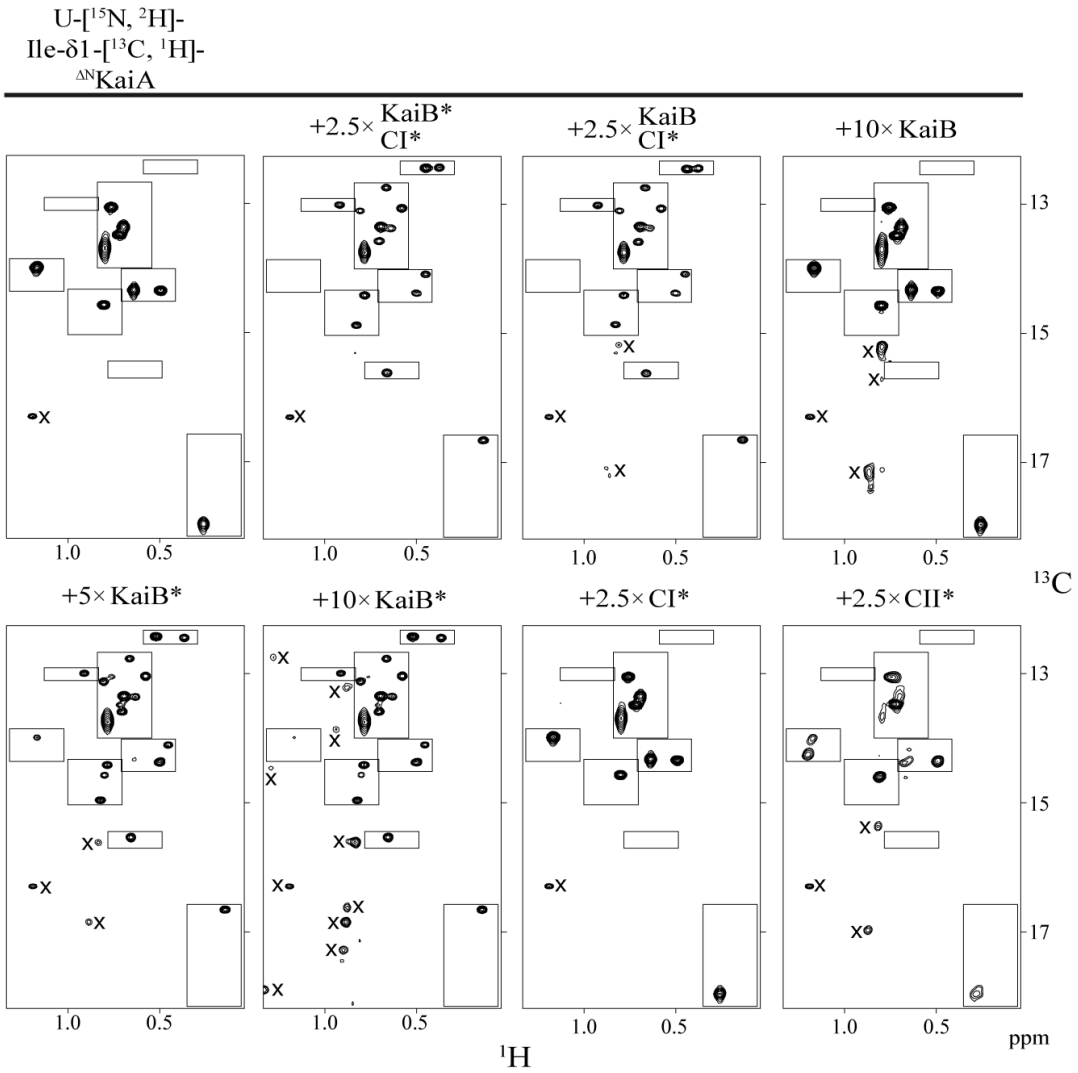


Figure 2.5. Full Methyl-TROSY spectra for Figure 2.4.

Methyl-TROSY spectra of U-[^{15}N, ^{2}H]-Ile-δ1-[^{13}C, ^{1}H]–labeled 1× $^{\Delta N}$KaiA alone (*top panel 1*), or in the presence of 2.5× (KaiB* + CI*) (*top panel 2*); 2.5× (KaiB + CI*) (*top panel 3*), 10× KaiB (*top panel 4*); 5× KaiB* (*bottom panel 1*); 10× KaiB (*bottom panel 2*); 2.5× CI* (*bottom panel 3*); or 2.5× CII* (*bottom panel 4*). 1× corresponds to 20 μM monomer concentration. Boxed regions are shown in Figure 2.4A. All spectra were recorded with identical parameters, processed identically, and plotted at the same contour level. “X”s in all spectra mark aliased peaks that were also present in spectra of unlabeled control samples (data not shown). Thus, they are not Ile-δ1 methyl peaks belonging to labeled $^{\Delta N}$KaiA.

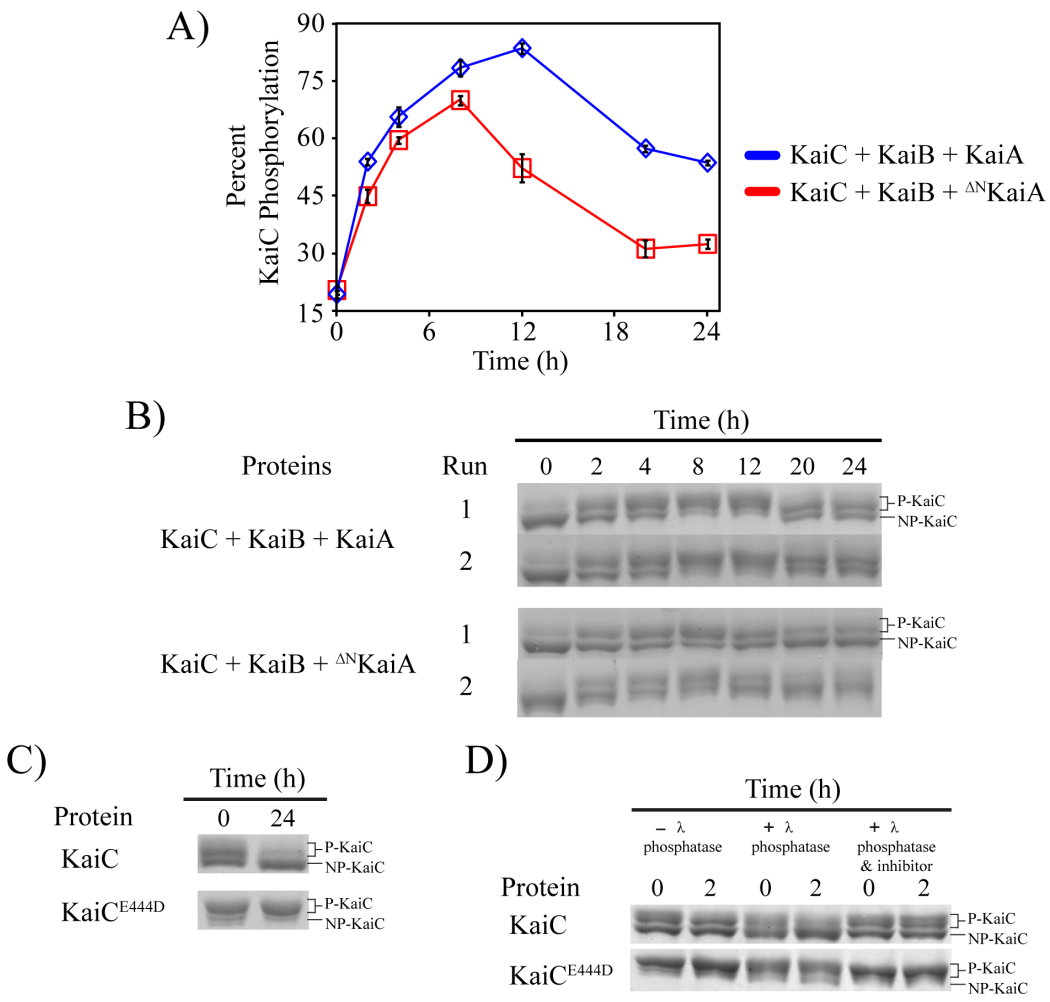


Figure 2.6. In vitro KaiC phosphorylation and dephosphorylation reactions.

(A) 24-hour phosphorylation profiles of wild-type KaiA + KaiB + KaiC (blue) and $^{\Delta N}$KaiA + KaiB + KaiC (red) mixtures. Standard error was estimated from duplicate runs. (B) SDS/PAGE of KaiC phosphorylation time points, which were used for densitometry analysis in (A). P-KaiC and NP-KaiC stand for phosphorylated and non-phosphorylated KaiC, respectively. (C) Dephosphorylation of KaiC and KaiC^{E444D} over 24 hours. (D) Dephosphorylation of KaiC and KaiC^{E444D} over 2 hours ± lambda phosphatase ± inhibitor.

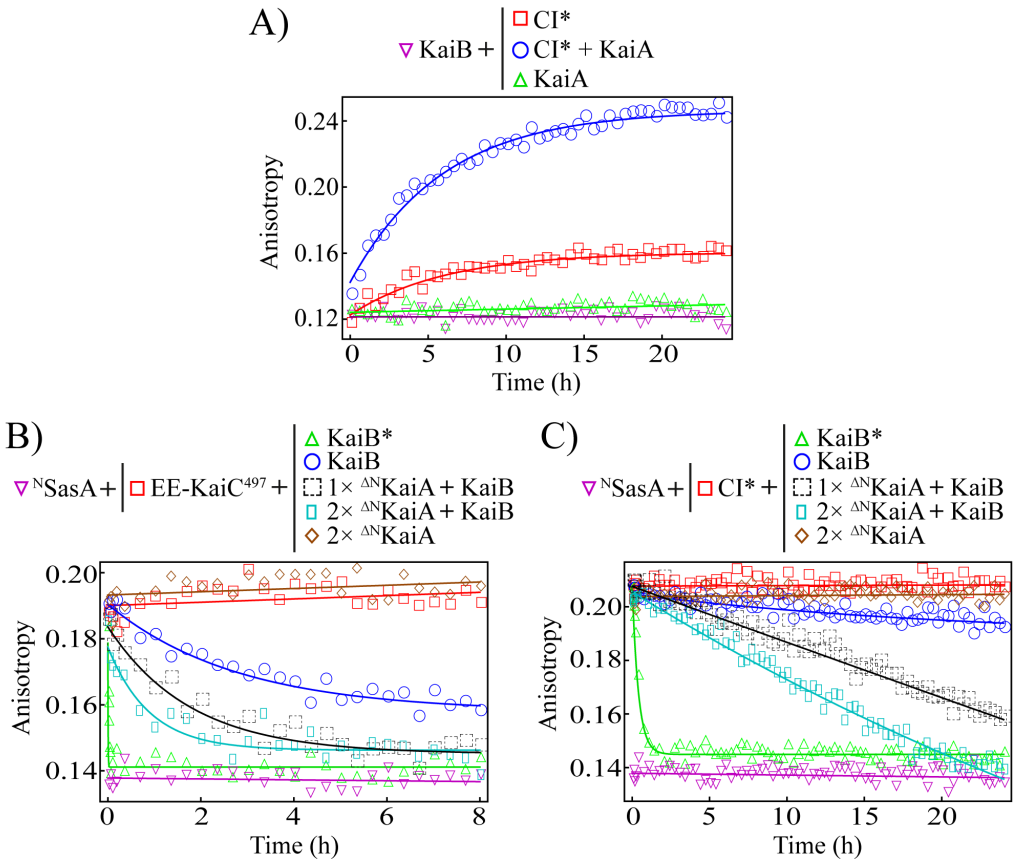


Figure 2.7. Fluorescence anisotropy measurements of protein binding kinetics.

(A) Kinetics of 6-IAF labeled KaiB binding to CI*, CI* + KaiA, and KaiA over 24 hours by fluorescence anisotropy. Data were fit to equation [3], yielding k_{obs} values (h-1) of 0.17 and 0.17 for CI* and CI* + KaiA, respectively. (B) and (C) are time courses of 6-IAF labeled ^{N}SasA-containing mixtures to which either EE-KaiC497 (left panel) or CI* (right panel) were added, as follows: buffer (red), 2× $^{\Delta N}$KaiA (brown), KaiB (blue), KaiB + 1× $^{\Delta N}$KaiA (black), KaiB + 2× $^{\Delta N}$KaiA (cyan), and KaiB* (green). Purple was used for data collected on a solution containing 6-IAF labeled ^{N}SasA as the sole protein component. Note that the only difference between (B) and (C) and Figure 2.11 is that $^{\Delta N}$KaiA was used instead of KaiA. Data were fit to equation [4], yielding k_{obs} values (h-1) for (B)/(C) as follows: 0.36/nd, 0.56/nd, 1.12/nd, and 179/2.47 for KaiB, KaiB + 1x $^{\Delta N}$KaiA, KaiB + 2x $^{\Delta N}$KaiA, and KaiB*, respectively (nd is not determined). 1× corresponds to 10 μM monomer concentration.

Previously, we showed that a construct of KaiA missing residues 1 – 179 could not be sequestered by KaiB (50), which together with the data here suggests that the linker connecting the N- and C-terminal domains, residues 147 – 179, plays a role in KaiA-KaiB interactions. $^{\Delta N}$KaiA was used for NMR experiments because of its smaller size, and labeled for fluorescence experiments because it retains only one of four naturally occurring cysteinyl residues (C272) for convenient labeling with thiol-reactive fluorophores.

Assembly of the KaiA-KaiB-KaiC complex is a cooperative process

Even after several hours of incubation there were no detectable interactions between $^{\Delta N}$KaiA and KaiB (Figure 2.4A, column 4). However, CI* significantly enhanced $^{\Delta N}$KaiA-KaiB binding (Figure 2.4A, column 3). Similarly, others have observed that KaiC enhances the otherwise weak interaction between KaiA and KaiB (42, 43, 80). In contrast to KaiB, KaiB* bound to $^{\Delta N}$KaiA even in the absence of CI* (Figure 2.2A, left panel; Figure 2.4A, columns 5-6). Recall that the stable forms of free KaiB and KaiB* are homotetramer (65, 106, 107) and homodimer (48), respectively. Based on our findings, we propose that the dimer is closer than the tetramer to the conformation that binds KaiC and KaiA. This idea is supported by recent findings that KaiB* binds more rapidly to CI* than does KaiB (48, 62).

As with KaiB, the weak $^{\Delta N}$KaiA-KaiB* interaction was significantly enhanced by CI*. Free $^{\Delta N}$KaiA or KaiB* were undetectable by gel-filtration chromatography in the presence of CI*, in contrast to when CI* was absent (Figure 2.2A, left panel). Similarly, NMR spectra of labeled $^{\Delta N}$KaiA showed complete KaiB* binding in the presence of CI* (Figure 2.4A, column 2), but incomplete binding without CI* (Figure 2.4A, columns 5-6). The absence of an observable interaction between just $^{\Delta N}$KaiA and CI* (Figure 2.4A, column 7 & Figure 2.3A) suggests that CI* facilitates KaiA-KaiB binding solely through interactions with KaiB. The observation that the majority of the NMR peaks of KaiB*-bound $^{\Delta N}$KaiA had the same chemical shifts in the presence and absence of CI* also supports this notion (Figure 2.4A, columns 2-3, 5-6 & Figure 2.3A). Although $^{\Delta N}$KaiA did not interact with CI*, NMR spectra showed, as expected, that $^{\Delta N}$KaiA interacted with CII*, a monomeric variant of the CII domain (50) (Figure 2.4A, column 8). These observations support the model in which KaiA has two distinct binding sites, the A-loop during KaiC phosphorylation (50), and KaiB during dephosphorylation (43, 48, 61).

Now that we have demonstrated that CI* facilitates KaiA-KaiB binding, we asked whether KaiA facilitates KaiB-KaiC binding. We measured the apparent dissociation constants, K_D^{app}, for the KaiB-CI* complex in the presence and absence of saturating amounts of KaiA (wild-type), by monitoring the

fluorescence anisotropy of labeled KaiB (Figure 2.4B). Indeed, we found that KaiA decreased the K_D^{app} almost 9-fold, from 22 ± 1 μM to 2.5 ± 0.4 μM, indicating positive cooperativity (cooperativity factor of 0.11 ± 0.02). Thus, our results suggest that the switch from KaiC phosphorylation to dephosphorylation is promoted by cooperative formation of the KaiA-KaiB-KaiC complex.

Estimating the stoichiometry of the KaiA-KaiB complex

So far, the structural details of the KaiA-KaiB complex in the dephosphorylation phase of KaiC remain largely unclear. The stoichiometry of the KaiA-KaiB complex was recently estimated by gel-filtration chromatography to be a dimer of KaiB bound to a dimer of KaiA (101). In a two-fold symmetric complex, each symmetry related pair of spin ½-labeled sites would produce a single observable NMR signal. However, the number of peaks in methyl-TROSY spectra of U-[^{15}N, ^{2}H]-Ile-δ1-[^{13}C, ^{1}H]-labeled $^{\Delta N}$KaiA almost doubled upon binding KaiB*, not counting the peaks that corresponded to unbound $^{\Delta N}$KaiA (Figure 2.5). In the spectrum of free $^{\Delta N}$KaiA there were nine methyl peaks, as was expected for a symmetric homodimer with nine isoleucinyl residues per subunit (108). However, 16 new $^{\Delta N}$KaiA peaks appeared upon adding unlabeled KaiB*. These new peaks persisted after forcing all $^{\Delta N}$KaiA into the complex by adding CI* (Figure 2.4A, column 1-2), indicating that the new peaks were from the bound form of $^{\Delta N}$KaiA. Peak doubling suggests that the symmetry of $^{\Delta N}$KaiA was broken by KaiB*. Although dramatic changes in the ^{15}N-TROSY spectra of [^{15}N, ^{2}H]-labeled KaiB* upon binding unlabeled $^{\Delta N}$KaiA suggested that KaiB* changed its structure globally, the number of peaks remained similar (Figure 2.8A). A model consistent with these observations is one in which KaiB binds as a monomer to one face of the $^{\Delta N}$KaiA dimer. This binding may distort the other face such that a second KaiB monomer cannot bind.

The idea of a monomeric subunit of KaiB binding to a KaiA dimer is also supported by size-exclusion chromatography (Figure 2.2A, right panel). We ran mixtures of $^{\Delta N}$KaiA:KaiB*:CI* at monomer molar ratios of 1:1:1, 2:1:1, and 3:1:1. At 1:1:1, KaiB*-CI and $^{\Delta N}$KaiA-KaiB*-CI complexes were both observed. However, in the 2:1:1 mixture, the KaiB*-KaiC complex disappeared, and like at 1:1:1 no free $^{\Delta N}$KaiA was detected. Only at the 3:1:1 ratio was free $^{\Delta N}$KaiA observed. Together, the data suggest that one subunit of KaiB binds one subunit of CI, and two subunits of KaiA at the C-terminal and linker regions. The hydrophobic nature of the dimer interface of KaiA (109) makes it improbable that KaiA disassociates into monomeric subunits with that interface exposed (64).

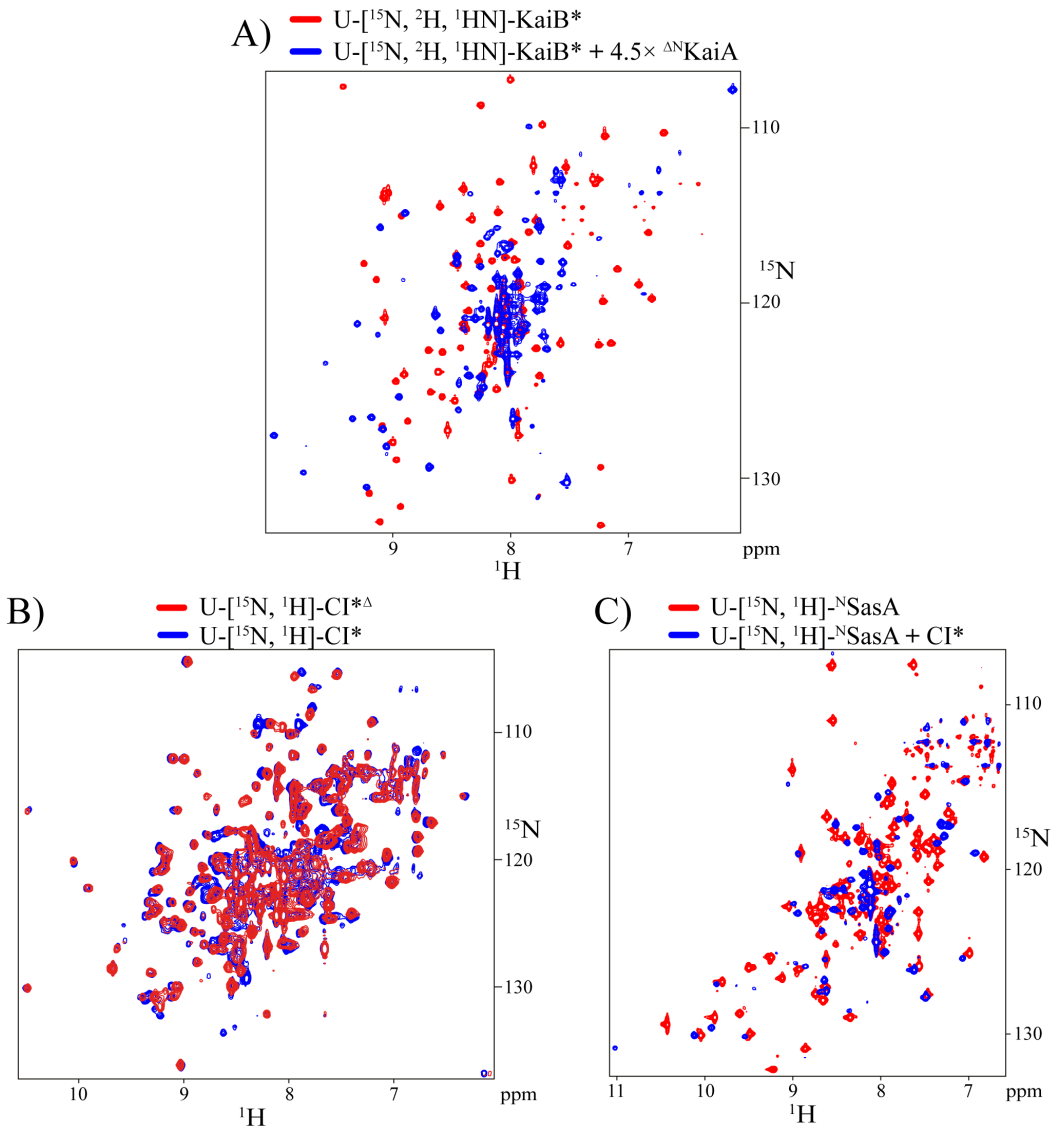


Figure 2.8. ^{15}N TROSY spectra of KaiB*, CI*$^{\Delta}$, CI* and ^{N}SasA.

(A) Overlay of ^{15}N, ^{1}H-TROSY spectra of U-[^{15}N, ^{2}H, ^{1}H^{N}]-KaiB* free, with approximately 94 peaks (red), and in complex with 4.5× unlabeled $^{\Delta N}$KaiA, with approximately with 87 peaks (blue). There are 95 possible peaks for KaiB* (102 minus 7 prolines). 1× corresponds to 150 μM monomer concentration. No peaks corresponding to free KaiB* were observed under this saturating condition. (B) Overlay of ^{15}N, ^{1}H-TROSY spectra of U-[^{15}N, ^{1}H] labeled CI*$^{\Delta}$ (red) and CI* (blue), demonstrating that the "B-loop" deletion mutant CI*$^{\Delta}$ is still well folded and similar to CI*. (C) Overlay of ^{15}N, ^{1}H-TROSY spectra of U-[^{15}N, ^{1}H] labeled ^{N}SasA free (red) and bound to CI* (blue).

Phosphoryl-S431 of KaiC imposes negative feedback on the KaiA-KaiC interaction

KaiA stimulates KaiC phosphorylation by binding to exposed A-loops (50, 58). Thus, dephosphorylation of KaiC requires sequestration of KaiA away from the A-loops in a KaiA-KaiB-KaiC complex. Interestingly, there seems to be a KaiB-independent negative feedback mechanism on KaiA-KaiC interactions that may facilitate sequestration. The affinity of KaiA for unphosphorylated KaiC is higher than that for phosphorylated KaiC (110, 111). A mathematical model has suggested that such negative feedback could play a role in maintaining phase coherence across an ensemble of Kai proteins (112). The high-resolution structure of a complex between KaiA and its binding site on KaiC, which includes the A-loop residues (residues 487 – 497), suggested that regulation of the A-loops between a KaiA-accessible exposed position and KaiA-inaccessible buried position (as found in the crystal structure (47)) is a possible mechanism for negative feedback (50). However, since the A-loops have only been observed in their buried position, the idea of a dynamic equilibrium between buried and exposed positions needed to be tested. As such, A-loop exposure was gauged by the kinetics of their proteolysis, k_{obs}, at a thrombin cut site (LVPRGS) that replaced a naturally occurring stretch of residues, I497-K502, immediately adjacent to the A-loop. We reasoned that the accessibility of the proteolysis site would depend significantly on the dynamic equilibrium of the A-loops between buried and exposed positions. We inserted the thrombin cut site in XY-KaiC variants where X and Y, at positions 431 and 432, respectively, were substituted with alanyl/glutamyl residues to mimic the unphosphorylated/phosphorylated states of S431 and T432 of KaiC. Thus, AA-, AE-, EE-, and EA-KaiC variants mimicked, respectively, ST- (unphosphorylated), SpT- (only T432 phosphorylated), pSpT- (both residues phosphorylated), and pST-KaiC (only S431 phosphorylated) phosphoforms (Figure 2.1B). Several studies have shown that these KaiC mutants are reasonably faithful mimics of the naturally occurring phosphoforms (95, 113, 114). Each of the four phosphomimics of KaiC containing the thrombin cut site was incubated with the protease under identical conditions. Time points were analyzed by SDS PAGE. EA- and EE-KaiC had slower rates of proteolysis, i.e., smaller values of k_{obs}, than AA- and AE-KaiC (Figure 2.9 & Figure 2.10A-B), suggesting that the A-loops made fewer excursions from the buried to exposed position in the pST- and pSpT-KaiC phosphomimics. Thus, phosphorylation at S431, which tightens the CII ring (52), likely shifts the A-loop equilibrium toward the buried position, where the A-loops can form a ring of hydrogen-bonding interactions that cooperate with phosphoryl-S431 to tighten the CII ring (50).

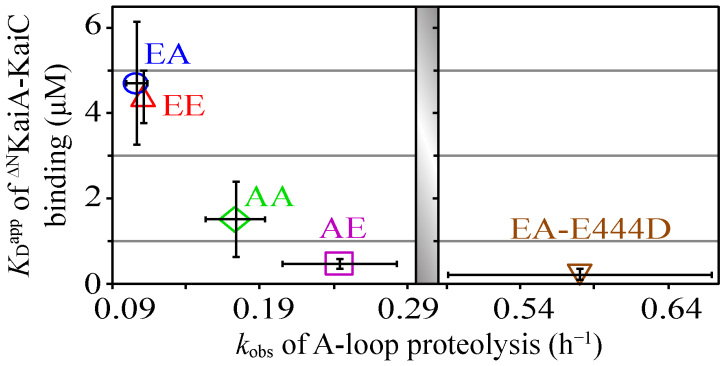


Figure 2.9. A-loop exposure is correlated with KaiA-KaiC affinity.

Kinetics of A-loop proteolysis of five variants of KaiC (AA-KaiC, green diamond; AE-KaiC, purple square; EE-KaiC, red triangle; EA-KaiC, blue circle; EA-KaiC^{E444D}, brown inverted triangle) with a thrombin cut site (LVPRGS) replacing residues I497 – K502. A-loop proteolysis rates were determined by fitting time points resolved by SDS/PAGE (Figure 2.10A-B) to equation [5]. K_{obs} for AA-, AE-, EE-, EA- and EA-KaiC^{E444D} were 0.17 ± 0.02, 0.24 ± 0.04, 0.111 ± 0.001, 0.106 ± 0.007 and 0.58 ± 0.09 h^{-1}, respectively. $^{\Delta N}$KaiA-KaiC affinity, K_D^{app}, was determined from the dependence of fluorescence anisotropy of 6-IAF labeled $^{\Delta N}$KaiA on the concentration of the same five 39hosphomimics variants of KaiC, but without the thrombin cut sites. The K_D^{app} values, obtained from fitting the data to equation [1], for AA-, AE-, EE-, EA- and EA-KaiC^{E444D} were 1.5 ± 0.9, 0.5 ± 0.1, 4.4 ± 0.6, 4.7 ± 1.4 and 0.2 ± 0.1 μM, respectively. Standard errors for k_{obs} and K_D^{app} were estimated from duplicate measurements as shown in Figure 2.10.

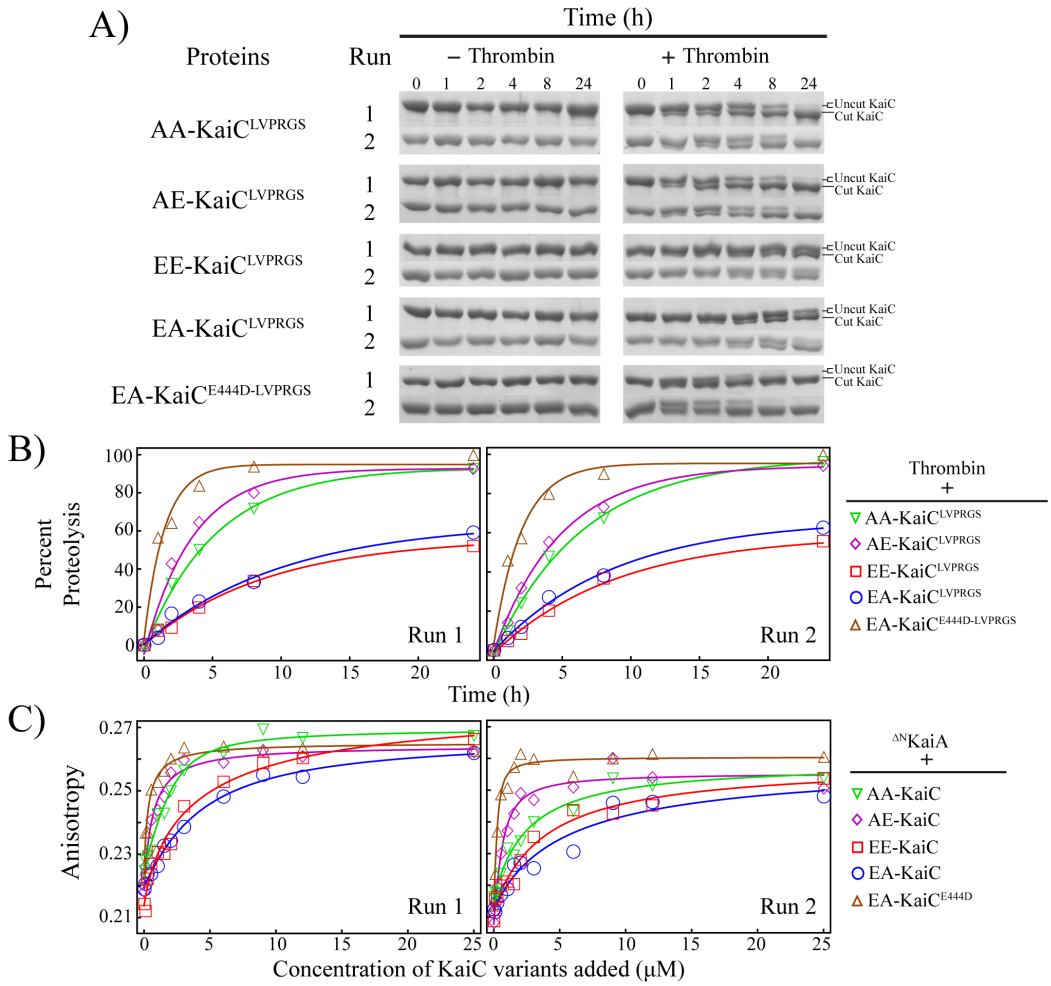


Figure 2.10. Raw data sets for Figure 2.9.

(A) A-loop proteolysis resolved by SDS/PAGE of five variants of KaiC (AA-KaiC, AE-KaiC, EE-KaiC, EA-KaiC, and EA-KaiC^{E444D}) with a thrombin cut site (LVPRGS) replacing residues I497 to K502. Twenty-four hour runs +/- thrombin were performed in duplicate. (B) Percent proteolysis determined by densitometry of the bands in (A). k_{obs} values were determined by fitting the data to equation [5] (C) Fluorescence anisotropy measurements of 6-IAF labeled $^{\Delta N}$KaiA interacting with five variants of KaiC (AA-KaiC, green; AE-KaiC, purple; EE-KaiC, red; EA-KaiC, blue; EA-KaiC^{E444D}, brown). Duplicate runs were performed as shown. Apparent dissociation constants, K_D^{app}, for $^{\Delta N}$KaiA-KaiC complexes were determined by fitting to equation [1]

Because the A-loops comprise ⅓ of the KaiA-binding site (49, 50) (the other 2/3 being the C-terminal tails of the CII domains of KaiC), promoting A-loop burial is predicted to weaken KaiA-KaiC binding. Thus, we also measured apparent K_D values, K_D^{app}, of the KaiA-KaiC interaction, as a function of concentration of KaiC phosphomimics, using fluorescence anisotropy of labeled $^{\Delta N}$KaiA proteins (Figure 2.9 & Figure 2.10C). As shown in Figure 2.9, the affinity decreases 10-fold from AE-KaiC to EA-KaiC. Although this change is modest, it should be noted that similarly modest changes were sufficient to reproduce KaiC phosphorylation rhythms in silico (111). In Figure 2.9, a trend was observed with smaller k_{obs} correlating with larger K_D^{app}, suggesting that the KaiA-stimulated phosphorylation of S431 imposes negative feedback by promoting burial of part of the KaiA-binding site on KaiC. Using just two phosphomimics (AA-KaiC and EE-KaiC), Qin et al (43) reported an opposite trend in KaiA-KaiC affinities using *S. elongatus* proteins. Thus, it was important that we validated our results. E444 appears to stabilize the buried form of the A-loops through side chain hydrogen bonds. Because destabilizing the buried position of the A-loops activates the autokinase activity of KaiC (50), we could gauge the affect of the E444D substitution on the A-loops by its affect on KaiC phosphorylation. In fact, the E444D substitution produced constitutively hyperphosphorylated mutants in both *S. elongatus* (50) and *T. elongatus* KaiC proteins (Figure 2.6C-D). To verify whether E444D indeed acted by destabilizing the A-loops, we measured k_{obs} and K_D^{app} on EA-KaiC containing the E444D substitution. Relative to EA-KaiC, for EA-KaiC^{E444D} we found a complete reversal of A-loop burial ⇔ exposure, and KaiA-KaiC affinities (Figure 2.9). These data suggest that phosphoryl S431-dependent burial of the A-loops imposes negative feedback regulation on the KaiA-KaiC interaction. This allosteric, KaiB-independent mechanism during the later part of the phosphorylation phase is likely to promote formation of the KaiA-KaiB-KaiC dephosphorylation complex.

KaiA-KaiB-KaiC cooperativity contributes to clock output

In order for circadian clocks to synchronize cellular processes to daily swings in ambient light and temperature, mechanisms must exist to transmit temporal information from the oscillator to other pathways. In cyanobacteria, two mutually antagonistic output pathways, initiated by the SasA and CikA proteins, transduce KaiC phosphorylation rhythms into genome-wide transcription rhythms (34, 40). SasA phosphorylates/activates the transcription factor RpaA, whereas CikA dephosphorylates/deactivates RpaA. SasA is a sensor histidine kinase (33) that, upon binding KaiC, phosphorylates (97), and then transfers the phosphoryl group to the transcription factor RpaA, thereby activating it (34, 40, 44). Once activated, RpaA induces transcription of genes such as *kaiBC*. In this way, circadian rhythms of SasA activity drive genome-wide transcription rhythms. SasA binds directly to KaiC using its N-terminal domain (33). Recently, it was found that KaiB and SasA compete for binding to KaiC (62, 64), suggesting that KaiB plays a direct role in regulating clock output (62). Since the formation of the KaiA-KaiB-KaiC complex is cooperative, we wanted to determine whether KaiA also plays a role in regulating SasA-KaiC interactions. We set up competition experiments between the isolated N-terminal domain of SasA, ^{N}SasA (68), and KaiB (or KaiB*) ± KaiA (or ± $^{\Delta N}$KaiA) (Figure 2.11 & Figure 2.7B-C). In order to minimize KaiA binding to the CII side (50, 83), we used a construct of EE-KaiC missing its C-terminal extensions, EE-KaiC^{497} (Figure 2.1A) (52). We also tested ^{N}SasA and KaiB competition for binding CI*, which was recently identified as their binding site (48, 52, 67). Because ^{N}SasA has no naturally occurring cysteinyl residues, a G57C substitution allowed labeling with a thiol-reactive fluorophore, so that competition could be monitored using fluorescence anisotropy. As seen in Figure 2.11 and Figure 2.7B-C, the fluorescence anisotropy of free ^{N}SasA was much lower than when it was bound to EE-KaiC^{497} or CI*, providing sufficient dynamic range for competition experiments. KaiB displaced ^{N}SasA slowly from EE-KaiC^{497} and CI*. The displacement was significantly faster in the presence of increasing concentration of KaiA [or $^{\Delta N}$KaiA (Figure 2.7B-C)]. This observation demonstrates that KaiA enhances the competitiveness of KaiB over ^{N}SasA for KaiC. Therefore, through its role in the cooperativity of KaiA-KaiB-KaiC interactions, KaiA appears to modulate SasA-KaiC interactions, with implications for regulating clock output. Since KaiB* binds KaiC much faster than KaiB (48), it was not surprising that, even in the absence of KaiA, KaiB* quickly displaced ^{N}SasA (Figure 2.11 & Figure 2.7B-C).

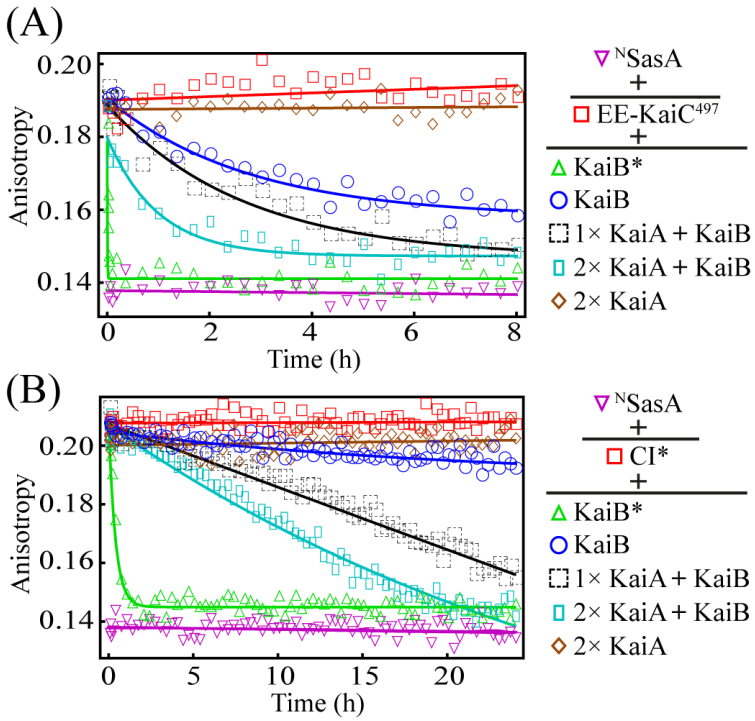


Figure 2.11. KaiA as a modulator of KaiB/SasA competition on KaiC and CI.

(A) Fluorescence anisotropy kinetics of 6-IAF labeled ^{N}SasA alone (purple), ^{N}SasA + EE-KaiC^{497} (red), ^{N}SasA + EE-KaiC^{497} + 2× KaiA (brown), ^{N}SasA + EE-KaiC^{497} + KaiB (blue), ^{N}SasA + EE-KaiC^{497} + KaiB + 1× KaiA (black), ^{N}SasA + EE-KaiC^{497} + KaiB + 2× KaiA (cyan), and ^{N}SasA + EE-KaiC^{497} + KaiB* (green). 1× corresponds to 10 μM monomer concentration. (B) Fluorescence anisotropy of competition kinetics of 6-IAF labeled ^{N}SasA alone (purple), ^{N}SasA + CI* (red), ^{N}SasA + CI* + 2× KaiA (brown), ^{N}SasA + CI* + KaiB (blue), ^{N}SasA + CI* + KaiB + 1× KaiA (black), ^{N}SasA + CI* + KaiB + 2× KaiA (cyan), and ^{N}SasA + CI* + KaiB* (green). 1× corresponds to 10 μM monomer concentration. Differences in the anisotropies between (A) and (B) suggest that the hexameric EE-KaiC^{497} and monomeric CI* do not interact identically with ^{N}SasA and KaiB.

SasA and KaiB bind to B-loops on CI

To better understand the basis of the competition between SasA and KaiB, we investigated whether they shared binding elements on CI. This identification was aided by sequence alignments between the highly similar (42) RecA-like (81) CI and CII domains of KaiC. As seen in Fig. 2.12a, CI contains an insertion of several residues (116-123) not found in CII. This insertion is exposed on the bottom of CI as part of a loop, which we named the B-loop (Figure 2.1A). The construct, $CI^{*\Delta}$ (Figure 2.1A), missing the B-loop insertion, retained the structure of CI*, as determined from a comparison of their NMR fingerprints (Figure 2.8B). However, unlike with CI*, $CI^{*\Delta}$ could not bind ^{N}SasA or KaiB* (Figure 2.12B). This result strongly suggests that the B-loop forms part of the binding site for both proteins. Although the binding sites of ^{N}SasA and KaiB on CI overlap, their interactions with CI are not identical, as suggested by a comparison of chemical shift perturbations induced by ^{N}SasA and KaiB* in methyl-TROSY spectra of U-[^{15}N, ^{2}H]-Ile-δ1-[^{13}C, ^{1}H]-labeled CI* (Figure 2.13, row 1-3).

Perturbations of CI* spectra by KaiB* were not significantly perturbed further by the addition of $^{\Delta N}$KaiA, and $^{\Delta N}$KaiA by itself did not perturb the spectrum of CI* (Figure 2.13, columns 4-5), providing further support for the notion that sequestration of KaiA on CI likely does not involve direct KaiA-KaiC interactions, but occurs indirectly through KaiA-KaiB interactions. Furthermore, it also suggests that CI adopts the same conformation in the binary KaiB-KaiC and ternary KaiA-KaiB-KaiC complexes. These observations contradict recent studies suggesting that KaiB and SasA bind to the CII side of KaiC (57, 58).

Figure 2.12. ^{N}SasA and KaiB* compete on CI near the B-loops.

(Figure is on the next page)

(A) CLUSTAL-W multiple sequence alignment of the two domains of KaiC from different species of cyanobacteria reveals an insertion in CI missing in CII. The insertion, termed the “B-loop”, is from residues 101 – 135 (highlighted in yellow). For the sequence alignment: BP-1, *Thermosynechococcus elongatus* BP-1 (used in this study); PCC-7942, *S. elongatus* PCC-7942; PCC-8801*, Synechococcus* sp*.* PCC-8801; PCC-9709, *Nostoc* sp. PCC-9709; IAM M-101, *L. boryana* IAM M-101; FACHB-438, *A. maxima* FACHB-438; PCC-7806, *M. aeruginosa* PCC-7806.

(B) Gel filtration profiles of CI* + KaiB* (red, from Figure 2.1A); CI* + ^{N}SasA (brown); CI*$^{\Delta}$ + KaiB* (orange); and CI*$^{\Delta}$ + ^{N}SasA (cyan). CI*$^{\Delta}$ is with B-loop (residues 116-123) deleted. Peaks denoted by *a*, *b*, *c*, *d*, *e*, *f*, *g* and *h* were checked by SDS/PAGE. Molecular weight markers in kDa are marked by black inverted triangles along the top of the gel filtration chromatograms.

(C) SDS/PAGE gel of peaks *a*, *b*, *c*, *d*, *e*, *f*, *g* and *h* in (B).

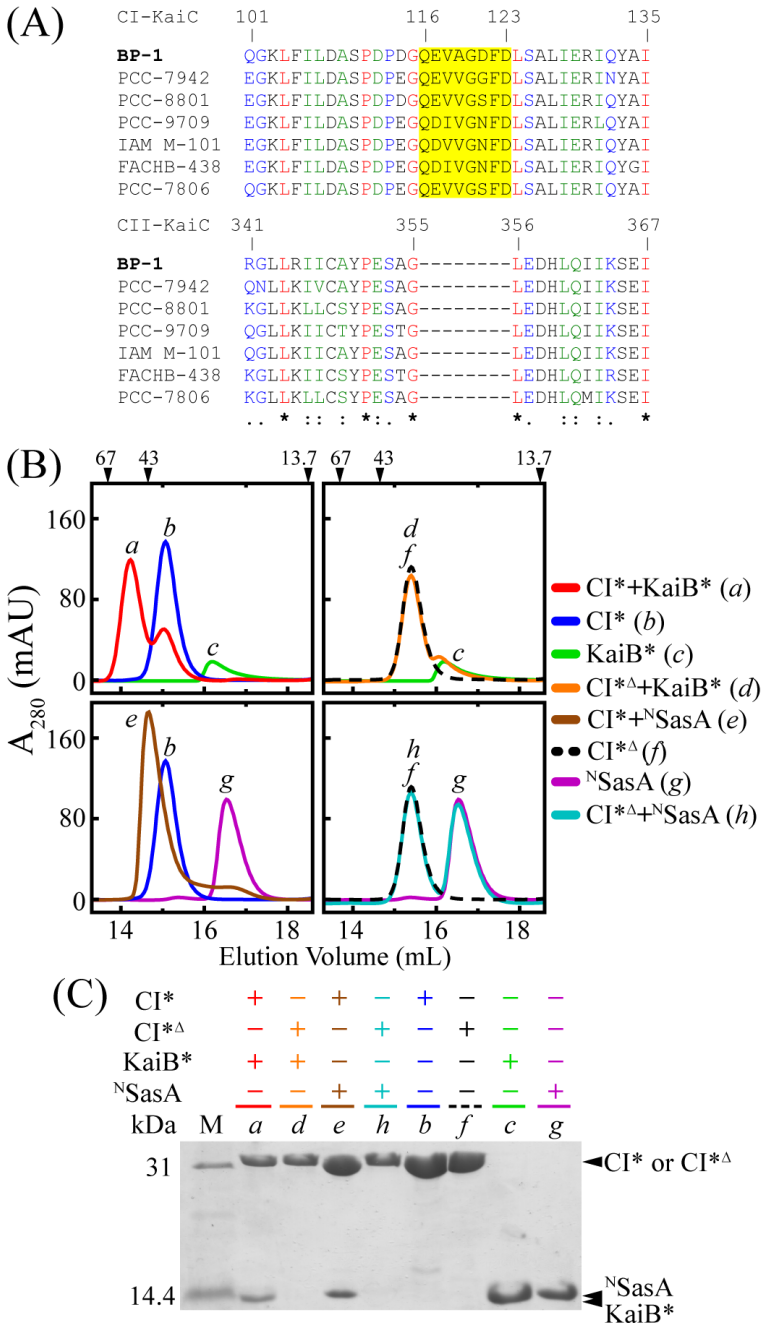


Figure 2.12. ^{N}SasA and KaiB* compete on CI near the B-loops.

(Figure caption is on the previous page)

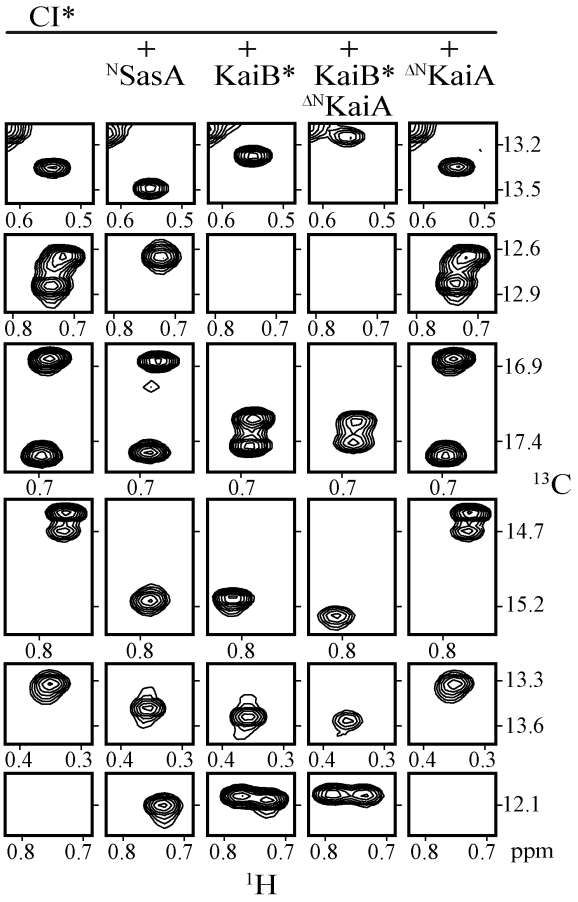


Figure 2.13. ^{N}SasA and KaiB* do not interact identically with CI.

Selected regions from methyl-TROSY spectra of U-[^{15}N, ^{2}H]-Ile-δ1-[^{13}C, ^{1}H]–labeled CI* alone (column 1), in the presence of ^{N}SasA (column 2), KaiB* (column 3), KaiB* + $^{\Delta N}$KaiA (column 4), and $^{\Delta N}$KaiA (column 5). Full spectra are shown in Figure 2.14.

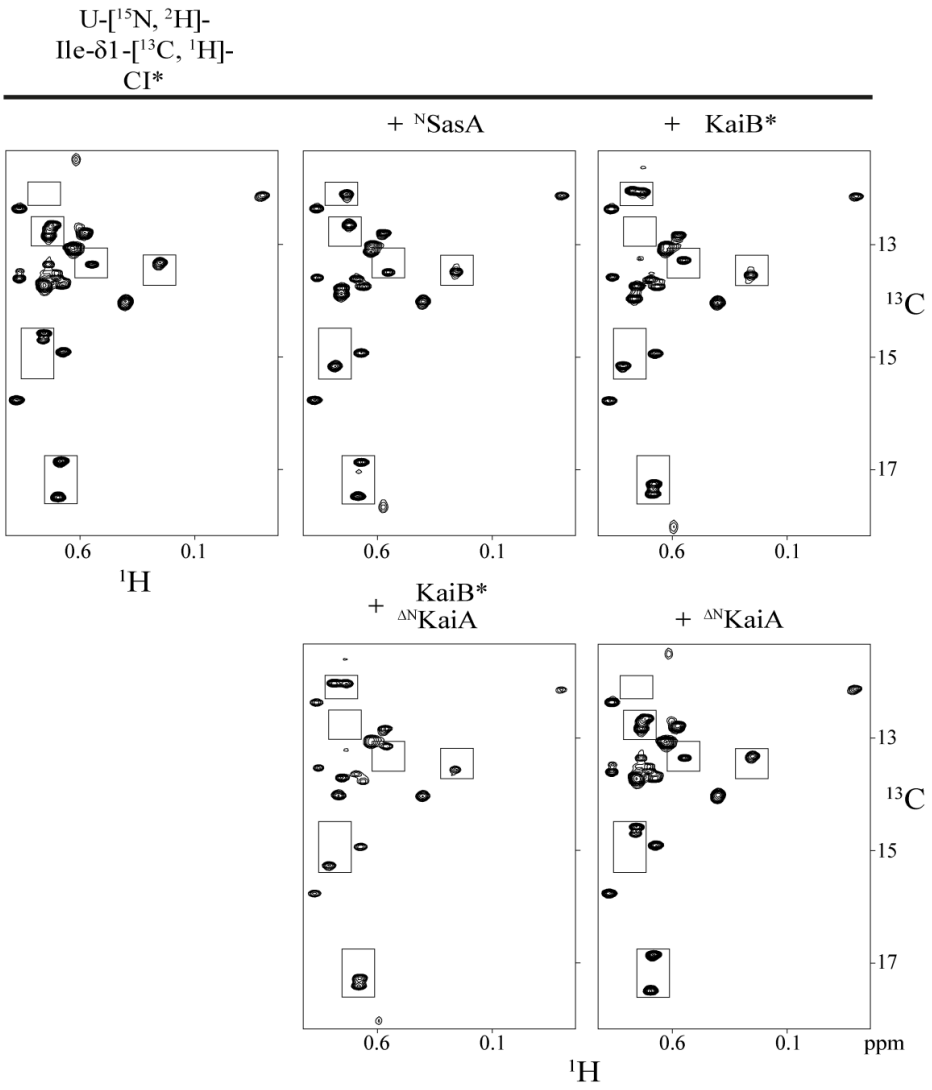


Figure 2.14. Full Methyl-TROSY spectra for Figure 2.13.

Methyl-TROSY spectra of U-[^{15}N, ^{2}H]-Ile-δ1-[^{13}C, ^{1}H]–labeled CI* alone (*top panel 1*), or in the presence of ^{N}SasA (*top panel 2*), KaiB* (*top panel 3*), KaiB* + $^{\Delta N}$KaiA (*bottom panel 1*), or $^{\Delta N}$KaiA (*bottom panel 2*). Boxed regions are shown Figure 2.13. All spectra were recorded with identical parameters, were processed identically and plotted at the same contour level.

Discussion

Based upon our findings, we think that the most stable form of KaiB, the homotetramer (65, 106, 107), cannot bind KaiA or KaiC, but is in a dynamic equilibrium with transiently disassociated subunits that can bind. Our previous demonstration of slow KaiB-KaiC binding, faster KaiB*-KaiC binding, and subunit shuffling between KaiB proteins supports this notion (48). Significant changes in NMR spectra of KaiB* after hours of incubation with KaiA (Figure 2.8A) also indicate that large changes in the tertiary structure of KaiB occur upon binding, reminiscent of KaiB-KaiC binding (48). KaiA speeds up KaiB-KaiC binding by shifting KaiB away from the tetramer and toward KaiB subunits active in binding KaiC. In the same way, KaiC enhances the KaiA-KaiB interaction. Thus, formation of the KaiA-KaiB-KaiC dephosphorylation complex on the CI side of KaiC is cooperative. This cooperativity is probably aided by KaiB-independent negative feedback imposed on the KaiA-KaiC interaction on the CII side of KaiC toward the end of the phosphorylation phase. We present our model of the KaiA-KaiB-KaiC interaction in Figure 2.15.

Based on EM data, it has been suggested that the competition between SasA and KaiB is on the CII domain of KaiC (58), instead of CI as we propose. The low resolution of the EM data of KaiC makes it difficult to unambiguously distinguish the CI from CII domains, in our opinion. In support of CI binding, another group has independently verified that SasA binds to CI (67). Our identification of the B-loop of CI as a common element of the KaiB and SasA binding sites adds further support that these interactions take place on CI. Also, NMR spectra of the KaiB*-CI* and ^{N}SasA-CI* complexes (Figure 2.13, Figure 2.8C & Figure 2.14) argue against artifactual interactions.

Since our data suggest that a single subunit of KaiB binds to a dimer of KaiA, it was important to see how well this stoichiometry agreed with published results, using back-of-the-envelope calculations. It has been shown that KaiB complexation (KaiA-KaiB-KaiC and KaiB-KaiC complexes) lags KaiC phosphorylation by 4 to 8 hours (80, 111, 115). When these complexes reach peak levels (at t_{max} in Figure 2.16), ~60% of total KaiB is bound (80). A new cycle of phosphorylation does not initiate until at least half of these complexes have disassociated (at $t_{1/2}$ in Figure 2.16A) (80). Thus, from $t_{max} \rightarrow t_{1/2}$, KaiA that is desequestered from decomposition of KaiA-KaiB-KaiC complexes would need to be resequestered by intact KaiBC complexes to prevent premature phosphorylation. Using the known dependency of the phosphorylation rhythm on KaiA and KaiB concentrations (116), by our estimation resequestration is possible when using a stoichiometry of one monomeric subunit of KaiB binding one dimer of KaiA (Table 2.2 & Figure 2.16). In contrast, resequestration of

desequestered KaiA from $t_{max} \rightarrow t_{1/2}$ was not possible when we assumed that a dimer of KaiB sequestered a KaiA dimer (101) or monomer (64). Thus, the literature lends support to our model. In conclusion, the circadian oscillator of cyanobacteria uses cooperativity to reciprocally regulate periodic formation of complexes on the CI and CII rings of KaiC. This cooperativity likely influences downstream output signaling as well. Thus, understanding the dynamic and allosteric mechanism of the oscillator is an important first step in elucidating how it transduces clock output signals.

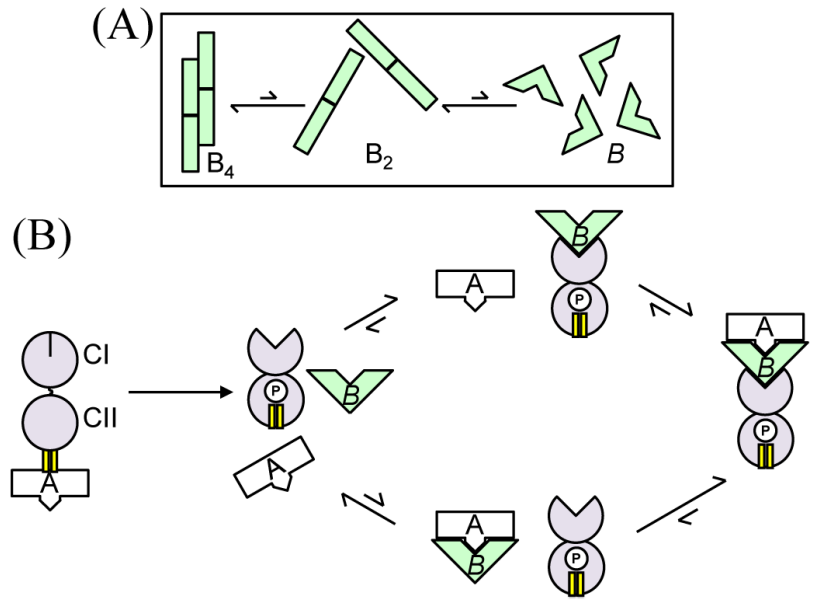


Figure 2.15. Model of cooperative formation of the KaiABC dephosphorylation complex.

(A) KaiB exists in a dynamic equilibrium between distinct quaternary states: tetramer (B_4), dimer (B_2), and active monomer (B). B_4 is the most stable form (65, 106, 107), whereas *B* is a highly transient state. (B) The CI domain of KaiC and KaiA can each stabilize the highly active monomer of KaiB, *B*, thereby cooperatively forming the KaiABC dephosphorylation complex. A-loops of KaiC are depicted as yellow bars. P indicates phosphorylation at S431 of KaiC. Upon S431 phosphorylation, two major events occur in KaiC: 1) The A-loops are shifted toward their buried position, due to CII-ring tightening (52), and KaiA consequently loses affinity for the CII domain. 2) The CI and CII domains of KaiC stack together, exposing the KaiB-binding site on CI (48). Now, the CI domain can selectively capture the highly transient *B*, enhancing KaiA sequestration. Alternatively, KaiA can selectively capture *B*, thereby promoting binding to KaiC.

Table 2.2. Back-of-envelope calculation of KaiA-KaiB complex stoichiometry compared with published experimental results.

Original experimental setup from Nakajima et al. 2010 (116) (μM monomer concentration).	KaiB:KaiA:KaiC 3.5 : 1.2 : 3.5	KaiB:KaiA:KaiC 3.5 : 2.4 : 3.5	KaiB:KaiA:KaiC 1.8 : 1.2 : 3.5	KaiB:KaiA:KaiC 3.5 : 3.6 : 3.5
Nakajima et al. 2010 (116) setup converted to dimer KaiA and hexamer KaiC (μM concentration).	KaiB:$KaiA_2$:$KaiC_6$ 3.5 : 0.6 : 0.6	KaiB:$KaiA_2$:$KaiC_6$ 3.5 : 1.2 : 0.6	KaiB:$KaiA_2$:$KaiC_6$ 1.8 : 0.6 : 0.6	KaiB:$KaiA_2$:$KaiC_6$ 3.5 : 1.8 : 0.6
Experimental result from Nakajima et al. 2010 (116).	Rhythmic	Rhythmic	Rhythmic	Arrhythmic
Assuming a monomer of KaiB sequesters a dimer of KaiA.	Rhythmic, Figure 2.16B	Rhythmic, Figure 2.16C	Rhythmic, Figure 2.16D	Arrhythmic, Figure 2.16E
Assuming a dimer KaiB sequesters a dimer of KaiA.	Rhythmic, Figure 2.16F	Arrhythmic, Figure 2.16G	Arrhythmic, not shown	Arrhythmic, not shown

Figure 2.16. Back-of-envelope calculations.

(Figure is on the next page)

(A) Adapted from Fig. 3d in Goda et al. 2012 (80) (*left panel*). Red oscillation is percent KaiC phosphorylation, and blue oscillation is percent KaiB complexation (KaiBC and KaiABC). T_{max} marks the time of maximal complexation of KaiB, i.e., ~60% of total KaiB (80). $T_{1/2}$ is the halfway point of disassociation of KaiB complexes, i.e., ~35% of total KaiB (80). $T_{1/2}$ also corresponds to the lowest point of KaiC phosphorylation (80). The goal is to estimate the level of desequestered KaiA [$KaiA_{2\ free}$] relative to that of [KaiBC] from $t_{max} \rightarrow t_{1/2}$, because (A) suggests that to be rhythmic desequestered KaiA needs to be resequestered during $t_{max} \rightarrow t_{1/2}$. Initial conditions are set at t_0.

The box in (A) illustrates our approach, which is applied to panels (B) – (G). KaiB:$KaiA_2$:$KaiC_6$ at t_0 were taken from Nakajima et al. 2010 (116) (see Table 2.2). Generally, $KaiB_{complex}$ at t_{max} equals t_0 KaiB × 0.6. $KaiB_{complex}$ at $t_{1/2}$ equals t_0 KaiB × 0.6 × 0.58. More specifically, we reason that at t_{max}, [$KaiABC(t_{max})$] = [$KaiA_2(t_0)$], and [$KaiBC(t_{max})$] = [$KaiB_{complex}(t_{max})$] – [$KaiABC(t_{max})$]. At $t_{1/2}$, [$KaiABC(t_{1/2})$] = [$KaiABC(t_{max})$] × 0.58 and [$KaiBC(t_{1/2})$] = [$KaiBC(t_{max})$] × 0.58. Likewise, [$KaiA_{2\ free}(t_{1/2})$] = [$KaiABC(t_{max})$] – [$KaiABC(t_{1/2})$]. In (B) – (E), we postulate a KaiB monomer sequesters a dimer of KaiA. In (F) and (G) we postulate that a KaiB dimer sequesters a dimer of KaiA.

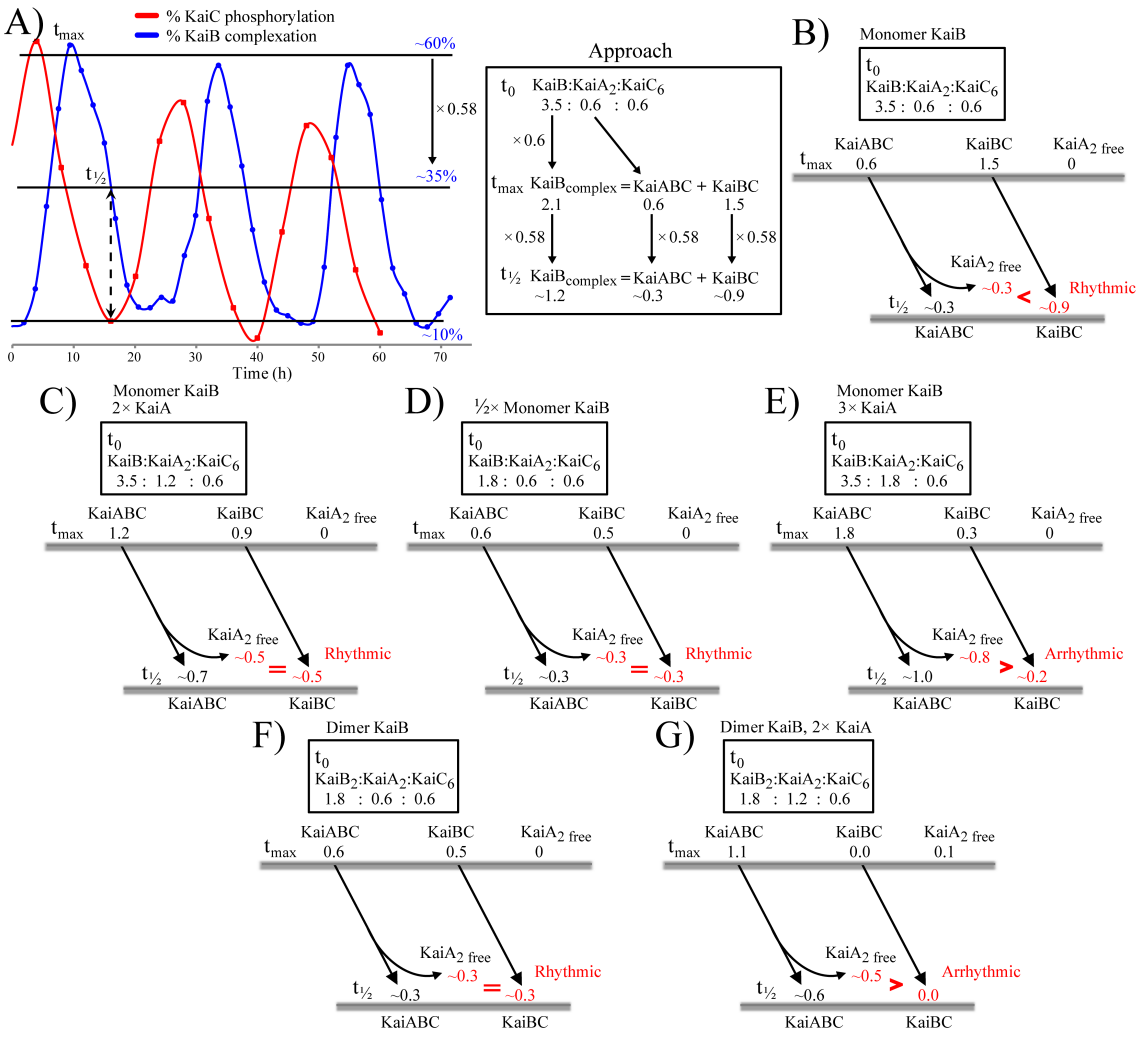


Figure 2.16. Back-of-envelope calculations.

(Figure caption is on the previous page)

CHAPTER III

CRYSTAL AND NMR STRUCTURES OF THE NIGHT TIME COMPLEXES[‡]

Synopsis

Circadian clocks, intracellular timekeeping systems, produce daily rhythms of activity and rest in organisms across all five kingdoms of life. Timekeeping mechanisms of these clocks remain elusive, as are the mechanisms by which clock signals are transmitted. In the cyanobacterial clock, the protein KaiA stimulates KaiC autophosphorylation during the day, and the protein KaiB induces KaiC (auto)dephosphorylation during the night. These proteins interact with clock-output proteins SasA and CikA to produce circadian rhythms in cell behavior. Using X-ray crystallography and NMR spectroscopy, we solved the structures of KaiB-KaiC, KaiA-KaiB-KaiC, and CikA-KaiB complexes, which are important nighttime complexes. We find that a metamorphic form of KaiB induces large-scale conformational changes in KaiA that are key to KaiC (auto)dephosphorylation, and that this same form of KaiB activates the clock signaling protein CikA. Our results feature the structural biology of a clock that utilizes the metamorphic properties of a protein to initiate a day-to-night phase transition in both the oscillator and clock-output.

[‡] X-ray crystallography experiments were performed by Nicolette Goularte and Jansen Lu in the laboratory of Dr. Carrie Partch (University of California, Santa Cruz, Department of Chemistry & Biochemistry). HDX-MS experiments were performed by David Lee in the laboratory of Dr. Sheng Li (University of California, San Diego, Department of Medicine). In vivo bioluminescence rhythms were performed by Dr. Susan E. Cohen in the laboratory of Dr. Susan S. Golden (University of California, San Diego, Division of Biological Sciences).

Introduction

Arising as an adaptation to Earth's persistent cycles of day and night, evolutionarily diverse organisms exhibit circadian (~24-h) rhythms in gene expression. These rhythms are driven by intracellular timekeeping systems known as circadian clocks (1). The oscillator of the cyanobacterial clock, composed of three proteins, KaiA, KaiB, and KaiC (30), generates a ~24-h rhythm of KaiC phosphorylation. Temporal interactions between the oscillator and two mutually antagonistic output proteins SasA, and CikA, regulate the phosphorylation level of the master transcription factor RpaA, which in turn drives global rhythms in gene expression (31, 40, 44).

KaiB binding to KaiC is key to the day/night transition in the cyanobacterial clock. The signal for this binding event is phosphorylation at Ser431 of KaiC. KaiC is a hexameric ATPase with two gene-duplicated domains termed CI (N-terminal) and CII (C-terminal) (47), with the latter harboring Ser431. Its CII domain undergoes ordered phosphorylation under the regulation of KaiA and KaiB (49, 50). During the subjective day, KaiA, a homodimer, promotes KaiC (auto)phosphorylation by binding to the so-called A loop of KaiC on the CII domain (50, 109). To transition from day to night, KaiB binds the B loop on the CI domain of KaiC (60). The binary KaiB-KaiC complex further interacts with KaiA to form a ternary KaiA-KaiB-KaiC complex that inhibits KaiA throughout the night, thereby inducing KaiC (auto)dephosphorylation (41, 42).

KaiB has recently been found to couple the oscillator to SasA and CikA-mediated output by switching between two distinctly different folds (59). KaiB fold switching is essential to pivoting the mode of the oscillator from daytime to nighttime and the mode of clock output as well. The fold-switched state of KaiB, fsKaiB, competes with SasA to disassociate the daytime KaiC-SasA complex, as KaiB and SasA bind mutually exclusively to the B loop on the CI domain of KaiC (60). Formation of the ternary KaiA-KaiB-KaiC complex is cooperative and exhibits a 10-fold increase in KaiBC binding as compared to in the binary complex (60). The cooperativity of the ternary complex was found to further promote SasA disassociation (60). Additionally, upon binding to KaiC, KaiB also interacted with CikA to stimulate its phosphatase activity toward RpaA, thereby transducing nighttime gene expression (40, 59).

The individual oscillator structures of KaiA, KaiB, KaiC are published (47, 107, 108). There are also high-resolution structures of KaiA in complex with A-loop peptides, underpinning KaiA's role in stimulating KaiC (auto)phosphorylation during the day (51, 109). However, structural details of KaiB-KaiC, KaiA-KaiB-

KaiC, and CikA-KaiB-KaiC interactions have been lacking. Thus, we pursued structures of KaiB bound to KaiC, KaiA, and CikA by X-ray crystallography and NMR.

Although the structures of full-length SasA and CikA are not available, the structures of their respective domains that interact with the oscillator, namely the N-terminal domain of SasA (^{N}SasA) and the pseudo receiver domain of CikA (CikA^{PSR}), are published (68, 72). Studying interactions between ^{N}SasA and KaiC provided insights for daytime output signal transduction (59). Here, using similar domain strategies, we solved high-resolution complex structures between domains of fold-switched KaiB (fsKaiB) and CI, $^{\Delta N}$KaiA^{C272S} and CikA^{PSR}. Specifically, we have solved the crystal structures of fsKaiB-CI and $^{\Delta N}$KaiA^{C272S}-CI-fsKaiB, at 1.8 Å and 2.7 Å resolution, respectively. Furthermore, we solved the NMR structure of the fsKaiB-CikA^{PSR} complex. These structures reveal that the oscillator proteins undergo dramatic conformational changes, which explain the clock's transition mechanism from day to night, and how that transition regulates clock output signaling.

Materials and methods

Cloning and Constructs of kaiA, *kaiB*, and *kaiC*: All genes were cloned into pET-28b using the Nde I/Hind III sites by PCR. Details of the cloning protocol have been described previously (52). Please see Table 3.1 for the list of protein constructs.

Protein expression and purification: Proteins were expressed in BL21(DE3) Escherichia coli (Novagen) and purified by Ni-NTA affinity chromatography and size-exclusion chromatography as described previously (48, 52, 60), except for A-loop peptides of KaiC was purified with C4 reversed-phase chromatography column (Vydac, Hesperia, CA) with sample loading buffer of H_2O + 0.1% trifluoroacetic acid (TFA) and elution buffer of 100% acetonitrile + 0.1% TFA. Gradient elution was set to reach 10% elution buffer at 1mL/min. Other proteins were purified by Ni-NTA affinity chromatography and gel filtration chromatography as previously described (52, 60).

Analytical gel-filtration chromatography: All gel-filtration chromatography experiments were performed with a Superdex 200 or Superdex S75 10/300 GL column (GE Healthcare) as described previously (48), except that the sample injection volume was 250 µL. Briefly, desired protein sample(s) and buffer were mixed to total of 250 µL, with or without incubation time in room temperature depending on experimental needs. The reaction sample(s) then is applied with a flow rate of 0.5 mL/min, at room temperature. The volume of the sample loop was 100 µL. Details of experimental setup and proteins used in each experiment are provided in Table 3.1.

Molecular weight markers used to calibrate the size-exclusion columns: Ferritin (440 kDa), albumin (67 kDa), ovalbumin (43 kDa), and ribonuclease A (13.7 kDa) from the gel-filtration HMW and LMW calibration kits (GE Healthcare) were used as molecular weight markers.

Fluorescence spectroscopy: All data were measured with an ISS PC1 spectrofluorometer. The thiol-reactive fluorophores used for all experiments were 6-iodoacetamidofluorescein (6-IAF) (Invitrogen). Detailed setup was followed as previously described (48), except a three-cuvette sample holder was used to measure triplicates simultaneously. The anisotropy measurements started immediately following the 5 minutes of incubation. One measurement was taken every minute for 5 minutes and the total averages were calculated. For experimental condition, please see Table 3.1.

Crystallization and X-ray crystallography structure determination of binary and ternary complexes: Initial screens were conducted using the Art Robbin

Instrument Crystal Gryphon for the KaiB-KaiC binary and KaiA-KaiB-KaiC ternary at ~20mg/ml and ~25mg/ml respectively with a three drop ratio method using the JCSG suite and protein complex screen (Qiagen) as well as the MCSG suite (Microlytic). Further crystal optimization using the hanging drop vapor diffusion at 22ºC yielded the best results using 0.2M Potassium Acetate, 10% (w/v) PEG3350, 2.5% glycerol for the binary and 0.15M Ammonium Sulfate, 0.1M Tris pH 7, 10% PEG 4000 for the ternary. Both the Binary and Ternary crystals were frozen using 12.5% glycerol cryo solution and liquid nitrogen.

Diffraction data was collected at beam line 8.3.1 at the Advance light source at Lawrence Berkley National Laboratory using the ADSC Quantum 315r detector. Both the binary and Ternary had to be deiced in order to get proper diffraction, but ice rings were still detected when collecting diffraction data for the Binary. The Binary diffraction data was processed using MOSFLM (117) within the CCP4 Suite (118), and scaled using SCALA (119, 120). The Ternary diffraction data was processed using HKL2000 (121) and converted using SCALEPACK2mtz within the CCP4 Suite. Phaser Molecular Replacement (122) in the CCP4 suite was used on both the binary and ternary complex to give initial PDB files. Refinement was conducted using PHENIX (123) refine with model building in COOT (124). Structural statistics is provided in Table 3.2.

In vivo bioluminescence rhythm: All *S. elongatus* strains were generated as previously described (125, 126). In order to test mutant variants of KaiA, KaiB and CikA strains, knockout *kaiA*, *kaiB* and *cikA* strains was created first, then secondly complemented with either original *kaiA*, *kaiB*, *cikA*, or with the mutant variants under investigation. Complemented *kaiA, kaiB* and their variants were expressed under native promoter. Complemented *cikA* and its variant were expression under *Ptrc* promoter.

Bioluminescence of *PkaiBC-luc* firefly luciferase reporter strains was used for *kaiA*, *kaiB* and their variants. *PpsbAI-luxAB* luciferase reporter was used for *cikA* and its variant. Bioluminescence was monitored at 30 °C under constant light conditions after two cycles of 12 h light: 12 h dark to synchronize the population as described previously (127). Data were analyzed with the Biological Rhythms Analysis Software System (http://millar.bio.ed.ac.uk/PEBrown/BRASS/BrassPage.htm) import and analysis program using Microsoft Excel.

Hydrogen deuterium exchange mass spectroscopy: The protocols for proteolysis, hydrogen/deuterium exchange and LC/MS analysis were similar to previously described (59), except that ratio of 2:1:1 mixture was used for $^{\Delta N}$KaiA:fsKaiB:CI* ternary sample to probe the relative hydrogen/deuterium exchange of $^{\Delta N}$KaiA free and mixed with fsKaiB-CI*, or fsKaiB-CI* in the absence

or the presence of $^{\Delta N}$KaiA. Standard 1:1 ratio was use for CikA^{PSR}-fsKaiB^{N29A} complex. The time points to quench the exchange were 10, 100 and 1000 seconds. All dataset were analyzed and the 100s dataset was selected to represent in the complex structures.

NMR spectroscopy, residual dipolar coupling and structure determination: All NMR experiments were run on a Bruker 600MHz AVANCE III spectrometer equipped with a TCI cryoprobe. Chemical shifts were referenced to internal DSS. Data were processed using NMRPipe and visualized using NMRDraw (100). Spectra were analyzed using PIPP (128) and XIPP (G. S. Garrett and G. M. Clore, unpublished data). For details on the protein constructs and experimental setup, please see Table 3.1.

For solving the single entity protein structures, all experiments were performed at 25°C. For solving binary complex, all experiments were performed at 50°C. 3D NMR experiments used to assign backbone, sidechain and intramolecular proton nuclear Overhauser effect (NOE) and residual dipolar couplings (RDC) were: HNCACB, HN(CO)CACB, HN(CA)CO, HNCO, HBHA(CO)NH, HCC(CO)NH, ^{15}N-NOESY-HSQC, ^{13}C-NOESY-HSQC, HCCH-COSY, HCAN, HCA(CO)N, IPAP-J-HNCO(CA)-measures ^{1}J(CA-HA), IPAP-J-HNCO-measures ^{1}J(N-H) and IPAP-J-HNCO-measures ^{1}J(CA-CO). For binary complex, additional HCCH-TOCSY, ^{13}C, ^{13}C edited 3D NOESY (in a sample containing ^{13}C-enriched CikA^{PSR} with unenriched fsKaiB^{N29A} or ^{13}C-enriched fsKaiB^{N29A} with unenriched CikA^{PSR}), ^{15}N-edited 3D NOESY (in a sample containing ^{15}N-enriched CikA^{PSR} with unenriched fsKaiB^{N29A} or ^{15}N-enriched fsKaiB^{N29A} with unenriched CikA^{PSR}), and 2D ^{15}N HSQC-IPAP (in a sample containing ^{15}N-enriched CikA^{PSR} with unenriched fsKaiB^{N29A} or ^{15}N-enriched fsKaiB^{N29A} with unenriched CikA^{PSR}), were collected for sidechain, intramolecular NOE and RDC. All NMR samples above were mixed with 1x complete midi protease inhibitor (Roche).

Interproton nuclear Overhauser effect (NOE) distance restraints for the binary complex were collected by: ^{15}N, ^{13}C edited 4D NOESY (in a sample containing ^{15}N-enriched CikA^{PSR} with ^{13}C-enriched fsKaiB^{N29A15} or ^{15}N-enriched fsKaiB^{N29A} with ^{13}C-enriched CikA^{PSR}), and ^{13}C-edited, ^{12}C-filtered 3D NOESY (in a sample containing ^{13}C-enriched CikA^{PSR} with unenriched fsKaiB^{N29A} or ^{13}C-enriched fsKaiB^{N29A} with unenriched CikA^{PSR}).

Residual dipolar coupling samples were soaked in 5.16% stretched polyacrylamide gel made with 6 mm chamber diameter (new era) for single entity protein samples and 5.5% stretched polyacrylamide gel made with 5 mm chamber diameter (new era) for binary complex protein samples. The buffer salt

amounts were reduced by half. The protocol for making an alignment gel has been described previously (129).

For NMR structure determination, dihedral angle values were predicted from chemical shifts using TALOS-N webserver (http://spin.niddk.nih.gov/bax/software/TALOS-N/), an upgraded version of TALOS+ (130). Structure calculation were made using XPLOR-NIH (131) following newer protocols from Tian et al. (132) and RDC refinement protocols from Chou et al. (133).

Structure validation statistics were generated by three webservers: PSVS (134), Molprobity (135), and DC (Restrained Fit of Dipolar Couplings to a PDB Structure Using Non-Linear Least-Squares) (http://spin.niddk.nih.gov/bax/nmrserver/dc/fit.html).

Table 3.1. Abbreviation and full name of *Thermosynechococcus elongatus* Kai proteins used in each experimental condition in Chapter Three.

Experiment Type	[Abbreviation]-Protein Full Name (N-terminal Tag_Protein_Mutation_Length of Protein_ C-terminal Tag) / Final Concentration (μM if not indicated)	Experimental Condition
X-ray Crystallography For fsKaiB-CI* Complex (Figure 3.2)	1) [CI*]-FLAG_KaiC_R41A-K173A_17-247_FLAG / ~20 mg/mL 2) [fsKaiB]-KaiB_Y8A-Y94A-G89A-D91R_1-99 / ~20 mg/mL	• Volume: three drop ratio method (1 drop is ~200nL) • Temperature: 22 °C • Buffer: JCSG suite, Qiagen (protein complex screen), and MCSG suite (Microlytic)
X-ray Crystallography For $^{\Delta N}$KaiA^{C272S}-fsKaiB-CI* Complex (Figure 3.3)	1) [CI*]-FLAG_KaiC_R41A-K173A_17-247_FLAG / ~25 mg/mL 2) [fsKaiB]-KaiB_Y8A-Y94A-G89A-D91R_1-99 / ~25 mg/mL 3) [$^{\Delta N}$KaiA^{C272S}]-FLAG_KaiA_C272S_147-283 / ~25 mg/mL	• Volume: three drop ratio method (1 drop is ~200nL) • Temperature: 22 °C • Buffer: JCSG suite, Qiagen (protein complex screen), and MCSG suite (Microlytic)
Fluorescence Spectroscopy For A loop Competition (Figure 3.5D)	1) [A loop]-Cystine_KaiC_488-518 (6-IAF-labeled) / 0.05 2) [KaiA]-KaiA_1-283 / 10 3) [CI*]-FLAG_KaiC_R41A-K173A_17-247_FLAG / 5, 10 , 15, 20, 30 4) [fsKaiB]-KaiB_Y8A-Y94A-G89A-D91R_1-99 / 5, 10 , 15, 20, 30	• Volume: 350 μL • Pre-experiment incubation time of mixture: 5 mins • Temperature: 25 °C • Buffer: 20 mM Tris, 50 mM NaCl, 0.2% NaN_3, pH 7.0
HDX-MS For $^{\Delta N}$KaiA^{C272S}-fsKaiB-CI* Complex (Figure 3.6A)	1) [$^{\Delta N}$KaiA]-FLAG_KaiA_147-283 2) [fsKaiB]-KaiB_FLAG_Y8A-Y94A-G89A-D91R_1-98 3) [CI*]-FLAG_KaiC_R41A-K173A_17-247_FLAG	• Buffer: 10mM Tris, 100mM NaCl, pH 7.2
HDX-MS For CikA^{PSR}-fsKaiB^{N29A} Complex (Figure 3.6B)	1) [CikA^{PSR}]-CikA_613-729 2) [fsKaiB^{N29A}]-KaiB_FLAG_Y8A-Y94A-G89A-D91R_N29A_1-98	• Buffer: 10mM Tris, 100mM NaCl, pH 7.2

S200 Analytical Gel Filtration Chromatography For $^{\Delta N}$KaiA-fsKaiB-CI*^{L} Complex (Figure 3.7A)	1) [fsKaiB]-KaiB_Y8A-Y94A-G89A-D91R_1-99 / 30 2) [CI*^{L}]-FLAG_KaiC_R41A-K173A_1-247_FLAG / 15 3) [$^{\Delta N}$KaiA]-FLAG_KaiA_147-283 / 30 4) [H153A-$^{\Delta N}$KaiA]-FLAG_KaiA_H153A_147-283 / 30 5) [L155A-$^{\Delta N}$KaiA]-FLAG_KaiA_L155A_147-283 / 30 6) [Q157A-$^{\Delta N}$KaiA]-FLAG_KaiA_Q157A_147-283 / 30 7) [K158A-$^{\Delta N}$KaiA]-FLAG_KaiA_K158A_147-283 / 30 8) [L159A-$^{\Delta N}$KaiA]-FLAG_KaiA_L159A_147-283 / 30 9) [E161A-$^{\Delta N}$KaiA]-FLAG_KaiA_E161A_147-283 / 30 10) [L163A-$^{\Delta N}$KaiA]-FLAG_KaiA_L163A_147-283 / 30	• Volume: 250 µL • Temperature: room temperature (~22 °C) • Buffer: 20 mM Tris, 100 mM NaCl, pH 7.0
S200 Analytical Gel Filtration Chromatography For $^{\Delta N}$KaiA- fsKaiB^{L}-CI*^{L} Complex (Figure 3.7B)	1) [$^{\Delta N}$KaiA]-FLAG_KaiA_147-283 / 30 2) [CI*^{L}]-FLAG_KaiC_R41A-K173A_1-247_FLAG / 16 3) [fsKaiB^{L}]-FLAG_KaiB_FLAG_G89A-D91R_1-108 / 30 4) [A41D-fsKaiB^{L}]-FLAG_KaiB_FLAG_G89A-D91R_A41D_1-108 / 30 5) [K43E-fsKaiB^{L}]-FLAG_KaiB_FLAG_G89A-D91R_K43E_1-108 / 30 6) [K43Q-fsKaiB^{L}]-FLAG_KaiB_FLAG_G89A-D91R_K43Q_1-108 / 30	• Volume: 250 µL • Temperature: room temperature (~22 °C) • Buffer: 20 mM Tris, 100 mM NaCl, pH 7.0
S75 Analytical Gel Filtration Chromatography For CikA^{PSR}-fsKaiB^{N29A} Complex (Figure 3.12A)	1) [CikA^{PSR}]-CikA_613-729 / 30 2) [fsKaiB^{N29A}]-KaiB_FLAG_Y8A-Y94A-G89A-D91R_N29A_1-98 / 30 3) [E33A-fsKaiB^{N29A}]-KaiB_FLAG_Y8A-Y94A-G89A-D91R_N29A_E33A_1-98 / 30 4) [A41D-fsKaiB^{N29A}]-KaiB_FLAG_Y8A-Y94A-G89A-D91R_N29A_A41D_1-98 / 30 5) [K43E-fsKaiB^{N29A}]-KaiB_FLAG_Y8A-Y94A-G89A-D91R_N29A_K43E_1-98 / 30	• Volume: 250 µL • Temperature: room temperature (~22 °C) • Buffer: 20 mM Tris, 100 mM NaCl, pH 7.0
S75 Analytical Gel Filtration Chromatography For CikA^{PSR}-fsKaiB^{N29A} Complex (Figure 3.12B)	1) [fsKaiB^{N29A}]-KaiB_FLAG_Y8A-Y94A-G89A-D91R_N29A_1-98 / 30 2) [CikA^{PSR}]-CikA_613-729 / 30 3) [T627R-CikA^{PSR}]-CikA_T627R_613-729 / 30 4) [C630R-CikA^{PSR}]-CikA_C630R_613-729 / 30 5) [R650D-CikA^{PSR}]-CikA_R650D_613-729 / 30	• Volume: 250 µL • Temp: room temp (~22 °C) • Buffer: 20 mM Tris, 100 mM NaCl, pH 7.0

Methyl-TROSY NMR For $^{\Delta N}$KaiA (Figure 3.4)	1) [$^{\Delta N}$KaiA]-FLAG_KaiA_147-283 (U-[^{15}N, ^{2}H]-Ile-1-[^{13}C, ^{1}H] labeled) / 20 2) [wt-KaiB]-KaiB_1-108 / 50 3) [CI^{*L}]-FLAG_KaiC_R41A-K173A_1-247_FLAG / 50 4) [fsKaiB]-KaiB_Y8A-Y94A-G89A-D91R_1-99 / 50	• Volume: 330 µL • Pre-experiment incubation time for mixtures: 24 hrs. • ^{1}H / ^{13}C sweep width (Hz): 3894.081 / 1209.247 • ^{1}H / ^{13}C carrier (Hz): -379.00 / 1626.00 • ^{1}H / ^{13}C acquisition time (ms): 63.9 / 79.3 • Temp: 25 °C • Buffer in 99.96 % D_2O: 20 mM Tris, 100 mM NaCl, 5 mM DTT, 1 mM ATP, 1 mM $MgCl_2$, 10 µM DSS, 0.02% NaN_3, pH 7 • Shaped tube
Methyl-TROSY NMR For CikA^{PSR} (Figure 3.10)	1) [CikA^{PSR}]-CikA_613-729 (U-[^{15}N, ^{2}H]-Ile-1-[^{13}C, ^{1}H] labeled) / 20 2) [CI^{*L}]-FLAG_KaiC_R41A-K173A_1-247_FLAG / 50 3) [wt-KaiB]-KaiB_1-108 / 50 4) [fsKaiB^{L}]-FLAG_KaiB_FLAG_G89A-D91R_1-108 / 50 5) [N29A-fsKaiB^{L}]-FLAG_KaiB_FLAG_G89A-D91R_N29A_1-108 / 50 6) [fsKaiB^{N29A}]-KaiB_FLAG_Y8A-Y94A-G89A-D91R_N29A_1-98 / 50 7) [CI^{*L}]-FLAG_KaiC_R41A-K173A_1-247_FLAG / 50 8) [fsKaiB]-KaiB_Y8A-Y94A-G89A-D91R_1-99 / 50 9) [$^{\Delta N}$KaiA]-FLAG_KaiA_147-283 / 110 10) [KaiA]-KaiA_1-283 / 110	• Volume: 330 µL • Pre-experiment incubation time for mixtures: 24 hrs. • ^{1}H / ^{13}C sweep width (Hz): 3894.081 / 2418.494 • ^{1}H / ^{13}C carrier (Hz): -379.00 / 1626.00 • ^{1}H / ^{13}C acquisition time (ms): 63.9 / 41.3 • Temp: 25 °C • Buffer in 99.96 % D_2O: 20 mM Tris, 100 mM NaCl, 5 mM DTT, 1 mM ATP, 1 mM $MgCl_2$, 10 µM DSS, 0.02% NaN_3, pH 7 • Shaped tube
^{15}N-HSQC NMR For $^{\Delta N}$KaiA and L155A-$^{\Delta N}$KaiA (Figure 3.8A)	1) [$^{\Delta N}$KaiA]- FLAG_KaiA_147-283 (U-[^{15}N, ^{1}H] labeled) / 40 2) [L155A-$^{\Delta N}$KaiA]-FLAG_KaiA_L155A_147-283 / 40	• Volume: 350 µL • ^{1}H / ^{15}N sweep width (Hz): 9615.385 / 1583.901 • ^{1}H / ^{15}N carrier (Hz): 2824.00 / 7278.86 • ^{1}H / ^{15}N acquisition time (ms): 68.8 / 63.1 • Temp: 25 °C • Buffer in 5% D_2O: 20 mM Tris, 100 mM NaCl, 00.02% NaN_3, 10 µM DSS, pH 7 • Shaped tube

^{15}N-HSQC NMR For CikA^{PSR} and C630R-CikA^{PSR} (Figure 3.8B)	1) [CikA^{PSR}]-CikA_613-729 (U-[^{15}N] labeled) / 100 2) [C630R-CikA^{PSR}]-CikA_C630R_613-729 (U-[^{15}N] labeled) / 100	• Volume: 350 µL • ^{1}H / ^{15}N sweep width (Hz): 9615.385 / 1583.901 • ^{1}H / ^{15}N carrier (Hz): 2824.00 / 7278.86 • ^{1}H / ^{15}N acquisition time (ms): 68.8 / 63.1 • Temp: 25 °C • Buffer in 5% D_2O: 20 mM Tris, 100 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7 • Shaped tube
3D HNCACB and HN(CO)CACB NMR for free fsKaiB^{N29A}	1) [fsKaiB^{N29A}]-KaiB_FLAG_Y8A-Y94A-G89A-D91R_N29A_1-98 (U-[^{15}N, ^{13}C] labeled) / 800	• Volume: 350 µL • ^{1}H / ^{15}N / ^{13}C sweep width (Hz): 7211.539 / 1583.899 / 9500.000 • ^{1}H / ^{15}N / ^{13}C carrier (Hz): 2828.00 / 7278.86 / 6575.41 • ^{1}H / ^{15}N / ^{13}C acquisition time (ms): 67.9 / 23.9 / 6.9 • Temp: 25 °C • Buffer in 5% D_2O: 20 mM Tris, 100 mM NaCl, 5 mM TCEP, 00.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Shaped tube
3D HN(CA)CO and HNCO NMR for free fsKaiB^{N29A}	1) [fsKaiB^{N29A}]-KaiB_FLAG_Y8A-Y94A-G89A-D91R_N29A_1-98 (U-[^{15}N, ^{13}C] labeled) / 800	• Volume: 350 µL • ^{1}H / ^{15}N / ^{13}C sweep width (Hz): 7211.539 / 1583.899 / 1500.000 • ^{1}H / ^{15}N / ^{13}C carrier (Hz): 2828.00 / 7278.86 / 26225.41 • ^{1}H / ^{15}N / ^{13}C acquisition time (ms): 67.9 / 23.3 / 30.0 • Temp: 25 °C • Buffer in 5% D_2O: 20 mM Tris, 100 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Shaped tube

3D HBHA(CO)NH NMR for free fsKaiB^{N29A}	1) [fsKaiB^{N29A}]-KaiB_FLAG_Y8A-Y94A-G89A-D91R_N29A_1-98 (U-[^{15}N, ^{13}C] labeled) / 800	• Volume: 350 µL • ^{1}H / ^{15}N / ^{1}H sweep width (Hz): 7211.539 / 1583.899 / 4000.000 • ^{1}H / ^{15}N / ^{1}H carrier (Hz): 2828.00 / 7278.86 / 2828.00 • ^{1}H / ^{15}N / ^{1}H acquisition time (ms): 67.9 / 23.9 / 13.0 • Temp: 25 °C • Buffer in 5% D_2O: 20 mM Tris, 100 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Shaped tube
3D HCC(CO)NH NMR for free fsKaiB^{N29A} Proton version	1) [fsKaiB^{N29A}]-KaiB_FLAG_Y8A-Y94A-G89A-D91R_N29A_1-98 (U-[^{15}N, ^{13}C] labeled) / 800	• Volume: 350 µL • ^{1}H / ^{15}N / ^{1}H sweep width (Hz): 7211.539 / 1583.899 / 4000.000 • ^{1}H / ^{15}N / ^{1}H carrier (Hz): 2828.00 / 7278.86 / 2828.00 • ^{1}H / ^{15}N / ^{1}H acquisition time (ms): 67.9 / 23.9 / 13.0 • Mixing time (ms): 12 • Temp: 25 °C • Buffer in 5% D_2O: 20 mM Tris, 100 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Shaped tube
3D HCC(CO)NH NMR for free fsKaiB^{N29A} Carbon version	1) [fsKaiB^{N29A}]-KaiB_FLAG_Y8A-Y94A-G89A-D91R_N29A_1-98 (U-[^{15}N, ^{13}C] labeled) / 800	• Volume: 350 µL • ^{1}H / ^{15}N / ^{13}C sweep width (Hz): 7211.539 / 1583.899 / 9674.291 • ^{1}H / ^{15}N / ^{13}C carrier (Hz): 2828.00 / 7278.86 / 6537.93 • ^{1}H / ^{15}N / ^{13}C acquisition time (ms): 67.9 / 23.9 / 8.0 • Mixing time (ms): 12 • Temp: 25 °C • Buffer in 5% D_2O: 20 mM Tris, 100 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Shaped tube

3D ^{15}N-NOESY-HSQC NMR for free fsKaiB^{N29A}	1) [fsKaiB^{N29A}]-KaiB_FLAG_Y8A-Y94A-G89A-D91R_N29A_1-98 (U-[^{15}N, ^{13}C] labeled) / 800	• Volume: 350 µL • ^{1}H / ^{15}N / ^{1}H sweep width (Hz): 7211.539 / 1583.899 / 6011.328 • ^{1}H / ^{15}N / ^{1}H carrier (Hz): 2828.00 / 7278.86 / 2828.00 • ^{1}H / ^{15}N / ^{1}H acquisition time (ms): 67.9 / 21.4 / 24.9 • Mixing time (ms): 18 • Temp: 25 °C • Buffer in 5% D_2O: 20 mM Tris, 100 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Shaped tube
3D ^{13}C-NOESY-HSQC NMR for free fsKaiB^{N29A}	1) [fsKaiB^{N29A}]-KaiB_FLAG_Y8A-Y94A-G89A-D91R_N29A_1-98 (U-[^{15}N, ^{13}C] labeled) / 800	• Volume: 350 µL • ^{1}H / ^{13}C / ^{1}H sweep width (Hz): 6009.615 / 5215.029 / 4027.582 • ^{1}H / ^{13}C / ^{1}H carrier (Hz): 1643.56 / 6011.05 / 1643.56 • ^{1}H / ^{13}C / ^{1}H acquisition time (ms): 67.8 / 9.9 / 32.2 • Mixing time (ms): 15 • Temp: 25 °C • Buffer in 99.96% D_2O: 20 mM Tris, 100 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Shaped tube
3D ^{13}C-HCCH-COSY NMR for free fsKaiB^{N29A} Proton version	1) [fsKaiB^{N29A}]-KaiB_FLAG_Y8A-Y94A-G89A-D91R_N29A_1-98 (U-[^{15}N, ^{13}C] labeled) / 800	• Volume: 350 µL • ^{1}H / ^{13}C / ^{1}H sweep width (Hz): 6009.615 / 5215.027 / 4027.584 • ^{1}H / ^{13}C / ^{1}H carrier (Hz): 1643.56 / 6011.05 / 1643.56 • ^{1}H / ^{13}C / ^{1}H acquisition time (ms): 67.8 / 9.9 / 32.2 • Temp: 25 °C • Buffer in 99.96% D_2O: 20 mM Tris, 100 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Shaped tube

3D ^{13}C-HCCH-COSY NMR for free fsKaiB^{N29A} Carbon version	1) [fsKaiB^{N29A}]-KaiB_FLAG_Y8A-Y94A-G89A-D91R_N29A_1-98 (U-[^{15}N, ^{13}C] labeled) / 800	• Volume: 350 µL • ^{1}H / ^{13}C / ^{13}C sweep width (Hz): 6009.615 / 5215.027 / 5215.027 • ^{1}H / ^{13}C / ^{13}C carrier (Hz): 1643.56 / 6011.05 / 6011.05 • ^{1}H / ^{13}C / ^{13}C acquisition time (ms): 67.8 / 9.9 / 9.9 • Temp: 25 °C • Buffer in 99.96% D_2O: 20 mM Tris, 100 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Shaped tube
3D HCAN and HCA(CO)N NMR for free fsKaiB^{N29A}	1) [fsKaiB^{N29A}]-KaiB_FLAG_Y8A-Y94A-G89A-D91R_N29A_1-98 (U-[^{15}N, ^{13}C] labeled) / 800	• Volume: 350 µL • ^{1}H / ^{13}C / ^{15}N sweep width (Hz): 6009.615 / 4534.871 / 1583.901 • ^{1}H / ^{13}C / ^{15}N carrier (Hz): 2850.40 / 8086.90 / 7278.98 • ^{1}H / ^{13}C / ^{15}N acquisition time (ms): 63.8 / 9.9 / 23.9 • Temp: 25 °C • Buffer in 99.96% D_2O: 20 mM Tris, 100 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Shaped tube
3D IPAP-J-HNCO(CA) NMR for 5.16% stretched polyacrylamide gel RDC of free fsKaiB^{N29A} measures: ^{1}J(CA-HA)	1) [fsKaiB^{N29A}]-KaiB_FLAG_Y8A-Y94A-G89A-D91R_N29A_1-98 (U-[^{15}N, ^{13}C] labeled) / 300	• ^{1}H / ^{15}N / ^{13}C sweep width (Hz): 7211.539 / 1583.899 / 1500.000 • ^{1}H / ^{15}N / ^{13}C carrier (Hz): 2823.20 / 7278.86 / 26187.80 • ^{1}H / ^{15}N / ^{13}C acquisition time (ms): 67.9 / 23.3 / 50.6 • Temp: 25 °C • Buffer in 10% D_2O: 10 mM Tris, 50 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Aligned sample: 4.2mm new era NMR tubes (NE-UP5-GT-7) • Isotropic sample: shaped tube

3D IPAP-J-HNCO NMR for 5.16% stretched polyacrylamide gel RDC of free fsKaiB^{N29A} measures: ^{1}J(N-H)	1) [fsKaiB^{N29A}]-KaiB_FLAG_Y8A-Y94A-G89A-D91R_N29A_1-98 (U-[^{15}N, ^{13}C] labeled) / 300	• ^{1}H / ^{15}N / ^{13}C sweep width (Hz): 7211.539 / 1583.899 / 1500.004 • ^{1}H / ^{15}N / ^{13}C carrier (Hz): 2823.20 / 7278.86 / 26187.80 • ^{1}H / ^{15}N / ^{13}C acquisition time (ms): 67.9 / 46.7 / 19.9 • Temp: 25 °C • Buffer in 10% D_2O: 10 mM Tris, 50 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Aligned sample: 4.2mm new era NMR tubes (NE-UP5-GT-7) • Isotropic sample: shaped tube
3D IPAP-J-HNCO NMR for 5.16% stretched polyacrylamide gel RDC of free fsKaiB^{N29A} measures: ^{1}J(CA-CO)	1) [fsKaiB^{N29A}]-KaiB_FLAG_Y8A-Y94A-G89A-D91R_N29A_1-98 (U-[^{15}N, ^{13}C] labeled) / 300	• ^{1}H / ^{15}N / ^{13}C sweep width (Hz): 7211.539 / 1583.899 / 1500.000 • ^{1}H / ^{15}N / ^{13}C carrier (Hz): 2823.20 / 7278.86 / 26187.80 • ^{1}H / ^{15}N / ^{13}C acquisition time (ms): 67.9 / 23.3 / 50.6 • Temp: 25 °C • Buffer in 10% D_2O: 10 mM Tris, 50 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Aligned sample: 4.2mm new era NMR tubes (NE-UP5-GT-7) • Isotropic sample: shaped tube
3D HNCACB and HN(CO)CACB NMR for free CikA^{PSR}	1) [CikA^{PSR}]-CikA_613-729 (U-[^{15}N, ^{13}C] labeled) / 1500	• Volume: 350 µL • ^{1}H / ^{15}N / ^{13}C sweep width (Hz): 9615.385 / 1583.899 / 11337.012 • ^{1}H / ^{15}N / ^{13}C carrier (Hz): 2823.00 / 7278.86 / 6537.93 • ^{1}H / ^{15}N / ^{13}C acquisition time (ms): 68.8 / 23.3 / 6.3 • Temp: 25 °C • Buffer in 5% D_2O: 20 mM Tris, 100 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Shaped tube

3D HN(CA)CO and HNCO NMR for free $CikA^{PSR}$	1) [$CikA^{PSR}$]-CikA_613-729 (U-[^{15}N, ^{13}C] labeled) / 1500	• Volume: 350 µL • ^{1}H / ^{15}N / ^{13}C sweep width (Hz): 9615.385 / 1583.899 / 1500.004 • ^{1}H / ^{15}N / ^{13}C carrier (Hz): 2823.00 / 7278.86 / 26225.41 • ^{1}H / ^{15}N / ^{13}C acquisition time (ms): 68.8 / 23.9 / 29.9 • Temp: 25 °C • Buffer in 5% D_2O: 20 mM Tris, 100 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Shaped tube
3D HBHA(CO)NH NMR for free $CikA^{PSR}$	1) [$CikA^{PSR}$]-CikA_613-729 (U-[^{15}N, ^{13}C] labeled) / 1500	• Volume: 350 µL • ^{1}H / ^{15}N / ^{1}H sweep width (Hz): 9615.385 / 1583.899 / 4000.000 • ^{1}H / ^{15}N / ^{1}H carrier (Hz): 2823.00 / 7278.86 / 2823.00 • ^{1}H / ^{15}N / ^{1}H acquisition time (ms): 68.8 / 23.9 / 13.0 • Temp: 25 °C • Buffer in 5% D_2O: 20 mM Tris, 100 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Shaped tube
3D HCC(CO)NH NMR for free $CikA^{PSR}$ Proton version	1) [$CikA^{PSR}$]-CikA_613-729 (U-[^{15}N, ^{13}C] labeled) / 1500	• Volume: 350 µL • ^{1}H / ^{15}N / ^{1}H sweep width (Hz): 9615.385 / 1583.899 / 4000.000 • ^{1}H / ^{15}N / ^{1}H carrier (Hz): 2823.00 / 7278.86 / 2823.00 • ^{1}H / ^{15}N / ^{1}H acquisition time (ms): 68.8 / 23.9 / 13.0 • Mixing time (ms): 12 • Temp: 25 °C • Buffer in 5% D_2O: 20 mM Tris, 100 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Shaped tube

3D HCC(CO)NH NMR for free CikA^{PSR} Carbon version	1) [CikA^{PSR}]-CikA_613-729 (U-[^{15}N, ^{13}C] labeled) / 1500	• Volume: 350 µL • ^{1}H / ^{15}N / ^{13}C sweep width (Hz): 9615.385 / 1583.899 / 11337.012 • ^{1}H / ^{15}N / ^{13}C carrier (Hz): 2828.00 / 7278.86 / 6537.93 • ^{1}H / ^{15}N / ^{13}C acquisition time (ms): 68.8 / 23.9 / 6.2 • Mixing time (ms): 12 • Temp: 25 °C • Buffer in 5% D_2O: 20 mM Tris, 100 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Shaped tube
3D ^{15}N-NOESY-HSQC NMR for free CikA^{PSR}	1) [CikA^{PSR}]-CikA_613-729 (U-[^{15}N, ^{13}C] labeled) / 1500	• Volume: 350 µL • ^{1}H / ^{15}N / ^{1}H sweep width (Hz): 9615.385 / 1583.899 / 5530.422 • ^{1}H / ^{15}N / ^{1}H carrier (Hz): 2823.00 / 7278.86 / 2823.00 • ^{1}H / ^{15}N / ^{1}H acquisition time (ms): 68.8 / 21.4 / 23.1 • Mixing time (ms): 18 • Temp: 25 °C • Buffer in 5% D_2O: 20 mM Tris, 100 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Shaped tube
3D ^{13}C-NOESY-HSQC NMR for free CikA^{PSR}	1) [CikA^{PSR}]-CikA_613-729 (U-[^{15}N, ^{13}C] labeled) / 1500	• Volume: 350 µL • ^{1}H / ^{13}C / ^{1}H sweep width (Hz): 6009.615 / 4534.808 / 4027.582 • ^{1}H / ^{13}C / ^{1}H carrier (Hz): 1643.56 / 6011.05 / 1643.56 • ^{1}H / ^{13}C / ^{1}H acquisition time (ms): 63.9 / 10.1 / 32.2 • Mixing time (ms): 15 • Temp: 25 °C • Buffer in 99.96% D_2O: 20 mM Tris, 100 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Shaped tube

3D ^{13}C-HCCH-COSY NMR for free CikA^{PSR} Proton version	1) [CikA^{PSR}]-CikA_613-729 (U-[^{15}N, ^{13}C] labeled) / 1500	• Volume: 350 µL • ^{1}H / ^{13}C / ^{1}H sweep width (Hz): 6009.615 / 4534.808 / 4027.582 • ^{1}H / ^{13}C / ^{1}H carrier (Hz): 1643.56 / 6011.05 / 1643.56 • ^{1}H / ^{13}C / ^{1}H acquisition time (ms): 63.9 / 10.1 / 32.2 • Temp: 25 °C • Buffer in 99.96% D_2O: 20 mM Tris, 100 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Shaped tube
3D ^{13}C-HCCH-COSY NMR for free CikA^{PSR} Carbon version	1) [CikA^{PSR}]-CikA_613-729 (U-[^{15}N, ^{13}C] labeled) / 1500	• Volume: 350 µL • ^{1}H / ^{13}C / ^{13}C sweep width (Hz): 6009.615 / 4534.808 / 4534.808 • ^{1}H / ^{13}C / ^{13}C carrier (Hz): 1643.56 / 6011.05 / 6011.05 • ^{1}H / ^{13}C / ^{13}C acquisition time (ms): 63.9 / 10.1 / 10.1 • Temp: 25 °C • Buffer in 99.96% D_2O: 20 mM Tris, 100 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Shaped tube
3D IPAP-J-HNCO(CA) NMR for 5.16% stretched polyacrylamide gel RDC of free CikA^{PSR} measures: ^{1}J(CA-HA)	1) [CikA^{PSR}]-CikA_613-729 (U-[^{15}N, ^{13}C] labeled) / 450	• ^{1}H / ^{15}N / ^{13}C sweep width (Hz): 9615.385 / 1583.899 / 1500.004 • ^{1}H / ^{15}N / ^{13}C carrier (Hz): 2823.20 / 7278.86 / 26187.80 • ^{1}H / ^{15}N / ^{13}C acquisition time (ms): 68.9 / 23.3 / 50.6 • Temp: 25 °C • Buffer in 10% D_2O: 10 mM Tris, 50 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Aligned sample: 4.2mm new era NMR tubes (NE-UP5-GT-7) • Isotropic sample: shaped tube

3D IPAP-J-HNCO NMR for 5.16% stretched polyacrylamide gel RDC of free $CikA^{PSR}$ measures: $^{1}J(N-H)$	1) $[CikA^{PSR}]$-CikA_613-729 (U-[^{15}N, ^{13}C] labeled) / 450	• ^{1}H / ^{15}N / ^{13}C sweep width (Hz): 9615.385 / 1583.899 / 1500.004 • ^{1}H / ^{15}N / ^{13}C carrier (Hz): 2823.20 / 7278.86 / 26187.80 • ^{1}H / ^{15}N / ^{13}C acquisition time (ms): 68.8 / 46.7 / 19.9 • Temp: 25 °C • Buffer in 10% D_2O: 10 mM Tris, 50 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Aligned sample: 4.2mm new era NMR tubes (NE-UP5-GT-7) • Isotropic sample: shaped tube
3D IPAP-J-HNCO NMR for 5.16%, stretched polyacrylamide gel RDC of free $CikA^{PSR}$ measures: $^{1}J(CA-CO)$	1) $[CikA^{PSR}]$-CikA_613-729 (U-[^{15}N, ^{13}C] labeled) / 450	• ^{1}H / ^{15}N / ^{13}C sweep width (Hz): 9615.385 / 1583.899 / 1500.004 • ^{1}H / ^{15}N / ^{13}C carrier (Hz): 2823.20 / 7278.86 / 26187.80 • ^{1}H / ^{15}N / ^{13}C acquisition time (ms): 68.8 / 23.3 / 50.6 • Temp: 25 °C • Buffer in 10% D_2O: 10 mM Tris, 50 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Aligned sample: 4.2mm new era NMR tubes (NE-UP5-GT-7) • Isotropic sample: shaped tube
3D HNCACB and HN(CO)CACB NMR for labeled $fsKaiB^{N29A}$ + 1.1x unlabeled $CikA^{PSR}$	1) $[fsKaiB^{N29A}]$-KaiB_FLAG_Y8A-Y94A-G89A-D91R_N29A_1-98 (U-[^{15}N, ^{13}C] labeled) / 800 2) $[CikA^{PSR}]$-CikA_613-729 / 880	• Volume: 380 µL • ^{1}H / ^{15}N / ^{13}C sweep width (Hz): 7211.539 / 1583.899 / 9500.000 • ^{1}H / ^{15}N / ^{13}C carrier (Hz): 2828.00 / 7278.86 / 6159.72 • ^{1}H / ^{15}N / ^{13}C acquisition time (ms): 62.8 / 23.9 / 7.0 • Temp: 50 °C • Buffer in 5% D_2O: 20 mM Tris, 100 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Shaped tube

3D HN(CA)CO and HNCO NMR for labeled fsKaiB^{N29A} + 1.1x unlabeled CikA^{PSR}	1) [fsKaiB^{N29A}]-KaiB_FLAG_Y8A-Y94A-G89A-D91R_N29A_1-98 (U-[^{15}N, ^{13}C] labeled) / 800 2) [CikA^{PSR}]-CikA_613-729 / 880	• Volume: 380 µL • ^{1}H / ^{15}N / ^{13}C sweep width (Hz): 7211.539 / 1583.899 / 1500.000 • ^{1}H / ^{15}N / ^{13}C carrier (Hz): 2828.00 / 7278.86 / 25809.72 • ^{1}H / ^{15}N / ^{13}C acquisition time (ms): 62.8 / 23.3 / 30.0 • Temp: 50 °C • Buffer in 5% D_2O: 20 mM Tris, 100 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Shaped tube
3D HBHA(CO)NH NMR for labeled fsKaiB^{N29A} + 1.1x unlabeled CikA^{PSR}	1) [fsKaiB^{N29A}]-KaiB_FLAG_Y8A-Y94A-G89A-D91R_N29A_1-98 (U-[^{15}N, ^{13}C] labeled) / 800 2) [CikA^{PSR}]-CikA_613-729 / 880	• Volume: 380 µL • ^{1}H / ^{15}N / ^{1}H sweep width (Hz): 7211.539 / 1583.899 / 4000.000 • ^{1}H / ^{15}N / ^{1}H carrier (Hz): 2828.00 / 7278.86 / 2828.00 • ^{1}H / ^{15}N / ^{1}H acquisition time (ms): 62.8 / 23.9 / 13.0 • Temp: 50 °C • Buffer in 5% D_2O: 20 mM Tris, 100 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Shaped tube
3D ^{15}N-edited-NOESY-HSQC NMR for labeled fsKaiB^{N29A} + 1.1x unlabeled CikA^{PSR} For intramolecular NOE of labeled fsKaiB^{N29A} in complex	1) [fsKaiB^{N29A}]-KaiB_FLAG_Y8A-Y94A-G89A-D91R_N29A_1-98 (U-[^{15}N, ^{13}C] labeled) / 800 2) [CikA^{PSR}]-CikA_613-729 / 880	• Volume: 380 µL • ^{1}H / ^{15}N / ^{1}H sweep width (Hz): 9615.385 / 1583.899 / 5590.535 • ^{1}H / ^{15}N / ^{1}H carrier (Hz): 2828.00 / 7278.86 / 2828.00 • ^{1}H / ^{15}N / ^{1}H acquisition time (ms): 67.8 / 21.4 / 20.0 • Mixing time (ms): 12 • Temp: 50 °C • Buffer in 5% D_2O: 20 mM Tris, 100 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Shaped tube

3D ^{13}C-edited-NOESY-HSQC NMR for labeled fsKaiB^{N29A} + 1.1x unlabeled CikA^{PSR} For intramolecular NOE of labeled fsKaiB^{N29A} in complex	1) [fsKaiB^{N29A}]-KaiB_FLAG_Y8A-Y94A-G89A-D91R_N29A_1-98 (U-[^{15}N, ^{13}C] labeled) / 800 2) [CikA^{PSR}]-CikA_613-729 / 880	• Volume: 380 µL • ^{1}H / ^{13}C / ^{1}H sweep width (Hz): 7211.539 / 3023.204 / 3727.018 • ^{1}H / ^{13}C / ^{1}H carrier (Hz): 1918.60 / 5944.17 / 1918.60 • ^{1}H / ^{13}C / ^{1}H acquisition time (ms): 59.9 / 10.6 / 19.0 • Mixing time (ms): 12 • Temp: 50 °C • Buffer in 99.96% D_2O: 20 mM Tris, 100 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Shaped tube
3D ^{13}C-HCCH-COSY NMR for labeled fsKaiB^{N29A} + 1.1x unlabeled CikA^{PSR} Proton version	1) [fsKaiB^{N29A}]-KaiB_FLAG_Y8A-Y94A-G89A-D91R_N29A_1-98 (U-[^{15}N, ^{13}C] labeled) / 800 2) [CikA^{PSR}]-CikA_613-729 / 880	• Volume: 380 µL • ^{1}H / ^{13}C / ^{1}H sweep width (Hz): 7211.539 / 3023.204 / 4809.062 • ^{1}H / ^{13}C / ^{1}H carrier (Hz): 1918.60 / 5944.17 / 1918.60 • ^{1}H / ^{13}C / ^{1}H acquisition time (ms): 59.9 / 10.6 / 12.9 • Temp: 50 °C • Buffer in 99.96% D_2O: 20 mM Tris, 100 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Shaped tube
3D ^{13}C-HCCH-COSY NMR for labeled fsKaiB^{N29A} + 1.1x unlabeled CikA^{PSR} Carbon version	1) [fsKaiB^{N29A}]-KaiB_FLAG_Y8A-Y94A-G89A-D91R_N29A_1-98 (U-[^{15}N, ^{13}C] labeled) / 800 2) [CikA^{PSR}]-CikA_613-729 / 880	• Volume: 380 µL • ^{1}H / ^{13}C / ^{13}C sweep width (Hz): 7211.539 / 3023.204 / 3023.204 • ^{1}H / ^{13}C / ^{13}C carrier (Hz): 1918.60 / 5944.17 / 5944.17 • ^{1}H / ^{13}C / ^{13}C acquisition time (ms): 59.9 / 10.6 / 10.6 • Temp: 50 °C • Buffer in 99.96% D_2O: 20 mM Tris, 100 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Shaped tube

3D ^{13}C-HCCH-TOCSY NMR for labeled fsKaiB^{N29A} + 1.1x unlabeled CikA^{PSR} Carbon version	1) [fsKaiB^{N29A}]-KaiB_FLAG_Y8A-Y94A-G89A-D91R_N29A_1-98 (U-[^{15}N, ^{13}C] labeled) / 800 2) [CikA^{PSR}]-CikA_613-729 / 880	• Volume: 380 µL • ^{1}H / ^{13}C / ^{13}C sweep width (Hz): 7211.539 / 3023.204 / 3023.204 • ^{1}H / ^{13}C / ^{13}C carrier (Hz): 1918.60 / 5944.17 / 5944.17 • ^{1}H / ^{13}C / ^{13}C acquisition time (ms): 59.9 / 10.6 / 10.6 • Temp: 50 °C • Buffer in 99.96% D_2O: 20 mM Tris, 100 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Shaped tube
3D HCAN and HCA(CO)N NMR for labeled fsKaiB^{N29A} + 1.1x unlabeled CikA^{PSR}	1) [fsKaiB^{N29A}]-KaiB_FLAG_Y8A-Y94A-G89A-D91R_N29A_1-98 (U-[^{15}N, ^{13}C] labeled) / 800 2) [CikA^{PSR}]-CikA_613-729 / 880	• Volume: 380 µL • ^{1}H / ^{13}C / ^{15}N sweep width (Hz): 7211.539 / 3023.247 / 1583.899 • ^{1}H / ^{13}C / ^{15}N carrier (Hz): 2850.40 / 8086.90 / 7278.98 • ^{1}H / ^{13}C / ^{15}N acquisition time (ms): 59.9 / 10.6 / 23.9 • Temp: 50 °C • Buffer in 99.96% D_2O: 20 mM Tris, 100 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Shaped tube
3D IPAP-J-HNCO(CA) NMR for 5.5% stretched polyacrylamide gel RDC of labeled fsKaiB^{N29A} + 1.1x unlabeled CikA^{PSR} measures: ^{1}J(CA-HA)	1) [fsKaiB^{N29A}]-KaiB_FLAG_Y8A-Y94A-G89A-D91R_N29A_1-98 (U-[^{15}N, ^{13}C] labeled) / 700 2) [CikA^{PSR}]-CikA_613-729 / 770	• ^{1}H / ^{15}N / ^{13}C sweep width (Hz): 7211.539 / 1583.899 / 1500.000 • ^{1}H / ^{15}N / ^{13}C carrier (Hz): 2823.20 / 7278.86 / 26187.80 • ^{1}H / ^{15}N / ^{13}C acquisition time (ms): 67.9 / 23.3 / 50.6 • Temp: 50 °C • Buffer in 10% D_2O: 10 mM Tris, 50 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Aligned sample: 4.2mm new era NMR tubes (NE-UP5-GT-7) • Isotropic sample: shaped tube

3D IPAP-J-HNCO NMR for 5.5% stretched polyacrylamide gel RDC of labeled fsKaiB^{N29A} + 1.1x unlabeled CikA^{PSR} measures: ^{1}J(N-H)	1) [fsKaiB^{N29A}]-KaiB_FLAG_Y8A-Y94A-G89A-D91R_N29A_1-98 (U-[^{15}N, ^{13}C] labeled) / 700 2) [CikA^{PSR}]-CikA_613-729 / 770	• ^{1}H / ^{15}N / ^{13}C sweep width (Hz): 7211.539 / 1583.899 / 1500.004 • ^{1}H / ^{15}N / ^{13}C carrier (Hz): 2823.20 / 7278.86 / 26187.80 • ^{1}H / ^{15}N / ^{13}C acquisition time (ms): 67.9 / 46.7 / 19.9 • Temp: 50 °C • Buffer in 10% D_2O: 10 mM Tris, 50 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Aligned sample: 4.2mm new era NMR tubes (NE-UP5-GT-7) • Isotropic sample: shaped tube
3D IPAP-J-HNCO NMR for 5.5% stretched polyacrylamide gel RDC of labeled fsKaiB^{N29A} + 1.1x unlabeled CikA^{PSR} measures: ^{1}J(CA-CO)	1) [fsKaiB^{N29A}]-KaiB_FLAG_Y8A-Y94A-G89A-D91R_N29A_1-98 (U-[^{15}N, ^{13}C] labeled) / 700 2) [CikA^{PSR}]-CikA_613-729 / 770	• ^{1}H / ^{15}N / ^{13}C sweep width (Hz): 7211.539 / 1583.899 / 1500.000 • ^{1}H / ^{15}N / ^{13}C carrier (Hz): 2823.20 / 7278.86 / 26187.80 • ^{1}H / ^{15}N / ^{13}C acquisition time (ms): 67.9 / 23.3 / 50.6 • Temp: 50 °C • Buffer in 10% D_2O: 10 mM Tris, 50 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Aligned sample: 4.2mm new era NMR tubes (NE-UP5-GT-7) • Isotropic sample: shaped tube

4D ^{15}N, ^{13}C-edited-NOESY-HSQC NMR for labeled fsKaiB^{N29A} + 1.1x unlabeled CikA^{PSR} For intermolecular NOE	1) [fsKaiB^{N29A}]-KaiB_FLAG_Y8A-Y94A-G89A-D91R_N29A_1-98 (U-[^{15}N] labeled) / 800 2) [CikA^{PSR}]-CikA_613-729 (U-[^{13}C] labeled) / 880	• Volume: 380 µL • ^{1}H / ^{15}N / ^{1}H / ^{13}C sweep width (Hz): 7211.539 / 1583.899 / 4809.062 / 3023.208 • ^{1}H / ^{15}N / ^{1}H / ^{13}C carrier (Hz): 2828.00 / 7278.86 / 2828.00 / 6121.94 • ^{1}H / ^{15}N / ^{1}H / ^{13}C acquisition time (ms): 62.8 / 10.1 / 11.8 / 5.2 • Mixing time (ms): 18 • Temp: 50 °C • Buffer in 5% D_2O: 20 mM Tris, 100 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Shaped tube
3D ^{13}C-edited, ^{12}C-filtered NOESY-HSQC NMR for labeled fsKaiB^{N29A} + 1.1x unlabeled CikA^{PSR} For intermolecular NOE	1) [fsKaiB^{N29A}]-KaiB_FLAG_Y8A-Y94A-G89A-D91R_N29A_1-98 (U-[^{15}N, ^{13}C] labeled) / 800 2) [CikA^{PSR}]-CikA_613-729 / 880	• Volume: 380 µL • ^{1}H / ^{13}C / ^{1}H sweep width (Hz): 7211.539 / 3023.204 / 3606.792 • ^{1}H / ^{13}C / ^{1}H carrier (Hz): 1918.60 / 5944.17 / 1918.60 • ^{1}H / ^{13}C / ^{1}H acquisition time (ms): 59.9 / 11.2 / 19.9 • Mixing time (ms): 15 • Temp: 50 °C • Buffer in 99.96% D_2O: 20 mM Tris, 100 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Shaped tube
2D IPAP-J-HSQC NMR for 5.5% stretched polyacrylamide gel RDC of labeled fsKaiB^{N29A} + 1.1x unlabeled CikA^{PSR} measures: ^{1}J(N-H)	1) [fsKaiB^{N29A}]-KaiB_FLAG_Y8A-Y94A-G89A-D91R_N29A_1-98 (U-[^{15}N] labeled) / 400 2) [CikA^{PSR}]-CikA_613-729 / 440	• ^{1}H / ^{15}N sweep width (Hz): 7211.539 / 1583.899 • ^{1}H / ^{15}N carrier (Hz): 2835.00 / 7278.86 • ^{1}H / ^{15}N acquisition time (ms): 62.6 / 202 • Temp: 50 °C • Buffer in 10% D_2O: 10 mM Tris, 50 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Aligned sample: 4.2mm new era NMR tubes (NE-UP5-GT-7) • Isotropic sample: shaped tube

3D HNCACB and HN(CO)CACB NMR for labeled $CikA^{PSR}$ + 1.1x unlabeled $fsKaiB^{N29A}$	1) [$CikA^{PSR}$]-CikA_613-729 (U-[^{15}N, ^{13}C] labeled) / 800 2) [$fsKaiB^{N29A}$]-KaiB_FLAG_Y8A-Y94A-G89A-D91R_N29A_1-98 / 880	• Volume: 380 µL • ^{1}H / ^{15}N / ^{13}C sweep width (Hz): 7211.539 / 1583.899 / 9500.000 • ^{1}H / ^{15}N / ^{13}C carrier (Hz): 2828.00 / 7278.86 / 6575.41 • ^{1}H / ^{15}N / ^{13}C acquisition time (ms): 62.8 / 23.9 / 7.0 • Temp: 50 °C • Buffer in 5% D_2O: 20 mM Tris, 100 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Shaped tube
3D HN(CA)CO and HNCO NMR for labeled $CikA^{PSR}$ + 1.1x unlabeled $fsKaiB^{N29A}$	1) [$CikA^{PSR}$]-CikA_613-729 (U-[^{15}N, ^{13}C] labeled) / 800 2) [$fsKaiB^{N29A}$]-KaiB_FLAG_Y8A-Y94A-G89A-D91R_N29A_1-98 / 880	• Volume: 380 µL • ^{1}H / ^{15}N / ^{13}C sweep width (Hz): 7211.539 / 1583.899 / 1500.004 • ^{1}H / ^{15}N / ^{13}C carrier (Hz): 2828.00 / 7278.86 / 26225.41 • ^{1}H / ^{15}N / ^{13}C acquisition time (ms): 62.8 / 23.3 / 30.0 • Temp: 50 °C • Buffer in 5% D_2O: 20 mM Tris, 100 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Shaped tube
3D HBHA(CO)NH NMR for labeled $CikA^{PSR}$ + 1.1x unlabeled $fsKaiB^{N29A}$	1) [$CikA^{PSR}$]-CikA_613-729 (U-[^{15}N, ^{13}C] labeled) / 800 2) [$fsKaiB^{N29A}$]-KaiB_FLAG_Y8A-Y94A-G89A-D91R_N29A_1-98 / 880	• Volume: 380 µL • ^{1}H / ^{15}N / ^{1}H sweep width (Hz): 7211.539 / 1583.899 / 4000.000 • ^{1}H / ^{15}N / ^{1}H carrier (Hz): 2828.00 / 7278.86 / 2828.00 • ^{1}H / ^{15}N / ^{1}H acquisition time (ms): 62.8 / 23.9 / 13.0 • Temp: 50 °C • Buffer in 5% D_2O: 20 mM Tris, 100 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Shaped tube

3D ^{15}N-edited-NOESY-HSQC NMR for labeled CikA^{PSR} + 1.1x unlabeled fsKaiB^{N29A} For intramolecular NOE of labeled CikA^{PSR} in complex	1) [CikA^{PSR}]-CikA_613-729 (U-[^{15}N, ^{13}C] labeled) / 800 2) [fsKaiB^{N29A}]-KaiB_FLAG_Y8A-Y94A-G89A-D91R_N29A_1-98 / 880	• Volume: 380 µL • ^{1}H / ^{15}N / ^{1}H sweep width (Hz): 9615.385 / 1583.899 / 5590.535 • ^{1}H / ^{15}N / ^{1}H carrier (Hz): 2828.00 / 7278.86 / 2828.00 • ^{1}H / ^{15}N / ^{1}H acquisition time (ms): 67.8 / 21.4 / 20.0 • Mixing time (ms): 12 • Temp: 50 °C • Buffer in 5% D_2O: 20 mM Tris, 100 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Shaped tube
3D ^{13}C-edited-NOESY-HSQC NMR for labeled CikA^{PSR} + 1.1x unlabeled fsKaiB^{N29A} For intramolecular NOE of labeled fsKaiB^{N29A} in complex	1) [CikA^{PSR}]-CikA_613-729 (U-[^{15}N, ^{13}C] labeled) / 800 2) [fsKaiB^{N29A}]-KaiB_FLAG_Y8A-Y94A-G89A-D91R_N29A_1-98 / 880	• Volume: 380 µL • ^{1}H / ^{13}C / ^{1}H sweep width (Hz): 7211.539 / 3023.204 / 4027.584 • ^{1}H / ^{13}C / ^{1}H carrier (Hz): 1918.60 / 5944.17 / 1918.60 • ^{1}H / ^{13}C / ^{1}H acquisition time (ms): 59.9 / 10.6 / 19.1 • Mixing time (ms): 12 • Temp: 50 °C • Buffer in 99.96% D_2O: 20 mM Tris, 100 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Shaped tube
3D ^{13}C-HCCH-COSY NMR for labeled CikA^{PSR} + 1.1x unlabeled fsKaiB^{N29A} Proton version	1) [CikA^{PSR}]-CikA_613-729 (U-[^{15}N, ^{13}C] labeled) / 800 2) [fsKaiB^{N29A}]-KaiB_FLAG_Y8A-Y94A-G89A-D91R_N29A_1-98 / 880	• Volume: 380 µL • ^{1}H / ^{13}C / ^{1}H sweep width (Hz): 7211.539 / 3023.204 / 4809.062 • ^{1}H / ^{13}C / ^{1}H carrier (Hz): 1918.60 / 5944.17 / 1918.60 • ^{1}H / ^{13}C / ^{1}H acquisition time (ms): 59.9 / 10.6 / 12.9 • Temp: 50 °C • Buffer in 99.96% D_2O: 20 mM Tris, 100 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Shaped tube

3D ^{13}C-HCCH-COSY NMR for labeled CikA^{PSR} + 1.1x unlabeled fsKaiB^{N29A} Carbon version	1) [CikA^{PSR}]-CikA_613-729 (U-[^{15}N, ^{13}C] labeled) / 800 2) [fsKaiB^{N29A}]-KaiB_FLAG_Y8A-Y94A-G89A-D91R_N29A_1-98 / 880	• Volume: 380 µL • ^{1}H / ^{13}C / ^{13}C sweep width (Hz): 7211.539 / 3023.204 / 3023.204 • ^{1}H / ^{13}C / ^{13}C carrier (Hz): 1918.60 / 5944.17 / 5944.17 • ^{1}H / ^{13}C / ^{13}C acquisition time (ms): 59.9 / 10.6 / 10.6 • Temp: 50 °C • Buffer in 99.96% D_2O: 20 mM Tris, 100 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Shaped tube
3D ^{13}C-HCCH-TCOSY NMR for labeled CikA^{PSR} + 1.1x unlabeled fsKaiB^{N29A} Carbon version	1) [CikA^{PSR}]-CikA_613-729 (U-[^{15}N, ^{13}C] labeled) / 800 2) [fsKaiB^{N29A}]-KaiB_FLAG_Y8A-Y94A-G89A-D91R_N29A_1-98 / 880	• Volume: 380 µL • ^{1}H / ^{13}C / ^{13}C sweep width (Hz): 7211.539 / 3023.204 / 3023.204 • ^{1}H / ^{13}C / ^{13}C carrier (Hz): 1918.60 / 5944.17 / 5944.17 • ^{1}H / ^{13}C / ^{13}C acquisition time (ms): 59.9 / 10.6 / 10.6 • Temp: 50 °C • Buffer in 99.96% D_2O: 20 mM Tris, 100 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Shaped tube
3D IPAP-J-HNCO(CA) NMR for 5.5% stretched polyacrylamide gel RDC of labeled CikA^{PSR} + 1.1x unlabeled fsKaiB^{N29A} measures: ^{1}J(CA-HA)	1) [CikA^{PSR}]-CikA_613-729 (U-[^{15}N, ^{13}C] labeled) / 800 2) [fsKaiB^{N29A}]-KaiB_FLAG_Y8A-Y94A-G89A-D91R_N29A_1-98 / 880	• ^{1}H / ^{15}N / ^{13}C sweep width (Hz): 7211.539 / 1583.899 / 1500.000 • ^{1}H / ^{15}N / ^{13}C carrier (Hz): 2823.20 / 7278.86 / 26187.80 • ^{1}H / ^{15}N / ^{13}C acquisition time (ms): 67.9 / 23.3 / 50.6 • Temp: 50 °C • Buffer in 10% D_2O: 10 mM Tris, 50 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Aligned sample: 4.2mm new era NMR tubes (NE-UP5-GT-7) • Isotropic sample: shaped tube

3D IPAP-J-HNCO NMR for 5.5% stretched polyacrylamide gel RDC of labeled $CikA^{PSR}$ + 1.1x unlabeled $fsKaiB^{N29A}$ measures: ^{1}J(N-H)	1) [$CikA^{PSR}$]-CikA_613-729 (U-[^{15}N, ^{13}C] labeled) / 800 2) [$fsKaiB^{N29A}$]-KaiB_FLAG_Y8A-Y94A-G89A-D91R_N29A_1-98 / 880	• ^{1}H / ^{15}N / ^{13}C sweep width (Hz): 7211.539 / 1583.899 / 1500.004 • ^{1}H / ^{15}N / ^{13}C carrier (Hz): 2823.20 / 7278.86 / 26187.80 • ^{1}H / ^{15}N / ^{13}C acquisition time (ms): 67.9 / 46.7 / 19.9 • Temp: 50 °C • Buffer in 10% D_2O: 10 mM Tris, 50 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Aligned sample: 4.2mm new era NMR tubes (NE-UP5-GT-7) • Isotropic sample: shaped tube
3D IPAP-J-HNCO NMR for 5.5% stretched polyacrylamide gel RDC of labeled $CikA^{PSR}$ + 1.1x unlabeled $fsKaiB^{N29A}$ measures: ^{1}J(CA-CO)	1) [$CikA^{PSR}$]-CikA_613-729 (U-[^{15}N, ^{13}C] labeled) / 800 2) [$fsKaiB^{N29A}$]-KaiB_FLAG_Y8A-Y94A-G89A-D91R_N29A_1-98 / 880	• ^{1}H / ^{15}N / ^{13}C sweep width (Hz): 7211.539 / 1583.899 / 1500.000 • ^{1}H / ^{15}N / ^{13}C carrier (Hz): 2823.20 / 7278.86 / 26187.80 • ^{1}H / ^{15}N / ^{13}C acquisition time (ms): 67.9 / 23.3 / 50.6 • Temp: 50 °C • Buffer in 10% D_2O: 10 mM Tris, 50 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Aligned sample: 4.2mm new era NMR tubes (NE-UP5-GT-7) • Isotropic sample: shaped tube

4D ^{15}N, ^{13}C-edited-NOESY-HSQC NMR for labeled $CikA^{PSR}$ + 1.1x unlabeled $fsKaiB^{N29A}$ For intermolecular NOE	1) [$CikA^{PSR}$]-CikA_613-729 (U-[^{15}N] labeled) / 800 2) [$fsKaiB^{N29A}$]-KaiB_FLAG_Y8A-Y94A-G89A-D91R_N29A_1-98 (U-[^{13}C] labeled) / 880	• Volume: 380 µL • ^{1}H / ^{15}N / ^{1}H / ^{13}C sweep width (Hz): 7211.539 / 1583.899 / 4809.062 / 3023.208 • ^{1}H / ^{15}N / ^{1}H / ^{13}C carrier (Hz): 2828.00 / 7278.86 / 2828.00 / 6121.94 • ^{1}H / ^{15}N / ^{1}H / ^{13}C acquisition time (ms): 62.8 / 10.1 / 11.8 / 5.2 • Mixing time (ms): 15 • Temp: 50 °C • Buffer in 5% D_2O: 20 mM Tris, 100 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Shaped tube
3D ^{13}C-edited, ^{12}C-filtered, NOESY-HSQC NMR for labeled $CikA^{PSR}$ + 1.1x unlabeled $fsKaiB^{N29A}$ For intermolecular NOE	1) [$CikA^{PSR}$]-CikA_613-729 (U-[^{15}N, ^{13}C] labeled) / 800 2) [$fsKaiB^{N29A}$]-KaiB_FLAG_Y8A-Y94A-G89A-D91R_N29A_1-98 / 880	• Volume: 380 µL • ^{1}H / ^{13}C / ^{1}H sweep width (Hz): 7211.539 / 3023.204 / 3606.792 • ^{1}H / ^{13}C / ^{1}H carrier (Hz): 1918.60 / 5944.17 / 1918.60 • ^{1}H / ^{13}C / ^{1}H acquisition time (ms): 59.9 / 11.2 / 19.9 • Mixing time (ms): 15 • Temp: 50 °C • Buffer in 99.96% D_2O: 20 mM Tris, 100 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Shaped tube
2D IPAP-J-HSQC NMR for 5.5% stretched polyacrylamide gel RDC of labeled $CikA^{PSR}$ + 1.1x unlabeled $fsKaiB^{N29A}$ measures: ^{1}J(N-H)	3) [$CikA^{PSR}$]-CikA_613-729 (U-[^{15}N] labeled) / 400 4) [$fsKaiB^{N29A}$]-KaiB_FLAG_Y8A-Y94A-G89A-D91R_N29A_1-98 / 440	• ^{1}H / ^{15}N sweep width (Hz): 7211.539 / 1583.901 • ^{1}H / ^{15}N carrier (Hz): 2835.00 / 7278.86 • ^{1}H / ^{15}N acquisition time (ms): 62.6 / 202 • Temp: 50 °C • Buffer in 10% D_2O: 10 mM Tris, 50 mM NaCl, 5 mM TCEP, 0.02% NaN_3, 10 µM DSS, pH 7, 1x protease inhibitor • Aligned sample: 4.2mm new era NMR tubes (NE-UP5-GT-7) • Isotropic sample: shaped tube

Table 3.2. Structural statistics by X-ray crystallography

	Binary fsKaiB-CI*	Ternary $^{\Delta N}$KaiA^{C272S}-fsKaiB-CI*
Data collection		
Space group	P2 2_1 2_1	P1
Unit cell parameters		
a,b,c (Å)	42.48, 56.59, 146.14	74.597, 79.104, 80.937
α,β,γ (º)	90, 90, 90	107.273,90.71,111.199
Wavelength (Å)	1.11	1.11
Resolution (Å)	73.08-1.689	50.00-2.70
Total reflections	109,956	229,504
Unique reflections	32,797	114,752
Redundancy	3.5 (3.6)	2
Completeness (%)	92.9 (99.3)	96.3
R_{Sym}	0.033 (0.447)	(0.443)
*R_{merge}	0.049*	R_{linear} .056* (.387)
Mean I/I_σ	6.5	23.64* (3.5)
$CC_{1/2}$	.987* (0.701)	0.997 inner (0.757)
Refinement		
Resolution (Å)	1.8	2.7
Number of reflections	32,745	42,121
Number of atoms	1601	9510
R-factor/R_{free}	21.88/25.65	20.06/24.49
rmsd bond lengths (Å)	.014	0.007
rmsd bond angles (°)	1.544	1.184
Ramachandran plot (%)		
Most Favored regions	96.7	96.7
Additional allowed regions	3	3
Disallowed regions	0.3	0.26

*Overall measurement
( ) outer measurement
Numbers in the parenthesis refer to the outer shell.

Table 3.3. Structural statistics by NMR

Restraints and Structural Statistics for top 10 structures	fsKaiB^{N29A}	CikA^{PSR}	CikA^{PSR}- fsKaiB^{N29A} complex
Restraints			
Total distance restraints	2298	2766	2698
Intra residue (|j-i|=0)	711	961	1109
Sequential (|j-i|=1)	530	628	527
Medium range (2 ≤ |j-i| ≥ 4)	466	494	349
Long range (|j-I| ≥ 5	591	683	517
Intermolecular restraints	n/a	n/a	196
Total dihedral restraints	199	217	416
Total RDC restraints	145	165	154
NOE violations > 0.5 Å	0	0	0
Dihedral violations > 0.5°	0	0	0
Q factor[1]	0.304	0.419	0.327
Deviations from ideal geometry[2]			
Bonds (Å)	0.004	0.004	0.004
Angles (deg)	0.63 ± 0.04	0.58 ± 0.04	0.58 ± 0.04
Molprobity Ramachandran and sidechain analysis[3]			
Most favored regions %	98.2 ± 0.8	97.0 ± 0.8	97.2 ± 1.3
Allowed regions %	1.7 ± 0.8	2.3 ± 0.8	2.3 ± 1.2
Disallowed regions %	0.1 ± 0.3	0.0 ± 0.0	0.5 ± 0.5
Favored rotamers %	89.3 ± 1.6	92.2 ± 1.9	89.3 ± 1.3
Poor rotamers %	6.6 ± 1.5	4.3 ± 1.5	6.2 ± 1.4
Atomic RMSDs from the mean structure	Residues 9-98	Residues 615-692, 702-724	Residues 615-692, 702-724 and 9-98
Backbone heavy atoms (Å)	0.17 ± 0.06	0.46 ± 0.16	0.88 ± 0.35
Non-hydrogen atoms (Å)	0.99 ± 0.33	1.25 ± 0.40	1.52 ± 0.52

[1] Q factor value is obtained by inputting the average structure into “DC: Servers for Dipolar Coupling Calculations” webserver (http://spin.niddk.nih.gov/bax/nmrserver/dc/fit.html).
[2] Calculated by PSVS structure validation webserver (http://psvs-1_5-dev.nesg.org) (134).
[3] Calculated by Molprobity structure validation webserver (http://molprobity.biochem.duke.edu) (135).

Results

Optimize constructs for structure determination of nighttime complexes by X-ray crystallography and NMR

Cyanobacteria are highly adaptive groups that inhabit diverse environments, such as freshwater, oceans, hot springs, desserts, and even the artic (136, 137). Two species of cyanobacteria, *Synechococcus elongatus* (freshwater) and *Thermosynechococcus elongatus* (hot spring) have emerged as two complementary model organisms, in that (1) they demonstrate bona fide circadian rhythms in vivo (24, 138); (2) their clock components share a high degree of homology as well as near-identical structures; and (3) the former has a plethora of genetic tools for functional studies, whereas the latter is more suitable for structural studies owing to their greater thermodynamic stability (24, 109). Therefore, structural investigations in this study were carried out using the clock proteins from the thermophile *T. elongatus* for its superior stability. Functional studies were carried out in *S. elongatus*. For clarity, proteins of *S. elongatus* are distinguished by adding "se" in front of the protein or gene name (e.g. seKaiA or *sekaiA*).

It has been recently reported that the nighttime binary KaiB-KaiC complex can be assembled using the CI domain of KaiC and fsKaiB (48, 60). Additionally, the pseudo-receiver domain of CikA, $CikA^{PSR}$, has been shown to be the minimal domain responsible for binding to KaiB (59). Therefore, these functional domains were used to produce high quality crystals and heat-tolerant NMR samples for structure determination by X-ray crystallography and NMR.

Figure 3.1 shows an overview of all the engineered protein constructs used in this study. Crystallization trials for the binary KaiB-KaiC complex used fsKaiB and CI* (a stable CI monomer variant) (59). It has been shown that the N-terminal domain of KaiA is not necessary for forming the ternary KaiA-KaiB-KaiC complex (52, 60). Thus, we used an N-terminally truncated KaiA ($^{\Delta N}$KaiA) for crystallography. To further improve sample homogeneity of the ternary complex for the crystallization trails, we introduced a C272S substitution into $^{\Delta N}$KaiA, $^{\Delta N}KaiA^{C272S}$, to prevent possible disulfide bond formation over time that has been observed in KaiA (109). Since $CikA^{PSR}$ has been shown to bind directly to fsKaiB by chemical shifts perturbation (59), we investigated the details of $CikA^{PSR}$-fsKaiB binding with NMR. Fortuitously, we observed that introducing an additional N29A mutation on fsKaiB enhances its binding affinity to $CikA^{PSR}$. Thus, $fsKaiB^{N29A}$ and $CikA^{PSR}$ were used for structure delineation of the output complex.

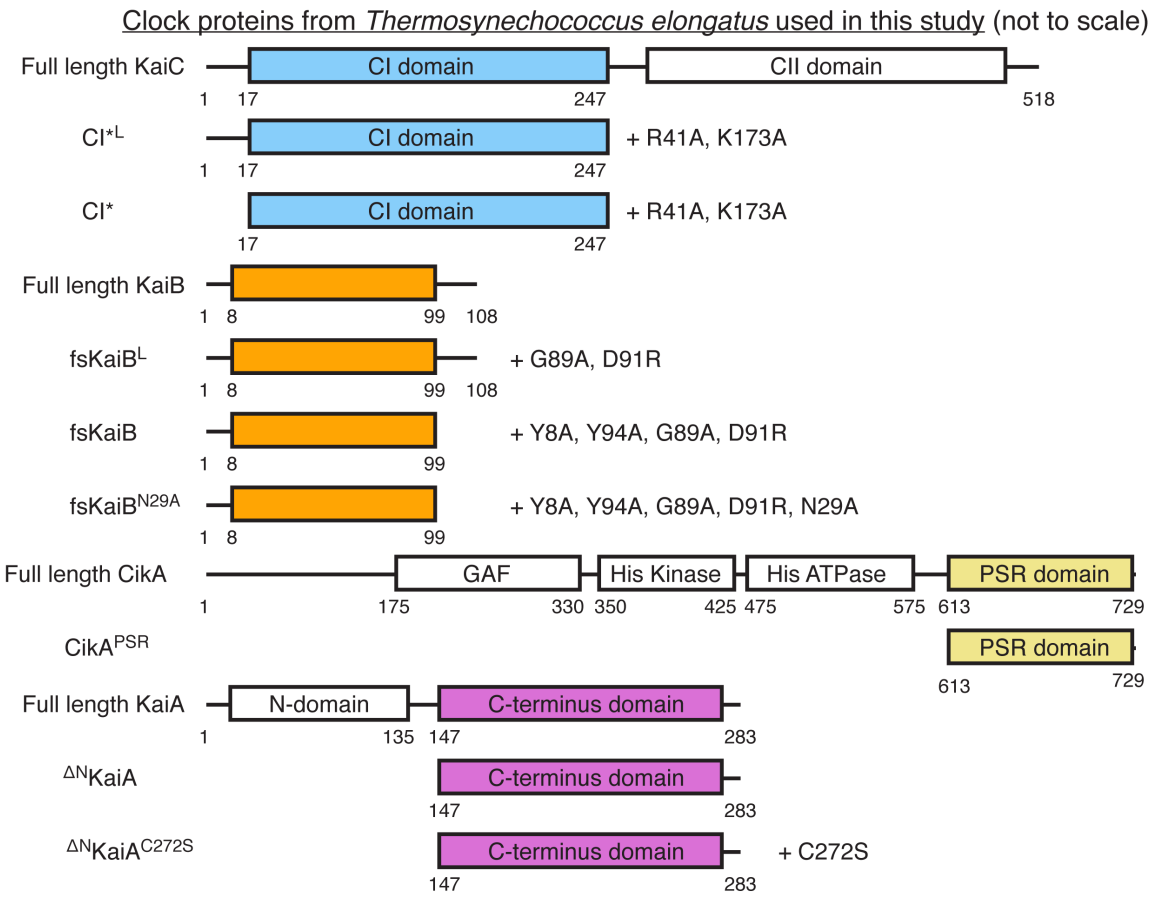


Figure 3.1. Overview of the engineered constructs for clock proteins used in this study.

The star symbol (*) on CI indicates a 16-residue N-terminal truncation and two mutations, R41A and K173A, which were introduced to prevent CI from hexamerization (52). CI*^{L}, a longer version of CI*, included the N-terminal 16-residues. fsKaiB was truncated after residue D99, had two stabilizing mutations, Y8A and Y94A, and two fold switch-inducing mutations, G89A and D91R (59). fsKaiB^{L} was a longer version of fsKaiB by having the C-terminal segment but without Y8A and Y94A substitutions. fsKaiB^{N29A} had one extra mutation, N29A, relative fsKaiB, and exhibited enhanced affinity for CikA^{PSR}. $^{\Delta N}$KaiA was without the N-terminal domain (residues 1-146). $^{\Delta N}$KaiA bound to KaiB-KaiC complexes normally (60). $^{\Delta N}$KaiA^{C272S} harbored an extra mutation, C272S, which was introduced to prevent disulfide bond formation (109). For construct details please see Table 3.1.

The binary CI-fsKaiB complex revealed that CI, in a state resembling the ADP-bound post-hydrolysis state, uses its B loop and immediate α5 C-terminal helix to engage fsKaiB.*

The first step of the day-to-night transition in the cyanobacterial clock is the formation of a KaiB-KaiC complex (41). The critical KaiB-KaiC complex forms a hub-like platform to 1) disassociate SasA from KaiC, 2) sequester KaiA away from the A loop on the CII domain of KaiC, and 3) bind to $CikA^{PSR}$ to activate its phosphatase activity (40). In line with this, the deletion of *kaiB* in cyanobacterial cells resulted in the phenotype that cells could not produce nighttime gene expression patterns. Furthermore, those mutants also exhibited cell division defects (95).

Here, we solved the nighttime initiating complex of fsKaiB-CI* at 1.8 Å resolution (Figure 3.2 & Table 3.2). fsKaiB bound to the α5 helix and B-loop region (residues 116-123) of CI*, utilizing hydrophobic and hydrogen bonding interactions (Figure 3.2B & Table 3.4). Residues V118 and F122 of the B loop formed hydrophobic interactions with A61 and I77 of fsKaiB; also E117 of the B loop and R130 on α5 of CI* paired electrostatically with K85 and E55 of fsKaiB (Figure 3.2B). These interactions explain the previous report that E117A, V118A, F122A, R130A mutations on CI strongly diminished KaiB-KaiC binding (59). The KaiB-KaiC structure reveals why KaiB needs to switch folds (59) to interact with KaiC. Indeed, C-terminal residues of fsKaiB, P51, Q52, I59, L60, and K85, interact with α5 residues of CI, Y133, A126, L127 and E117, respectively.

ATPase hydrolysis on CI has been suggested to be essential for KaiB binding as abolishing CI ATPase activity and non-hydrolysable ATP analogs abolished KaiB-KaiC binding (53, 61). Superimposing the fsKaiB-CI* complex on published CI ring structures in both pre- and post-ATP hydrolysis states revealed that the tilt of α7 in CI* of the fsKaiB-CI* complex matched well to that of the post-hydrolysis state (139). This suggests that the tilt of α7 on CI may regulate KaiB-KaiC complex formation. However, the backbone and the sidechain conformations of α5 and B loop residues, which make direct interactions with fsKaiB, differ only marginally between pre- and post-hydrolysis states (Figure 3.2C), suggesting that quaternary structural changes of the entire CI ring, obviously not observable here, may play a large role in KaiB-KaiC binding. Thus a structure of KaiB bound to full-length KaiC may help reveal details of how ATP hydrolysis affects inter-subunit interfaces of the CI ring that serve to favor KaiB binding.

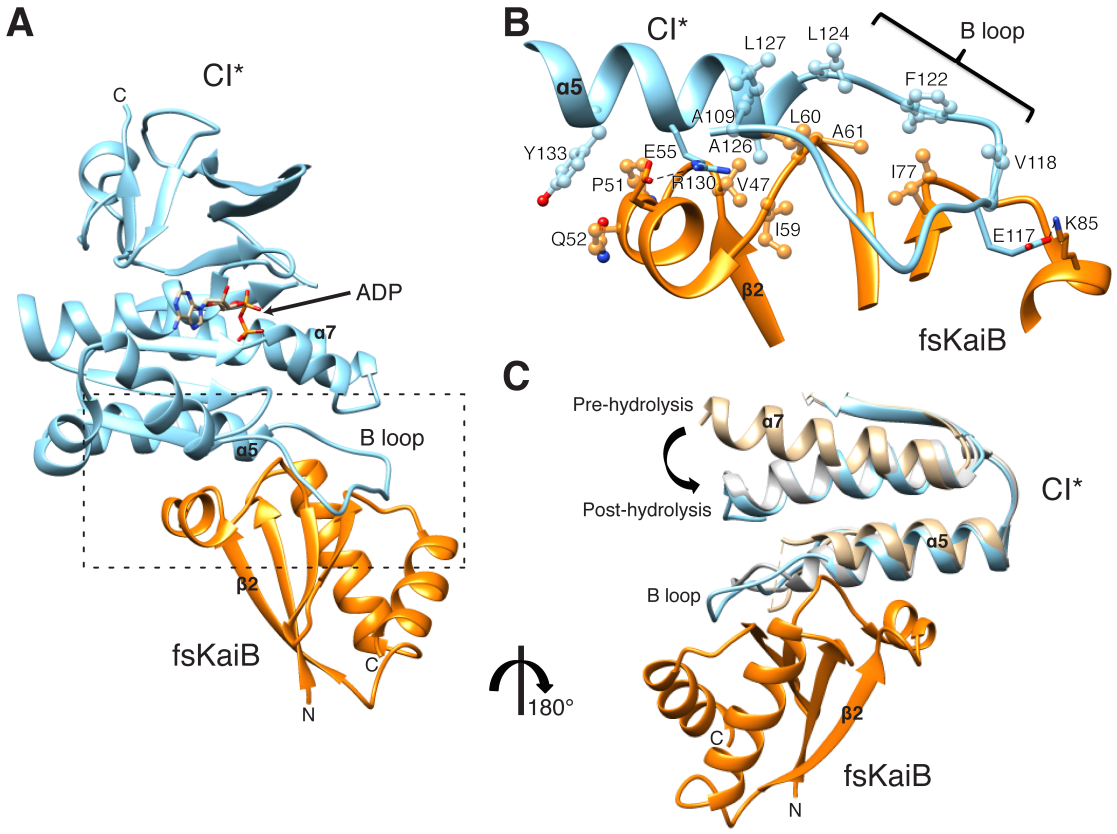


Figure 3.2. Crystal structure of the binary fsKaiB-CI* complex at 1.8 Å resolution.

(A) Structural overview of the binary fsKaiB-CI* complex. CI* (sky blue) has ADP bound and interacts with fsKaiB (orange) near B loop (residues 116-123) (60). (B) A close-up view of the interactions for the region boxed in (A). Hydrophobic and hydrophilic interactions are shown as balls and sticks, and sticks, respectively. (C) Superposition of the binary fsKaiB-CI* complex onto homologous seCI structures from *S. elongatus*. ATP pre-hydrolysis (brown) and post-hydrolysis (light gray) states are shown for chain E of PDB ID: 4TLA and 4TLB (139), respectively.

Table 3.4. Interface statistics of fsKaiB-CI*, $^{\Delta N}$KaiA^{C272S}-fsKaiB-CI* and CikA^{PSR}-fsKaiB^{N29A} complexes by PDBePISA.

Complex name	Interface partner (A:B)	Number of interface residues (A:B)	Interface area (Å^2)	$\Delta G_{hydropho}$ (kcal/M)[1]	Number of potential interface hydrogen bonds[2]	Number of potential interface salt bridges[3]
fsKaiB-CI*	fsKaiB:CI*	30:28	1000.0	-8.7	13	5
$^{\Delta N}$KaiA^{C272S}-fsKaiB-CI*	fsKaiB:CI*	28:31	1075.3	-6.5	19	5
	fsKaiB: $^{\Delta N}$KaiA^{C272S}	27:24	821.9	-7.0	12	2
	CI*: $^{\Delta N}$KaiA^{C272S}	2:1	43.3	0.8	2	1
CikA^{PSR}-fsKaiB^{N29A} [2]	CikA^{PSR}:fsKaiB^{N29A}	35:34	1357.3	-2.2	13	4

[1] $\Delta G_{hydropho}$ indicates the solvation free energy gain upon formation of the interface. Negative $\Delta G_{hydropho}$ corresponds to hydrophobic interfaces. This value does not include the effect of satisfied hydrogen bonds and salt bridges across the interface.
[2] The lowest energy model was used as input to PDBePISA.

PDBePISA official website (http://www.ebi.ac.uk/pdbe/prot_int/pistart.html) (140).

$^{\Delta N}$KaiA^{C272S}-fsKaiB-CI complex revealed the mechanism of cooperative KaiABC binding*

To transition from day to night, KaiB inhibits KaiA by forming a ternary KaiA-KaiB-KaiC complex. Ternary formation is cooperative, as supported by findings that KaiA associates with KaiB strongly only in the presence of CI (42, 59, 60) and that the presence of KaiA enhances KaiB-CI interactions (60). This cooperative binding also promotes disassociation of SasA from KaiC to ensure day/night transition in the output pathway (60). The structural basis of this cooperativity was revealed by solving the ternary $^{\Delta N}$KaiA^{C272S}-fsKaiB-CI* complex at 2.7 Å resolution (Figure 3.3 & Table 3.2).

As can be seen in Figure 3.3, fsKaiB is sandwiched between $^{\Delta N}$KaiA^{C272S} and CI*. This result is consistent with the previous NMR analysis that $^{\Delta N}$KaiA did not perturb the chemical shifts of CI domain significantly after KaiB-KaiC complex had already formed (60). In the structure of the complex, it can be seen that $^{\Delta N}$KaiA^{C272S} binds asymmetrically to KaiB, as was also suggested in a previous NMR report (60). Using a longer version of CI* (CI*^{L}), methyl-TROSY NMR (105) spectra of U-[^{15}N, ^{2}H]-Ile-δ1-[^{13}C, ^{1}H]-labeled $^{\Delta N}$KaiA showed that $^{\Delta N}$KaiA only interacts with the complex of wt-KaiB and CI*^{L}, but not with either wt-KaiB or CI*^{L} alone (Figure 3.4A & B). The chemical shifts of $^{\Delta N}$KaiA + fsKaiB + CI*^{L} matched almost identically to those of $^{\Delta N}$KaiA + wt-KaiB + CI*^{L}, verifying that the highly engineered fsKaiB and wt-KaiB bind to KaiA and CI similarly (Figure 3.4B). The spectrum of $^{\Delta N}$KaiA + fsKaiB showed both free and bound peak populations as compared to fully bound $^{\Delta N}$KaiA + wt-KaiB + CI*^{L} or $^{\Delta N}$KaiA + fsKaiB + CI*^{L} spectra (Figure 3.4B & C). This suggests that even with KaiB locked in the fold-switched structure (59), CI domain of KaiC still contributes to the cooperative binding of KaiA. Indeed, residues Y133 and K137 of α5 on CI* made three hydrogen bonds and one salt bridge to Q52 on fsKaiB and E161 on $^{\Delta N}$KaiA^{C272S}, forming a ternary interaction junction (Figure 3.3B). Previously, the presence of KaiA had been shown to enhance the binding affinity of KaiB and CI by ~10 fold (60). Consistent with this observation, the KaiB-CI interface adds six additional hydrogen bonds and 7% more interfacial area in the presence of KaiA (Figure 3.3B & Table 3.2). The presence of KaiA can strengthens the fsKaiB-CI* interface, providing a structural explanation for the cooperative binding between KaiA, KaiB and KaiC.

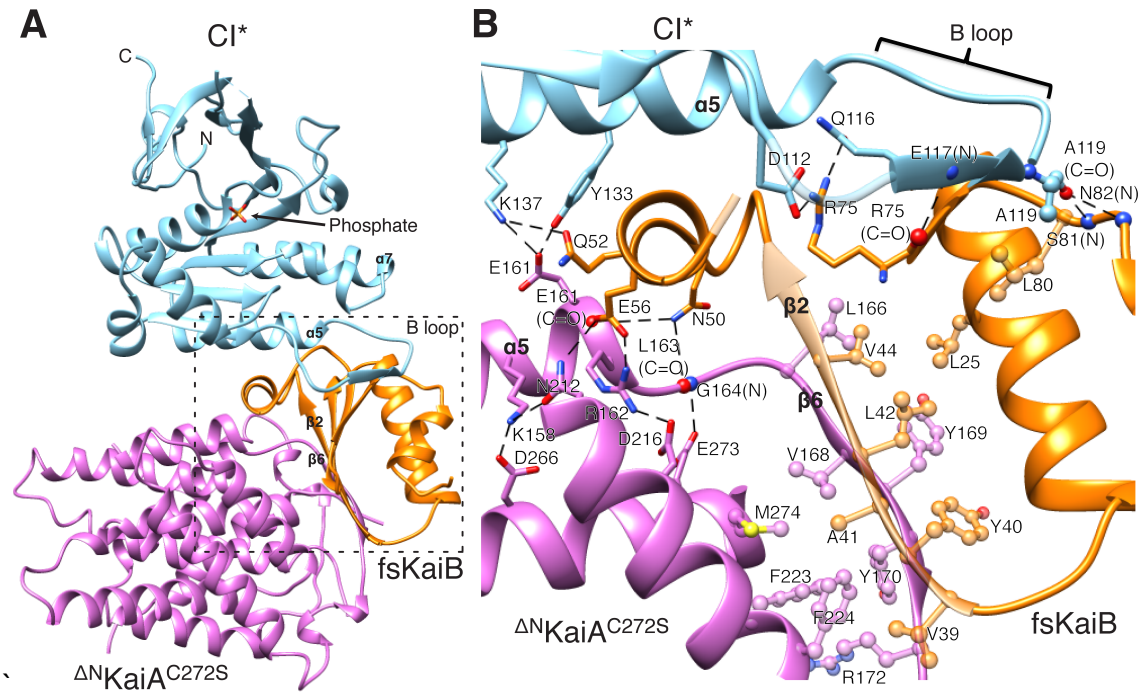


Figure 3.3. Crystal structure of the ternary $^{\Delta N}$KaiA^{C272S}-fsKaiB-CI* complex at 2.7 Å resolution.

(A) Structural overview of the ternary $^{\Delta N}$KaiA^{C272S}-fsKaiB-CI* complex. CI* (sky blue) has a phosphate bound and interacts with fsKaiB (orange), and $^{\Delta N}$KaiA^{C272S} (purple). (B) Detailed interactions for the region boxed in (A). At the fsKaiB-CI* interface, the interactions that are identical with those in the binary complex are omitted for clarity. Only new interactions in the ternary complex are shown. Hydrophobic and hydrophilic interactions are shown as balls and sticks, and sticks, respectively. Black dashed lines represent putative hydrogen bonds.

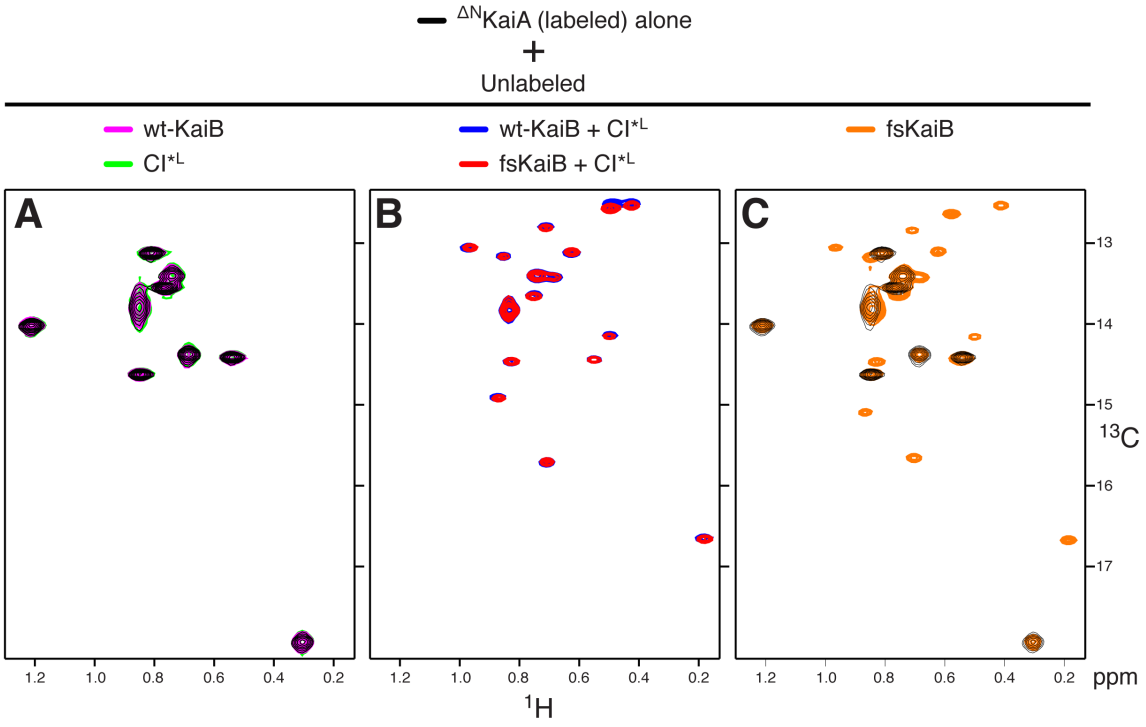


Figure 3.4. CI*^{L} enhances the $^{\Delta N}$KaiA-fsKaiB interaction.

(A) Methyl-TROSY spectra of U-[^{15}N, ^{2}H]-Ile-δ1-[^{13}C, 1H]–labeled 1× $^{\Delta N}$KaiA alone (black), or in the presence of 2.5× wt-KaiB (magenta), or 2.5× CI*^{L} (green). (B) Methyl-TROSY spectra of U-[^{15}N, ^{2}H]-Ile-δ1-[^{13}C, ^{1}H]–labeled 1× $^{\Delta N}$KaiA alone (black), or in the presence of 2.5× wt-KaiB + 2.5x CI*^{L} (blue) or 2.5× fsKaiB + 2.5x CI*^{L} (red). (C) Methyl-TROSY spectra of U-[^{15}N, ^{2}H]-Ile-δ1-[^{13}C, ^{1}H]–labeled 1× $^{\Delta N}$KaiA alone (black), or in the presence of 2.5× fsKaiB (orange). 1× corresponds to 20 μM monomer concentration. For experimental details please see Table 3.1.

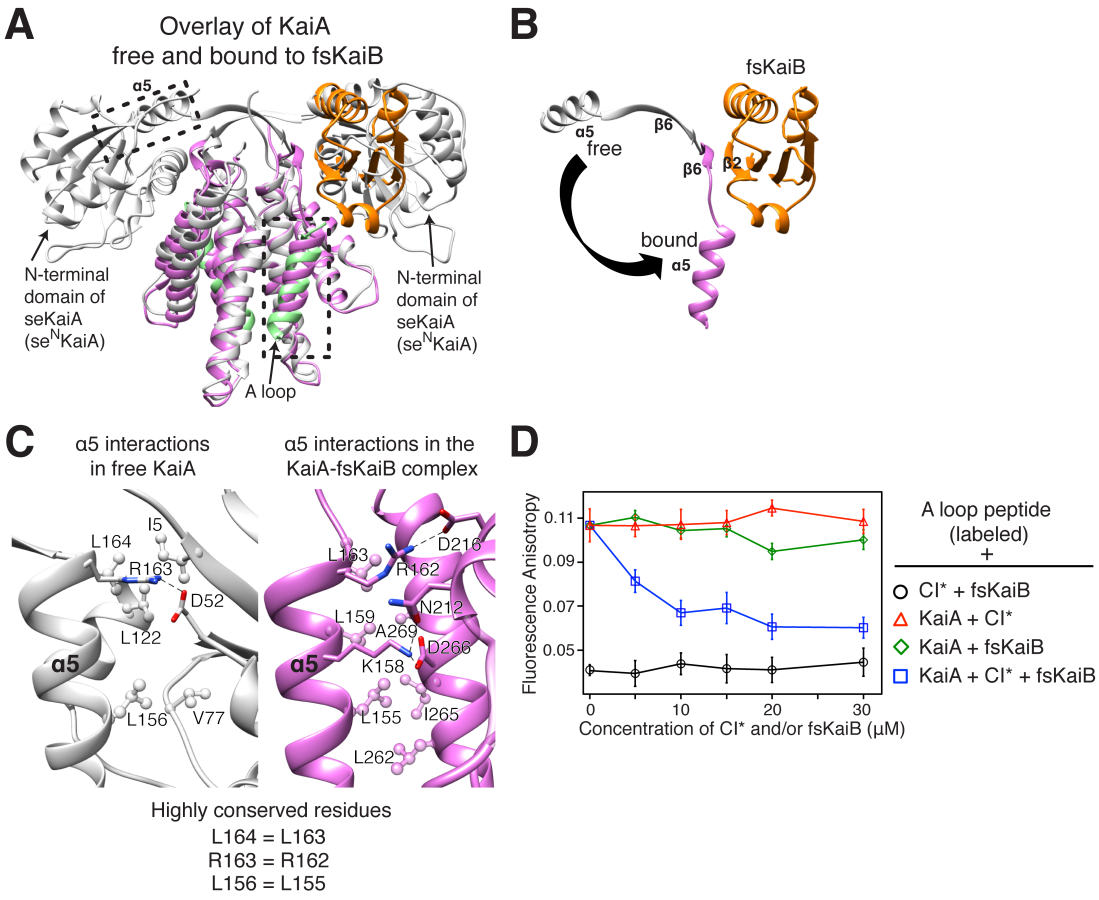


Figure 3.5. fsKaiB induces self-inhibiting conformation in $^{\Delta N}$KaiA^{C272S}.

(A) Overlay of the structure of seKaiA (light gray) bound to A-loop region peptide (green) (51) and the ternary complex (purple) in this study. N-terminal domain of seKaiA is termed se^{N}KaiA. PDB ID for seKaiA-A loop complex is 5C5E. (B) fsKaiB (orange), α5 helix of seKaiA (light gray) and $^{\Delta N}$KaiA^{C272S} (purple) are shown with the same orientation as in (A). (C) Zoomed-in of the two boxed regions in (A), showing homologous α5 interactions in free KaiA and in the KaiA-fsKaiB complex. Hydrophobic and hydrophilic interactions are shown as balls and sticks, and sticks, respectively. (D) Fluorescence anisotropy measurements of 6-IAF-labeled A-loop region peptide from KaiC + 10 μM of KaiA in the presence of increasing amounts of CI* (red) and/or fsKaiB (blue/green). Labeled A-loop region peptide in the presence of titrant CI* and fsKaiB but without KaiA is black. Error bar indicates standard deviation from triplicate measurements.

The ternary complex reveals the mechanism of KaiA inhibition by KaiB-induced conformational rearrangements of $^{\Delta N}$KaiA^{C272S}

In the ternary complex, $^{\Delta N}$KaiA^{C272S} interacted through a highly conserved β6 strand (108) to the β2 strand of fsKaiB (Figure 3.3B). This interface was the most hydrophobic in the ternary structure (Table 3.4). The β6-β2 interaction induced a large conformational change in α5 of $^{\Delta N}$KaiA^{C272S} (Figure 3.3B & 3.5). This new conformation produced polar interactions between α5 of KaiA and fsKaiB (Figure 3.3B & Table 3.4). No significant conformational differences were exhibited by fsKaiB between the binary and the ternary structures, except that Y40 flipped to make hydrophobic interactions with Y169 on $^{\Delta N}$KaiA^{C272S}. Superimposing the homologous seKaiA structure in complex with A-loop region peptides (51) showed that the α5 helix of $^{\Delta N}$KaiA^{C272S} and A-loop region peptide occupy the same binding site along the C-terminal domain of KaiA (Figure 3.5A), indicating that the new α5 position induced by fsKaiB is the mechanism of KaiA inactivation during subjective night (Figure 3.5). Interestingly, L156, R163 and L164 in α5 in seKaiA make intramolecular interactions to the N-terminal domain of seKaiA (Figure 3.5C). In the ternary complex, these highly conserved residues made key interactions to stabilize the new α5 position in autoinhibited $^{\Delta N}$KaiA^{C272S} (Figure 3.5C). As a test of the implied inactivation mechanism, we asked whether fsKaiB could induce disassociation of the A-loop region peptide (residues 488-518) from KaiA. Fluorescently labeled A-loop region peptide was displaced from KaiA upon titrating a mixture of CI* and fsKaiB (Figure 3.5D), as indicated by decreasing fluorescence anisotropy values. Independent verification of the ternary structure was obtained by hydrogen deuterium exchange mass spectroscopy (HDX-MS), which showed strong exchange protection on α5 of $^{\Delta N}$KaiA and β2 of fsKaiB in the $^{\Delta N}$KaiA-fsKaiB-CI* complex (Figure 3.6A). Structure-based mutagenesis on α5 of $^{\Delta N}$KaiA and β2 of fsKaiB^{L} (a longer version of fsKaiB) (Figure 3.1) identified four mutations, L155A, K158A, L159A and L163A that abolished ternary complex formation, as observed by gel-filtration chromatography (Figure 3.7). Thus, α5 allosterically regulates fsKaiB-KaiA interactions. A41D and K43E on β2 of fsKaiB^{L} also eliminated $^{\Delta N}$KaiA binding to CI*-fsKaiB (Figure 3.7). An ^{15}N, ^{1}H-HSQC spectrum of L155A-$^{\Delta N}$KaiA showed similarly dispersed peaks as compared with the spectrum of $^{\Delta N}$KaiA, suggesting that the L155A mutation did not affect the structural integrity of $^{\Delta N}$KaiA (Figure 3.8A). Therefore, we made the corresponding substitutions in *S. elongatus* homologs (Figure 3.9). As expected, L156A in *sekaiA*, and A40D/K42E in *sekaiB* disrupted cellular rhythms, matching our in vitro results (Figure 3.9A & B). Collectively, our structural, in vitro, and in vivo experiments verified and supported the proposed mechanism of KaiA self-inhibition induced by KaiB during the nighttime phase of the rhythm.

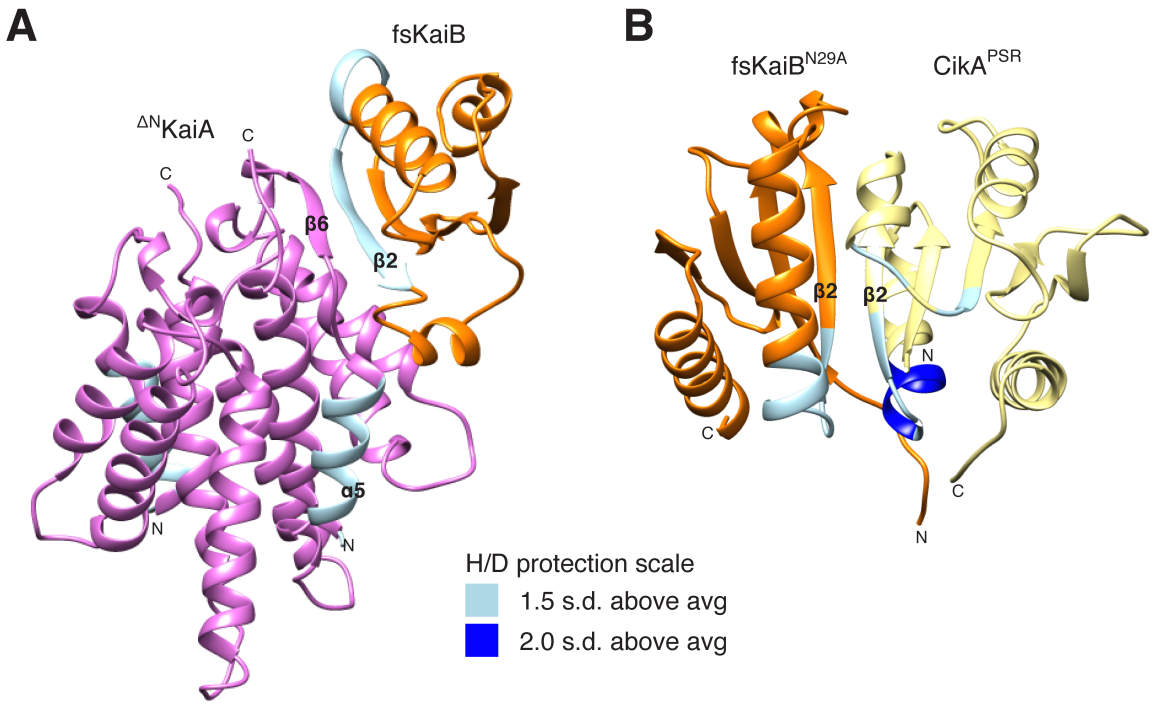


Figure 3.6. HDX-MS protection profiles for the complexes mapped onto the crystal and NMR structures.

(A) Mapping of relative protection against hydrogen-deuterium exchange of $^{\Delta N}$KaiA free and bound to the fsKaiB-CI* complex (purple); and fsKaiB-CI* complex free and bound to $^{\Delta N}$KaiA (orange). CI* protection profiles for the binary and ternary complexes are very similar and are thus omitted. (B) Mapping of relative deuterium protection levels of fsKaiB^{N29A} free and bound to CikA^{PSR} (orange); and CikA^{PSR} free and bound to fsKaiB^{N29A} (khaki). Protection levels that are 1.5 and 2.0 standard deviations above the average are in light blue and blue, respectively.

Figure 3.7. Mutagenesis of α5 in $^{\Delta N}$KaiA and β2 in fsKaiB^{L} to investigate their effects on the ternary complex formation by gel-filtration chromatography.

(Figure is on the next page)

(A) Gel-filtration profiles of ternary mixtures (top panel), $^{\Delta N}$KaiA mutants and control samples (bottom panel). Control samples include: $^{\Delta N}$KaiA (sky blue), H153A-$^{\Delta N}$KaiA (green), L155A-$^{\Delta N}$KaiA (pink), Q157A-$^{\Delta N}$KaiA (blue), K158A-$^{\Delta N}$KaiA (cyan), L159A-$^{\Delta N}$KaiA (black), E161A-$^{\Delta N}$KaiA (red), L163A-$^{\Delta N}$KaiA (coffee), fsKaiB (orange), CI*^{L} (dark green), $^{\Delta N}$KaiA + fsKaiB (purple), $^{\Delta N}$KaiA + CI*^{L} (dark green, dashed), and fsKaiB + CI*^{L} (orange, dashed). The label BC represents fsKaiB + CI*^{L} (top panel). Ternary samples include: $^{\Delta N}$KaiA + BC (sky blue, doubly dashed), H153A-$^{\Delta N}$KaiA + BC (green, doubly dashed), L155A-$^{\Delta N}$KaiA + BC (pink, doubly dashed), Q157A-$^{\Delta N}$KaiA + BC (blue, doubly dashed), K158A-$^{\Delta N}$KaiA + BC (cyan, doubly dashed), L159A-$^{\Delta N}$KaiA + BC (black, doubly dashed), E161A-$^{\Delta N}$KaiA + BC (red, doubly dashed), and L163A-$^{\Delta N}$KaiA + BC (coffee, doubly dashed). (B) Gel-filtration profiles of ternary mixtures (top panel), fsKaiB^{L} mutants and control samples (bottom panel). Control samples include: $^{\Delta N}$KaiA (sky blue), fsKaiB^{L} (orange), A41D-fsKaiB^{L} (black), K43E-fsKaiB^{L} (red), K43Q-fsKaiB^{L} (blue), CI*^{L} (dark green), $^{\Delta N}$KaiA + fsKaiB^{L} (purple), fsKaiB^{L} + CI*^{L} (orange, dashed), A41D-fsKaiB^{L} + CI*^{L} (black, dashed), K43E-fsKaiB^{L} + CI*^{L} (red, dashed), K43Q-fsKaiB^{L} + CI*^{L} (blue, dashed), and $^{\Delta N}$KaiA + CI*^{L} (dark green, dashed). The label AC represents $^{\Delta N}$KaiA + CI*^{L} (top panel). Ternary samples include: fsKaiB^{L} + AC (sky blue, doubly dashed), A41D-fsKaiB^{L} + AC (black, doubly dashed), E43E-fsKaiB^{L} + AC (red, doubly dashed), and K43Q-fsKaiB^{L} + AC (blue, doubly dashed). Elution positions of molecular weight markers are indicated by inverted triangles and molecular weights in kDa along the top of the gel-filtration chromatograms. For experimental details, please see Table 3.1.

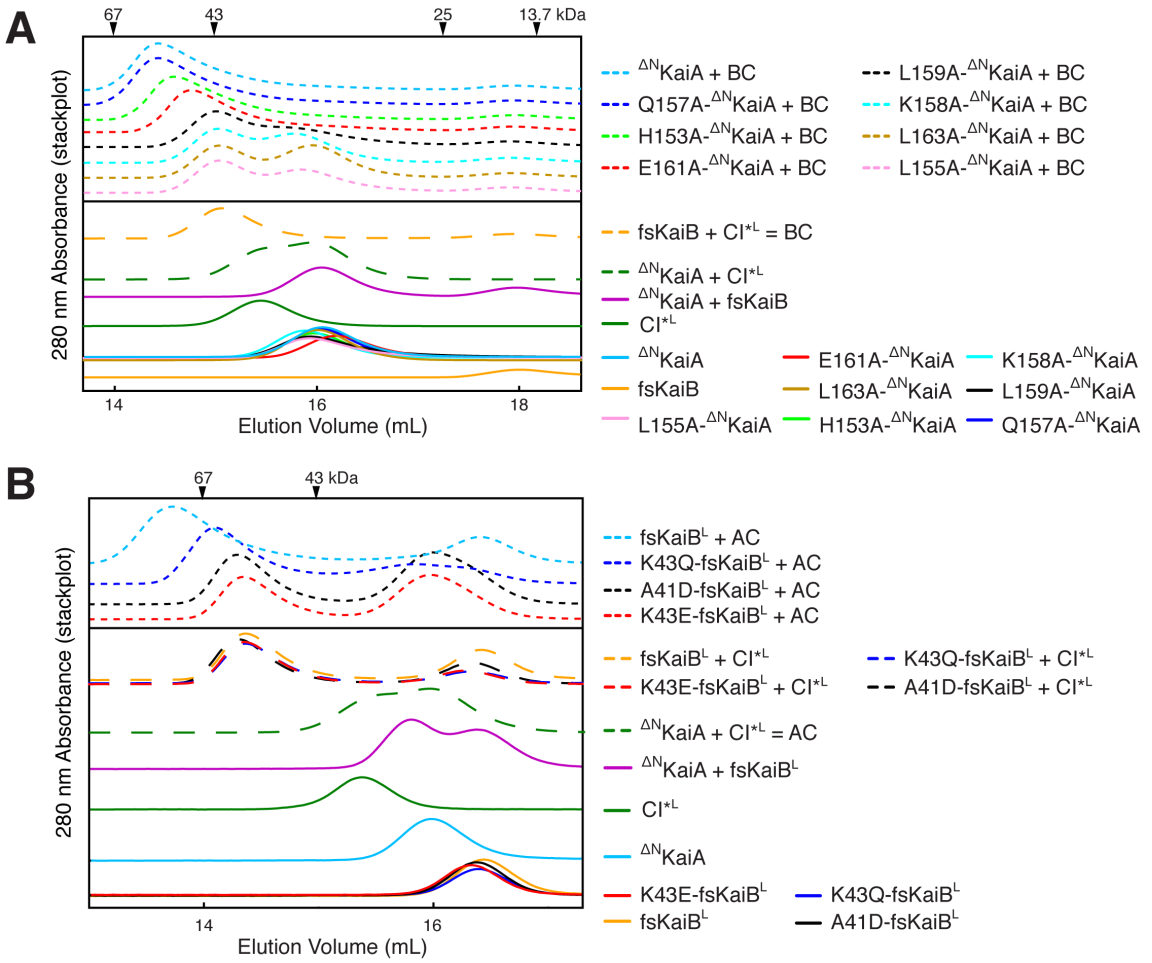


Figure 3.7. Mutagenesis of α5 in $^{\Delta N}$KaiA and β2 in fsKaiB^{L} to investigate their effects on the ternary complex formation by gel-filtration chromatography.

(Figure caption on the previous page)

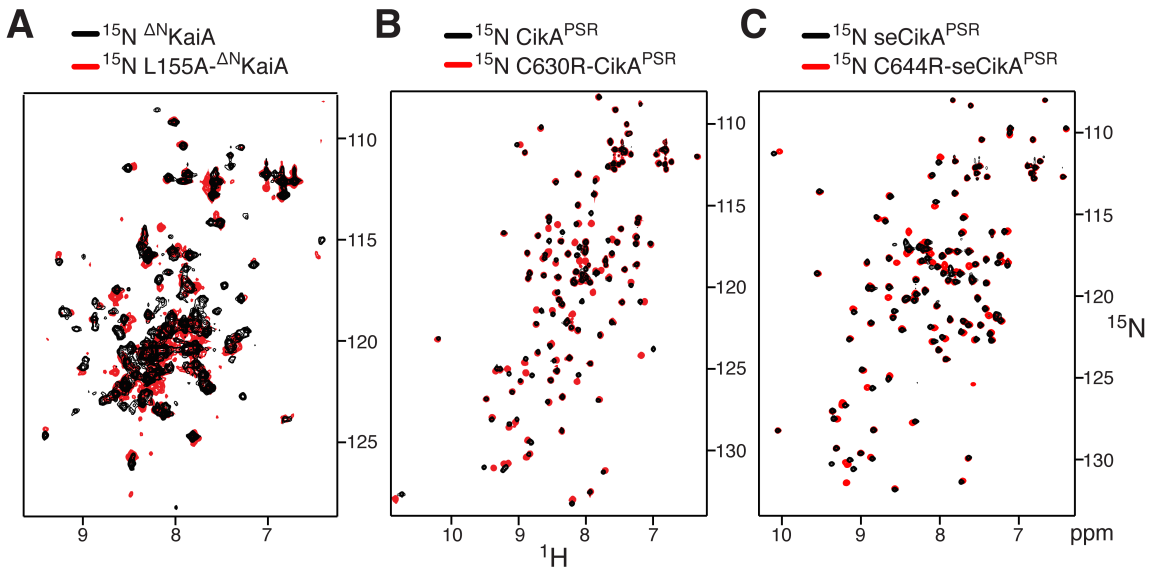


Figure 3.8. ^{15}N, ^{1}H-HSQC spectra of $^{\Delta N}$KaiA, CikA^{PSR}, seCikA^{PSR} and their variants.

(A) Overlay of ^{15}N, ^{1}H-HSQC spectra of U-[^{15}N, ^{1}H]-$^{\Delta N}$KaiA alone (black) and L155A-$^{\Delta N}$KaiA alone (red). (B) Overlay of ^{15}N, ^{1}H-HSQC spectra of U-[^{15}N, ^{1}H]-CikA^{PSR} alone (black) and C630R-CikA^{PSR} alone (red). (C) Overlay of ^{15}N, ^{1}H-HSQC spectra of U-[^{15}N, ^{1}H]-seCikA^{PSR} alone (black) and C644R-seCikA^{PSR} alone (red). C630R and C644R are homologous mutations between CikA^{PSR} and seCikA^{PSR}, respectively.

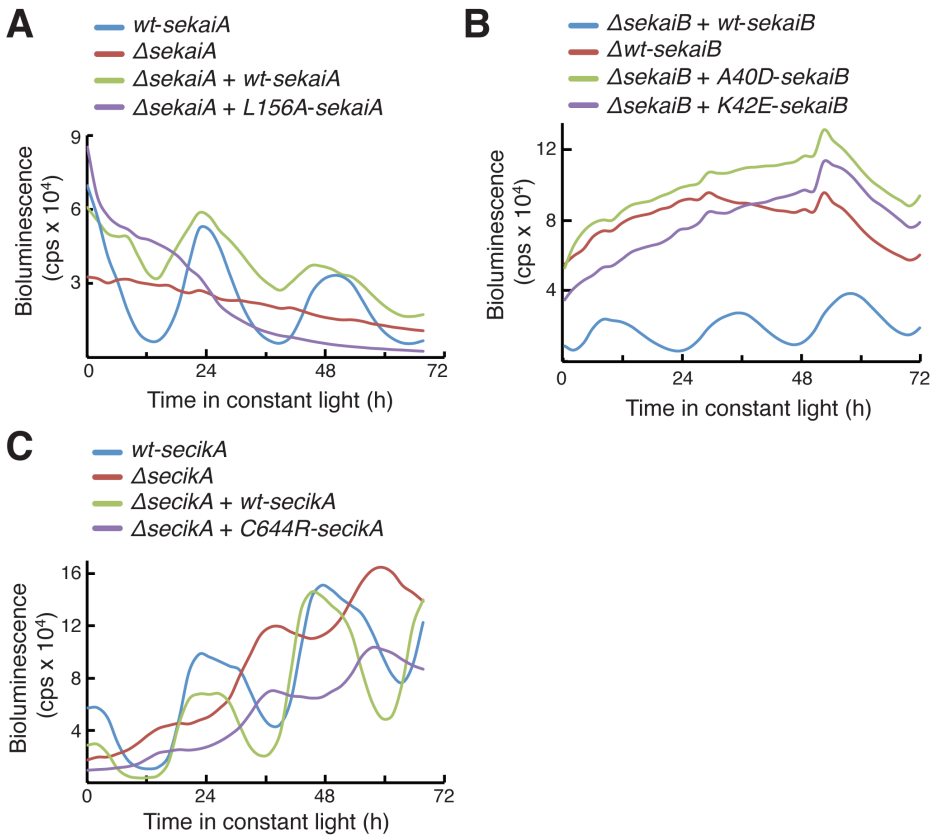


Figure 3.9. Effect of KaiA, KaiB, CikA mutations on in vivo bioluminescence rhythms.

(A) Bioluminescence rhythms of cyanobacterial strains harboring *wt-sekaiA* (blue), complemented with *wt-sekaiA* (green), or *kaiA* mutant *L156A-sekaiA* (purple), and a *seKaiA* knockout strain (brown). (B) Bioluminescence rhythms of strains harboring complemented *wt-sekaiB* (blue), or *kaiB* mutants *A40D-sekaiB* (green), *k42E-sekaiB* (purple), and a *seKaiB* knockout strain (brown). (C) Bioluminescence rhythms of strains harboring *wt-secikA* (blue), complemented with *wt-secikA* (green), or *cikA* mutants, *C644R-secikA* (purple), and a *seCikA* knockout strain (brown).

The CikA^{PSR}-fsKaiB^{N29A} complex reveals that fsKaiB uses the same β2 strand in binding $^{\Delta N}$KaiA and CikA^{PSR}.

CikA was discovered as a clock input component (35), in that it acts as a chemical sensor of redox metabolites such as quinones through its pseudo-receiver domain, CikA^{PSR}, which influences the period of the cyanobacterial clock (69, 71, 87). Deletion of this domain of CikA in cells resulted in circadian period shortening by 2-3 hours (69). Furthermore, overexpression of the CikA^{PSR} domain caused more severe period shortening by as long as 5 hours (69). Recently, CikA has also been found to play a role in clock nighttime output (31, 40), as its phosphatase activity is activated upon binding to the nighttime fsKaiB-KaiC complex (59). CikA decreases the phosphorylation level of the transcription factor RpaA, in contrast to the daytime output protein SasA, a kinase that phosphorylates RpaA (44). The CikA^{PSR} domain was identified to be key in mediating CikA-fsKaiB binding (59). Addition of this domain shortened the period of KaiC phosphorylation rhythms in vitro, similar to in vivo results (59, 69). Thus, we investigated the CikA^{PSR} interaction to the KaiB-CI*^{L} complex by NMR to understand molecular interactions regulating nighttime output signaling (Figure 3.10).

Assigned methyl-TROSY NMR (105) spectra of U-[^{15}N, ^{2}H]-Ile-δ1-[^{13}C, ^{1}H]-labeled CikA^{PSR} showed that binding of the expected nine isoleucine peaks occurred only upon mixing with wt-KaiB and CI*^{L} together, but not separately (Figure 3.10A & B). This suggests that CikA^{PSR} likely formed a ternary complex with the KaiB-CI*^{L} complex. CikA^{PSR} in a fsKaiB^{L} + CI*^{L} mixture showed similar chemical shifts as compared when mixed with wt-KaiB + CI*^{L} (Figure 3.10B & F). Interestingly, in a mixture with CikA^{PSR} , residues I616, I641, I657 and I708 showed two bound peaks (major and minor), indicating the existence of two different ternary conformations coexisting in the same sample (Figure 3.10B). However, when mixed with fsKaiB^{L} + CI*^{L}, the spectrum of CikA^{PSR} showed only one bound form, matching that of the minor population when mixed with wt-KaiB^{L} + CI*^{L} (Figure 3.10B). Chemical shift analysis showed that I641 was the most perturbed residue. This suggests that it likely resides near the binding interface (Figure 3.10F). fsKaiB^{L} interacted with CikA^{PSR} even in the absence of CI*^{L}, albeit weakly as both free and bound peaks were present (Figure 3.10C). These results again verify that the thioredoxin-like fsKaiB structure is the active nighttime form (Figure 3.4) (59). However, the presence of KaiC is also important as the CI domain stabilizes the fsKaiB state, allowing the formation of ternary KaiA-KaiB-KaiC and CikA-KaiB-KaiC interactions between wild-type proteins (Figure 3.4 & 3.10). Fortuitously, we identified a single N29A substitution on fsKaiB^{L} that allowed the formation of a stable complex with CikA^{PSR} in the absence of CI, as indicated by absence of detectable NMR peaks corresponding to unbound proteins (Figure 3.10D). Therefore, I made a more stable construct, fsKaiB^{N29A}.

The NMR spectrum of $CikA^{PSR}$ in the presence of $fsKaiB^{N29A}$ is highly similar to that of $CikA^{PSR}$ in the presence of $fsKaiB^{L}$, indicating very similar binding modes (Figure 3.10). Using these enhanced protein constructs, we solved the solution NMR structure of a $CikA^{PSR}$-$fsKaiB^{N29A}$ complex with 196 intermolecular NOEs and 154 RDC restraints using XPLOR-NIH (Figure 3.11 & Table 3.3) (131, 132).

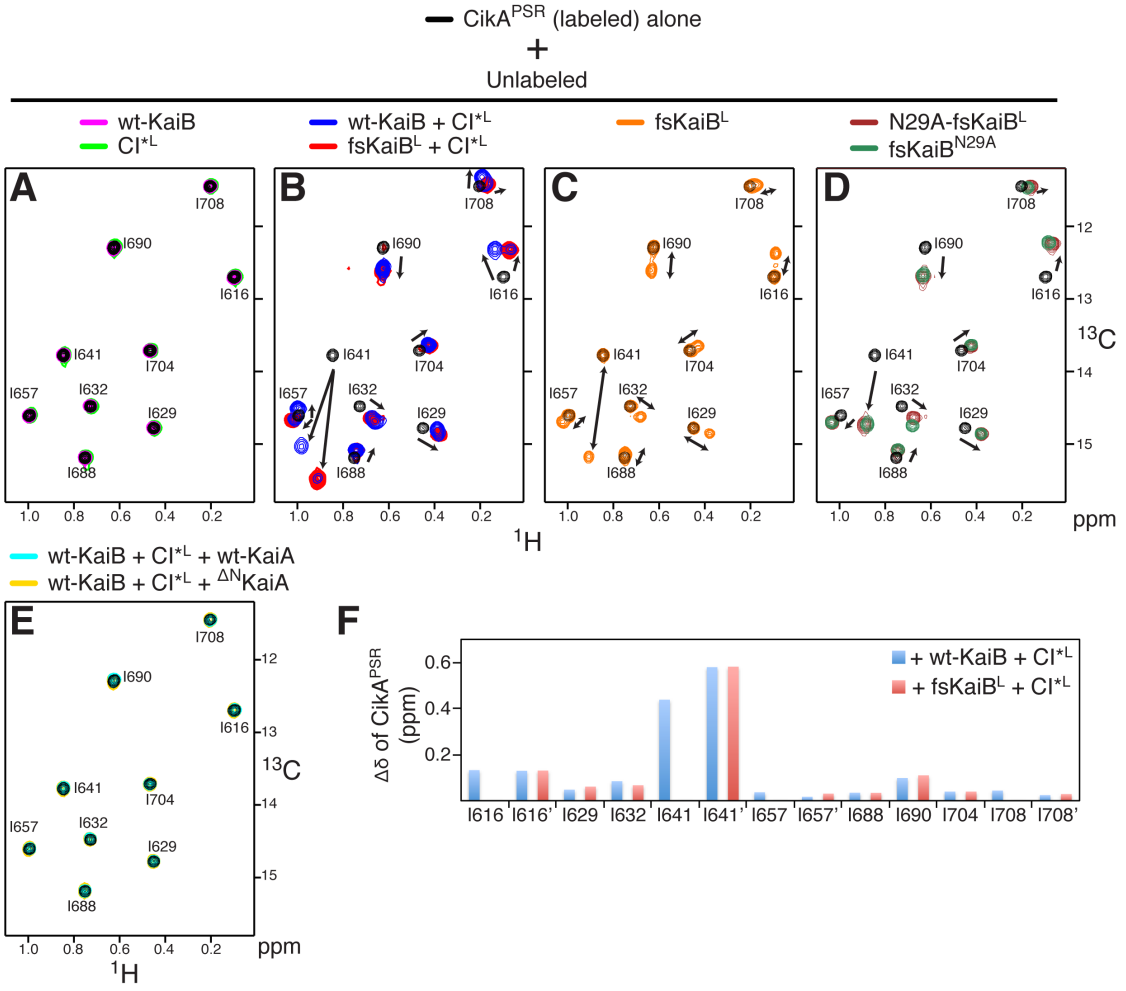


Figure 3.10. CikA^{PSR} forms ternary complex with KaiB and CI*^{L} near residue I641.

(A) Assigned methyl-TROSY spectra of U-[^{15}N, ^{2}H]-Ile-δ1-[^{13}C, ^{1}H]–labeled 1× CikA^{PSR} alone (black); or in the presence of 2.5× wt-KaiB (magenta); 2.5× CI*^{L} (green); or (B) in the presence of 2.5× wt-KaiB + 2.5× CI*^{L} (blue); 2.5× fsKaiB + 2.5× CI*^{L} (red); or (C) in the presence of 2.5× fsKaiB^{L} (orange); or (D) in the presence of 2.5× N29A-fsKaiB^{L} (brown); 2.5× fsKaiB^{N29A} (dark green); or (E) in the presence of 2.5× wt-KaiB + 2.5× CI*^{L} + 5.5x wt-KaiA (cyan); or 2.5× wt-KaiB + 2.5× CI*^{L} + 5.5x ^{ΔN}KaiA (yellow). (F) ^{1}H, ^{13}C combined chemical shift difference plot from (B) of CikA^{PSR} perturbed by 2.5× wt-KaiB + 2.5× CI*^{L} (blue) or 2.5× fsKaiB + 2.5× CI*^{L} (red). Apostrophes next to residue names indicate minor peak (e.g. I616'). 1× corresponds to 20 μM monomer concentration. For experimental details please see Table 3.1.

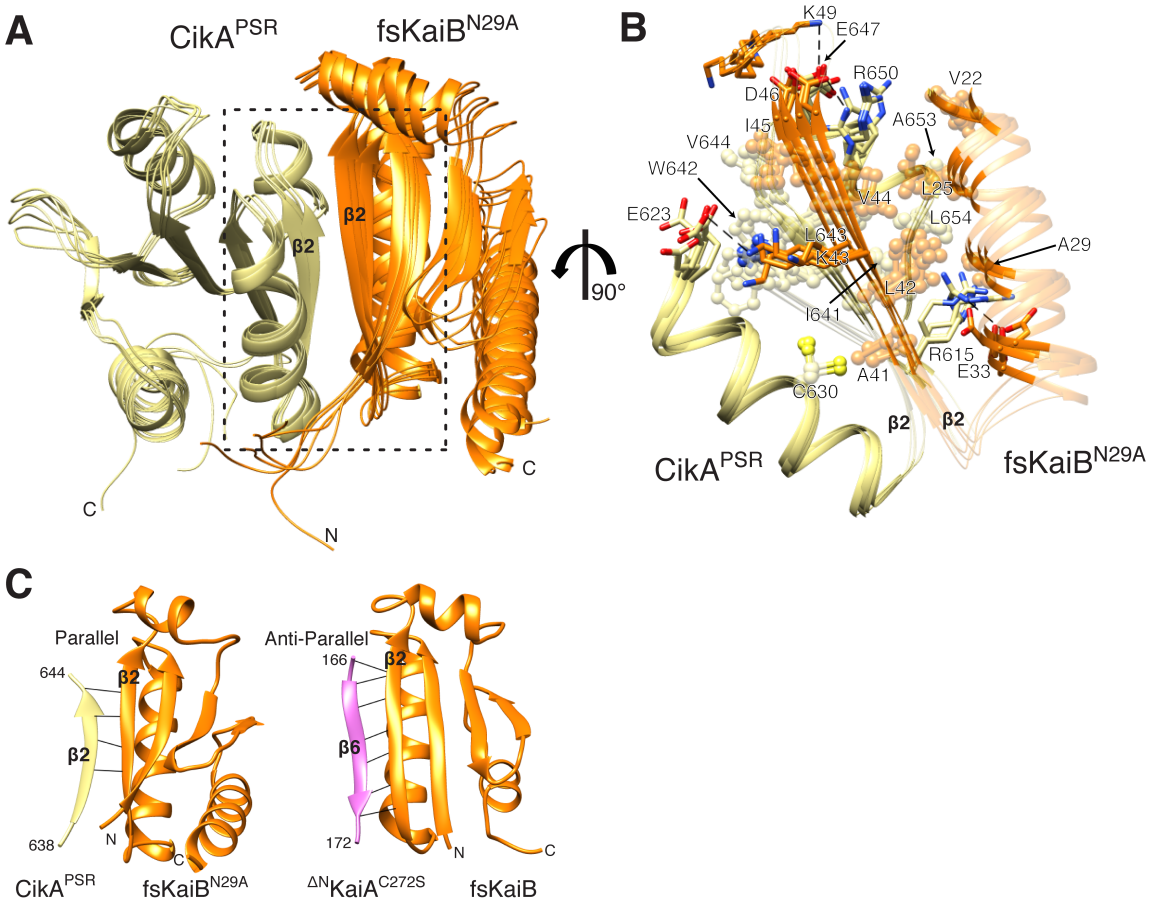


Figure 3.11. NMR structure of CikA^{PSR} in complex with fsKaiB^{N29A} reveals the structural basis for CikA and KaiA competition.

(A) Ensemble of the five lowest energy structures of the complex of CikA^{PSR} (khaki) and fsKaiB^{N29A} (orange), calculated by XPLOR-NIH (131). Flexible FLAG-tag tail of fsKaiB^{N29A} is not shown. (B) Zoomed-in interactions for the region boxed in (A). Hydrophobic and hydrophilic interactions are shown as balls and sticks, and sticks, respectively. (C) Simplified view showing interactions of β2 of fsKaiB^{N29A} with β2 of CikA^{PSR} (left) and β2 of fsKaiB with β6 of $^{\Delta N}$KaiA^{C272S} (right). Black lines are backbone-backbone hydrogen bonds.

The CikA^{PSR}-fsKaiB^{N29A} complex revealed a large binding interface between β2 of CikA^{PSR} and β2 of fsKaiB^{N29A} (Figure 3.11 & Table 3.4). There are also several polar interactions surrounding the β-sheet interface (Figure 3.11B). Residue I641 was at the center of the β2-β2 interface (Figure 3.10F & 3.11B). HDX-MS protection profile of the complex independently supported my NMR structure by showing strong protection against H/D exchange at the heterodimer interface (Figure 3.6B). The structure also showed that residue N29A, the affinity enhancing substitution on fsKaiB^{N29A}, was surrounded by hydrophobic residues I641 and L654 on CikA^{PSR} (Figure 3.11B). This suggests that the increased affinity of fsKaiB^{N29A} over fsKaiB for CikA^{PSR} is due to a hydrophobic effect since the asparagine is more polar than alanine. Interestingly, fsKaiB^{N29A} utilized the same residues, A41 and K43, to interact with C630 and E623 on CikA^{PSR}, respectively, as are used to make key contacts with $^{\Delta N}$KaiA. A side-by-side comparison of the parallel β2-β2 interface between CikA^{PSR} and fsKaiB^{N29A}, and the anti-parallel β6-β2 interface between $^{\Delta N}$KaiA and fsKaiB shows how fsKaiB can either sequester KaiA or engage the clock-output protein CikA, explaining how KaiA and CikA^{PSR} compete directly for KaiB (Figure 3.11C). Also, addition of wt-KaiA or $^{\Delta N}$KaiA into a sample of CikA^{PSR} + wt-KaiB + CI*^{L} restored the spectrum of free CikA^{PSR}, lending further support to the mutual exclusiveness of fsKaiB-KaiA and fsKaiB-CikA^{PSR} interactions. (Figure 3.10E). This competitive binding likely explains previous results that the presence of CikA^{PSR} perturbed circadian rhythms in vitro and in vivo (59, 69). My studies suggest that CikA^{PSR} liberated some amount of KaiA from KaiB, which could then stimulate KaiC (auto)phosphorylation. Indeed, CikA^{PSR} overexpression in vivo only has an effect during the subjective night, and not during the subjective day (69). This observation supports that notion that CikA^{PSR} only competes with KaiA during the night.

To further validate the structure of the CikA^{PSR}-fsKaiB^{N29A} complex, we carried out structure-guided mutagenesis studies of CikA^{PSR} and fsKaiB^{N29A}. We found that A41D/K43E substitutions in fsKaiB^{N29A} strongly affected binding by gel-filtration chromatography (Figure 3.12A). A41D/K43E were also the same sites that previously disrupted fsKaiB binding to $^{\Delta N}$KaiA (Figure 3.7B). Thus, this result further supports the structural finding that CikA^{PSR} and KaiA interact with the same β2 strand on KaiB. Because C630 on CikA^{PSR} was found to be buried deeply in the intermolecular hydrophobic interface and interacted with A41 on fsKaiB^{N29A}, we substituted it with a bulky and charged argininyl residue. This C630R amino acyl substitution was very effective in eliminating complex formation (Figure 3.12B). The ^{15}N, ^{1}H-HSQC spectrum of C630R-CikA^{PSR} was very similar to that of CikA^{PSR}, suggesting that the C630R substitution did not significantly affect the structural integrity of CikA^{PSR} (Figure 3.8B). Likewise, the corresponding substitution in seCikA^{PSR}, C644R, also produced minimal spectral perturbations. (Figure 3.8C). Thus, we tested the functional effect of C644R in

full-length seCikA on the circadian gene expression rhythm (Figure 3.9C). Cyanobacterial strains that harbored the C644R mutation in seCikA exhibited defective circadian rhythms that were similar to *secikA* knockout strains (Figure 3.9C). This functional assay validates that $CikA^{PSR}$ and KaiB interact across the β2-β2 interface.

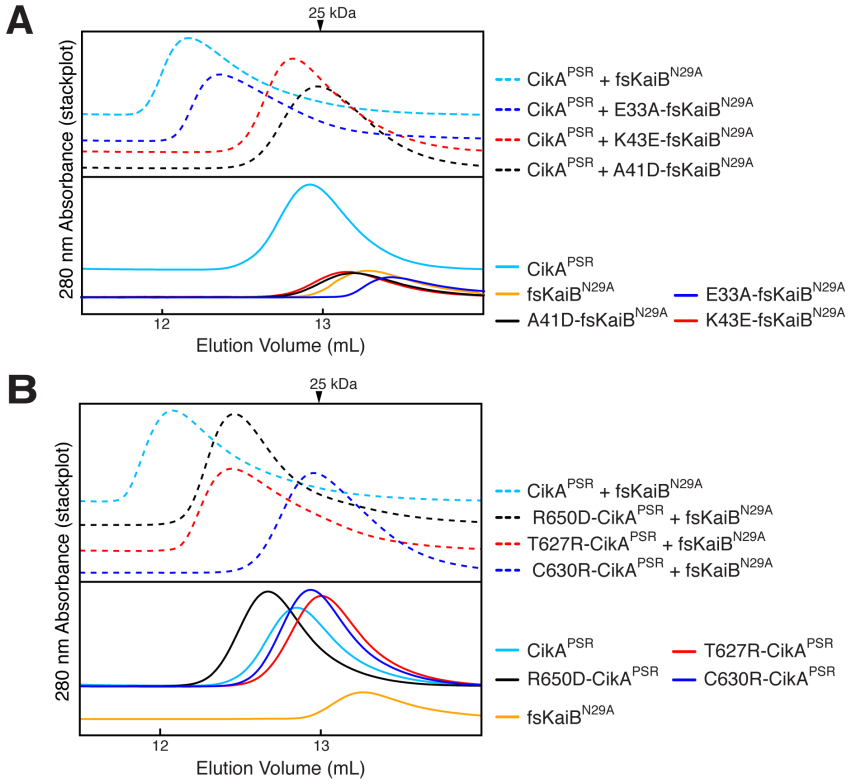


Figure 3.12. Mutagenesis of fsKaiB^{N29A} and CikA^{PSR} to interrogate complex formation by gel-filtration chromatography.

(A) Gel-filtration profiles of binary mixtures (top panel), and control samples (bottom panel). Control samples include: CikA^{PSR} (sky blue), fsKaiB^{N29A} (orange), E33A-fsKaiB^{N29A} (blue), A41D-fsKaiB^{N29A} (black), and K43E-fsKaiB^{N29A} (red). The binary samples include: CikA^{PSR} + fsKaiB^{N29A} (sky blue, dashed), CikA^{PSR} + E33A-fsKaiB^{N29A} (blue, dashed), CikA^{PSR} + A41D-fsKaiB^{N29A} (black, dashed), and CikA^{PSR} + K43E-fsKaiB^{N29A} (red, dashed). (B) Gel-filtration profiles of binary mixtures (top panel), and control samples (bottom panel). Control samples include: CikA^{PSR} (sky blue), T627R-CikA^{PSR} (red), C630R-CikA^{PSR} (blue), R650D-CikA^{PSR} (black) and fsKaiB^{N29A} (orange). The binary samples include: CikA^{PSR} + fsKaiB^{N29A} (sky blue, dashed), T627R-CikA^{PSR} + fsKaiB^{N29A} (red, dashed), C630R-CikA^{PSR} + fsKaiB^{N29A} (blue, dashed), and R650D-CikA^{PSR} + fsKaiB^{N29A} (black, dashed). Elution positions of molecular weight markers are indicated by inverted triangles. For experimental details, please see Table 3.1.

Discussion

The three high-resolution protein complexes determined in this study have revealed several important mechanisms that are necessary for day-to-night transitions by the cyanobacterial circadian oscillator and output signaling of the cyanobacterial clock (Figure 3.13). To initiate the transition, the binary fsKaiB-CI* complex showed that it is the thioredoxin fold of KaiB that interacts with the post-ATP hydrolysis state of the CI domain of KaiC. Secondly, the ternary $^{\Delta N}$KaiA^{C272S}-fsKaiB-CI* complex revealed that a single KaiB monomer inactivates both KaiC-binding sites of KaiA by stabilizing a conformation in which an α-helix from each protomer of the KaiA dimer physically block the KaiC-binding sites. Finally, I showed that the mechanism by which KaiB activates CikA is through forming a stable complex with the *pseudo*-receiver domain of CikA. Furthermore, I proposed that the mutually exclusive binding by KaiA and CikA^{PSR} to KaiB is likely the molecular basis of circadian period modulation by CikA^{PSR}.

CI domain of KaiC stabilizes fsKaiB to facility cooperative formation of the ternary $^{\Delta N}$KaiA-fsKaiB-CI and CikA^{PSR}-fsKaiB-CI complexes, even when KaiB is already in thioredoxin fold (Figure 3.4 & 3.10). Interestingly, $^{\Delta N}$KaiA and CikA^{PSR} both bind to the part of KaiB that does not change local structure between the two different folds of KaiB. Superimposing only the conformation of the β2 strand of ground-state KaiB (gsKaiB) (59, 107) onto the structures of the $^{\Delta N}$KaiA^{C272S}-fsKaiB-CI* and CikA^{PSR}-fsKaiB^{N29A} complexes showed that E56 and D57 on gsKaiB are 1.9 and 3.7 Å from E213 on $^{\Delta N}$KaiA^{C272S} and E623 on CikA^{PSR} (Figure 3.14D & F) but are 6.4 and 6.5 Å away in fsKaiB (Figure 3.14C & E). Additionally, E95 and E96 in gsKaiB physically clash with α7 and α1 helices of $^{\Delta N}$KaiA^{C272S} and CikA^{PSR}, respectively (Figure 3.14D & F). Thus, it can be concluded that the structure of gsKaiB is not compatible with binding KaiA or CikA.

In vitro reactions with $^{\Delta N}$KaiA showed that KaiC (auto)dephosphorylation occurred sooner that than with full-length KaiA (60), suggesting that the N-terminal domain of KaiA prevents premature autoinhibition by the α5 helix during the subjective day. Indeed, NMR spectra of $^{\Delta N}$KaiA suggest a stable well-folded structure (Figure 3.8A). However, an L155A-$^{\Delta N}$KaiA variant had line-broadened spectra, indicative of large-scale conformational fluctuations (Figure 3.8A). These spectra suggest that in $^{\Delta N}$KaiA the α5 helix is stably packed against the main body of $^{\Delta N}$KaiA, which in the absence of the N-terminal domain is most likely the C-terminal domain. The N-terminal domain of KaiA is not as conserved as the C-terminal domain, which is functionally important in binding to A loop. Perhaps not surprisingly, strand β6 of KaiA which binds KaiB is highly conserved (Figure 3.3). Interestingly, D52 in se^{N}KaiA is also a highly conserved residue that is shown to make the potential electrostatic interaction to R163 in α5 of seKaiA (Figure 3.5C).

To separate day and nighttime events, the cyanobacterial clock utilizes post-translational modifications such as phosphorylation, that is the hallmark of clock regulation in diverse organisms (1). Surprisingly, the day/night regulation involves a metamorphic protein, KaiB, which uses two district folds to control binding events temporally (59). Upon switching to the thioredoxin fold and binding Ser431-phosphorylated KaiC, KaiB acts as a hub to regulate a series of events to initiate day/night transition in the oscillator and output signaling. KaiB disassembles SasA (44, 60, 67), induces KaiA self-inhibition and activates the CikA phosphatase (40, 59). KaiB forms a signaling hotspot on the CI domain of KaiC. ATP hydrolysis by CI and rigidity induced CII-CI ring interactions likely play critical roles in controlling KaiB binding to CI. In the mammalian clock, the phenomenon of clock proteins competing for binding to a few crucial signaling hotspots was also observed (141). Thus, the competition mechanisms between SasA/KaiB and KaiA/CikA found here also illustrate common strategies adapted by other clock system.

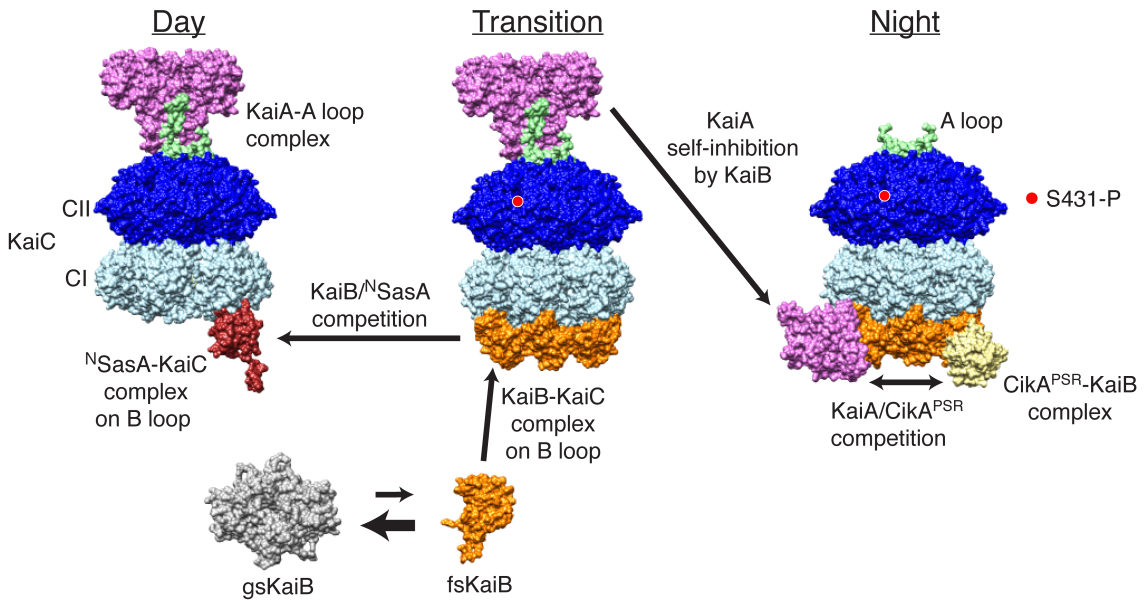


Figure 3.13. Day to night clock protein complex remodeling in cyanobacteria.

Concluding from structural investigations in this study, three complexes: fsKaiB-CI*, $^{\Delta N}$KaiA^{C272S}-fsKaiB-CI*, and CikA^{PSR}-fsKaiB^{N29A} were overlaid to all current structures of KaiC, KaiB, KaiA-A loop and ^{N}SasA-CI complexes to illustrate the summary of day to night transition of the clock proteins. The oscillator and output events during the subjective day are: 1) KaiA stimulates of KaiC (auto)phosphorylation on S431 and T432 residues by binding to A loop, and 2) KaiC stimulated SasA kinase activity on RpaA by ^{N}SasA and CI complex on B loop of KaiC. The oscillator and output events transitioning throughout the subjective night are: 1) the active fsKaiB transformed from the inactive ground-state fold of KaiB (gsKaiB) to bind to B loop of KaiC upon CII S431 phosphorylation (S431-P), and initiates nighttime complex disassembly of ^{N}SasA and assembly with KaiA and CikA^{PSR}, 2) fsKaiB induces large conformational change on KaiA and blocks A loop binding site via KaiA self-inhibition, thereby allowing KaiC (auto)dephosphorylation, and 3) CikA^{PSR} forms ternary complex with KaiB and KaiC as the basis for KaiB-KaiC stimulated CikA phosphatase activity on RpaA. This interaction antagonizes SasA's effect and possibly also modulate circadian period through competition with KaiA, as KaiA and CikA^{PSR} bind mutually exclusively on KaiB at night.

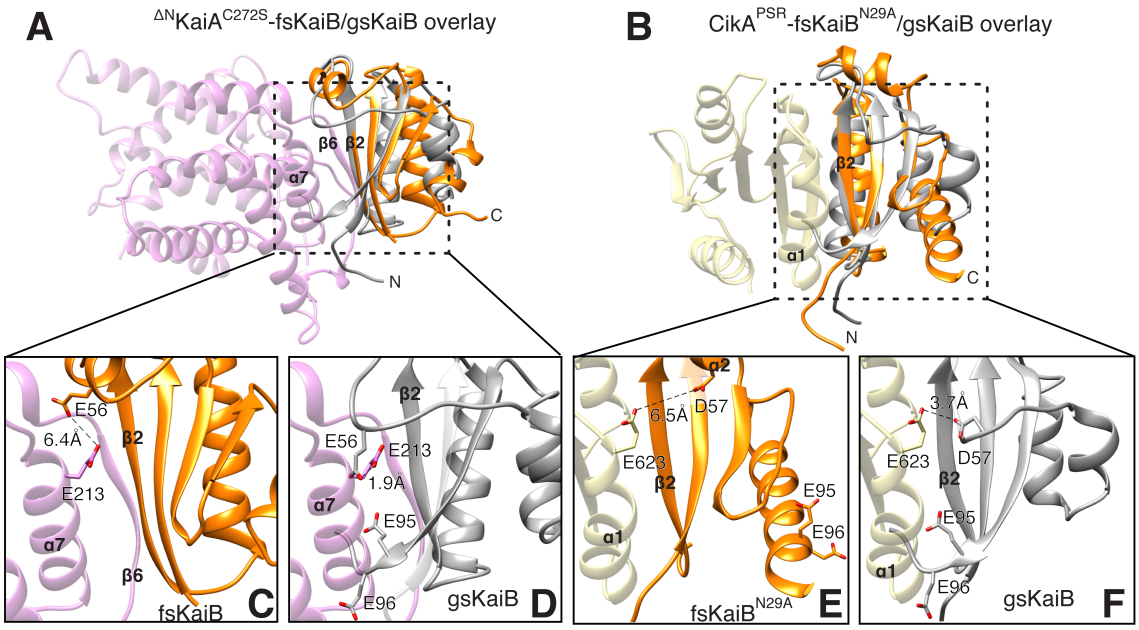


Figure 3.14. Ground-state fold of KaiB repels $^{\Delta N}$KaiA and CikA^{PSR} with acidic residues.

(A) Superposition of ground-state fold of KaiB (gray) to the ternary $^{\Delta N}$KaiA^{C272S}-fsKaiB-CI* structure. CI* is not shown. (B) Overlay of ground-state fold of KaiB (gsKaiB) (gray) to the CikA^{PSR}-fsKaiB^{N29A} structure. (C) and (D) are zoomed-in regions from the region boxed in (A). (E) and (F) are detailed zoomed-in interactions for the region boxed in (B). PDB ID for gsKaiB is 2QKE.

CHAPTER IV

CONCLUSIONS

The study set out to address how endogenous circadian clocks of diverse organisms generate ~24-h rhythms to match their physiology and behavior with daily environmental changes such as light and temperature. The cyanobacterial circadian clock has been chosen for detailed mechanistic investigation in that it is considerably less complex in terms of the number of components and regulatory circuitry, and notably, can be reconstituted in vitro. As the assembly of daytime and nighttime complexes is key to understanding a 24-h cycle, extensive studies have been carried out to elucidate how cyanobacterial clock components achieve temporal assembly and clock output. Although daytime complex assembly in the cyanobacterial clock and daytime clock output has been well examined, little was known about how a nighttime component, KaiB, through fold-switching, serves as linchpin of nighttime complex assembly and clock output. Thus, this study sought to answer two specific questions:

1. What role does the daytime component, KaiA, play in the assembly of the nighttime complex, KaiA-KaiB-KaiC?

2. How does the nighttime component, KaiB, inhibit KaiA and engage clock output at night?

The results of the studies addressing each question are presented in Chapters II (*Cooperative KaiA-KaiB-KaiC Interactions Affect KaiB/SasA Competition in the Circadian Clock Of Cyanobacteria*) and III (Crystal and NMR Structures Of The Nighttime Complexes), respectively. These results led to the following conclusions.

Conclusions regarding Question 1:

a. *Day-to-night transition involves long-range signal transduction.*

   The first step in testing the potential role of KaiA in regulating KaiB-KaiC interactions is to understand how KaiB senses the day-to-night transition signal on KaiC, i.e. phosphorylation at Ser431 in the CII domain of KaiC. The study identified the binding site of KaiB to be the CI domain of KaiC, contrary to an opposing model in which KaiB binds to CII. Although the two domains are homologous, the CI domain has an insertion relative to

CII. It is this insertion that constitutes a key binding element for KaiB, and is thus termed B loop (KaiB-binding loop). This finding that day-to-night signal transmission is long range, (> 60 Å) from CII to CI, explains our observation that CII-CI ring stacking is essential for the onset of the night phase, i.e., KaiB binding.

b. *The daytime component, KaiA, plays a cooperative role in nighttime complex formation and also promotes dissociation of the daytime output complex.*

It has been a general view that KaiA functions during the day, and at night is inhibited by being sequestered by the KaiB-KaiC complex. However, this study found that KaiB-KaiC is insufficient in disassembling the daytime output complex SasA-KaiC, raising the question how daytime output is terminated as the clock transitions from day to night. The finding that KaiA enhances KaiB-KaiC association reveals an active role in regulating day-to-night transition and clock output.

Conclusions regarding Question 2:

a. *KaiB is the first metamorphic protein known to play a key role in biological timekeeping*.

KaiB has been known to regulate clock output by inactivating the daytime clock-output protein SasA (31, 40). The basis for this inactivation is that KaiB blocks SasA binding via ^{N}SasA to KaiC. Despite having high sequence similarity, free KaiB and ^{N}SasA adopt distinctly different folds, with the former having a rare fold and the latter adopting the thioredoxin fold. This study revealed that KaiB can also adopt a thioredoxin-like fold, which is stabilized when bound to KaiC. This finding that KaiB can switch reversibly between two distinct folds reveals the structural basis of how KaiB regulates SasA-mediated output. Furthermore, KaiB, along with other metamorphic proteins, serves as evidence that invokes rethinking about the dogma “one sequence, one structure” of proteins.

b. *Metamorphosis of a single protein pivots day-to-night transitions in a clock*.

This study solved the structures of three KaiB-based nighttime complexes: KaiB-KaiC, KaiA-KaiB-KaiC, and KaiB-CikA. These structures show that KaiB adopts the thioredoxin fold in disassociating daytime complexes KaiA-KaiC and SasA-KaiC, and integrating nighttime entities for the night phase. As KaiA and CikA binding to KaiB are mutually

exclusive, KaiB is a hub, facilitating the interplay between clock rhythm generation and clock output.

This study offers insights into the structural basis of how KaiB engages KaiC, inactivates KaiA and SasA, and activates CikA. We used the divide-and-conquer strategy, which studies interactions between isolated protein domains. This methodology proved successful for achieving the goal. However, it did not capture the state of KaiB binding to the hexameric full-length KaiC particle. In this state, six KaiBs have been shown to bind cooperatively and this binding is regulated by the CI ATPase activity (56, 61). Thus, future research needs to be directed at solving the structure of KaiB bound to full-length KaiC, which would be expected to elucidate how ATPase activity of CI adds a layer of regulation in day-to-night transition.

My work filled in a major gap in the structural biology of the cyanobacterial clock by determining the structures of all the nighttime complexes, which center on a single metamorphic protein. In addition, it uncovered an active role of a daytime component in night-phase generation. The knowledge gained helped create a new framework in which clock components can assume dynamic roles during the day and at night, and long-range allosteric effects and large conformational changes underpin the transition from day to night. This framework is expected to influence research on eukaryotic circadian clocks including humans, which could eventually help improve health quality.

Other important implications of this work on future research are listed below.

1. What is the mechanism of CII-CI ring communication? Although we know that the CII and CI rings stack upon phosphorylation of residue S431, we do not know the mechanism by which this happens. One hypothesis is that the ATPase activity of the CI ring periodically drives the CI ring into an excited state, where it checks the status of the CII ring. When the CII ring is unphosphorylated at residue S431, it is loose and breathing (52), and not receptive to the excited state of CI. However, upon phosphorylation of S431, the CII ring tightens and becomes selectively receptive to the excited state of CI. The excited CI ring captures the CII ring, and CI-CII binding energy prevents CI from relaxing back to the ground state.

2. What happens to the CI ring upon CII-CI ring communication? We know that in the absence of CII-CI ring communication, KaiB is unable to bind to the CI ring. Upon ring stacking, KaiB easily binds to the CI ring. When KaiB is added to isolated CI rings, it disrupts the ring and binds to individual monomeric domains (48). Thus, we proposed that CI-CII ring stacking opens or pries apart the subunits of the CI ring. However,

modeling the crystal structure of the KaiB-CI complex onto that of full length KaiC hypothesizes that ring stacking may elongate the CI ring to provide access to KaiB (Figure 3.13).

3. What determines how the A loop C-terminal tail of KaiC binds KaiA? A recent crystal structure showed that a peptide representing the C-terminal tail of KaiC binds to the C-terminal domain of full-length KaiA as an α-helix, which is similar to how the α5 helix of KaiA binds there (51). However, an NMR structure showed the C-terminal tail of KaiC binding in an extended conformation to a KaiA construct consisting of only the C-terminal domain (residues 180-283) (109). The discrepancy between the conformations of the KaiC peptide bound to two different forms of KaiA suggests that the α5 helix and N-terminal domain of KaiA regulate KaiA-KaiC interactions. Here, we would resolve the discrepancy between NMR and crystal structures of the KaiA-KaiC peptide complex. The importance of this interaction demands that we figure out the reason behind these two different modes of binding. Future experiments should be designed that report on the conformational states of the KaiC A loop region peptide, and manipulate the conditions of binding to identify the determining factors.

4. KaiC interact with ATP and ADP. What are the energetics of ATP & ADP binding by different phosphorylation states of KaiC and isolated CII and CI rings? Deeper understanding of KaiC function could be obtained by measuring heats of ATP and ADP binding by the following KaiC constructs:

    - wild-type KaiC
    - E318Q, E319Q KaiC (no CII ATPase activity)
    - E77Q, E78Q KaiC (no CI ATPase activity)
    - E77Q, E78Q, E318Q, E319Q KaiC (no CI & CII ATPase activity)
    - E318Q, E319Q, S431A, T432A KaiC (no CII ATPase activity for dephosphorylated mimic)
    - E318Q, E319Q, S431A, T432E KaiC (no CII ATPase activity for phospho-T432 mimic)
    - E318Q, E319Q, S431E, T432E KaiC (no CII ATPase activity for phospho-S431,phospho-T432 mimic)
    - E318Q, E319Q, S431E, T432A KaiC (no CII ATPase activity for phospho-S431 mimic)
    - E77Q, E78Q, S431A, T432A KaiC (no CI ATPase activity for dephosphorylated mimic)
    - E77Q, E78Q, S431A, T432E KaiC (no CI ATPase activity for phospho-T432 mimic)

- E77Q, E78Q, S431E, T432E KaiC (no CI ATPase activity for phospho-S431,phospho-T432 mimic)
- E77Q, E78Q, S431E, T432A KaiC (no CI ATPase activity for phospho-S431 mimic)

5. How do the N- & C-terminal domains of KaiA regulate the position of the α5 helix of KaiA, and thereby fine-tune KaiA-KaiC interactions? We found that the α5 helix of KaiA has two major conformations, one of which permits KaiC to bind to KaiA and one that prohibits such binding. The hypothesis is that the permissive state is stabilized by interactions between α5 and the N-terminal domain of KaiA, and the prohibitive state is stabilized by interactions between α5 and the C-terminal domain of KaiA. In other words, the N- and C-terminal domains of KaiA compete with each other for binding α5, and that this competition tunes the clock. To test this hypothesis, future experiments would be designed that would report on the conformational and dynamical states of α5 under a variety of conditions and correlate those states with KaiA-KaiC interactions.

6. How does KaiA binding to the A loop of KaiC negatively regulate SasA and KaiB binding to the CI domain of KaiC? I recently showed that adding KaiA to a complex between ET-KaiC and SasA causes the displacement of SasA from KaiC, suggesting exquisitely sensitive coupling between pairwise protein-protein interactions are transduced across the opposing rings of KaiC, and this coupling is essential to segregating daytime interactions from nighttime interactions.

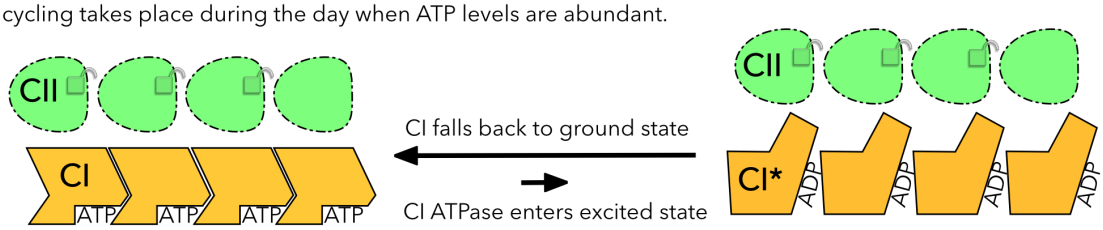
7. Unfortunately, my work has not led to additional insights into the mechanism of temperature compensation, a defining criterion of all biological circadian clocks. This property is perhaps the most poorly understood in circadian clocks, and future experiment should address this highly prized and mysterious mechanism in the cyanobacterial system.

Last but not least, an intriguing hypothesis regarding the function of CI ATPase activity is illustrated on the next page.

NEW HYPOTHESIS ON CI ATPASE ACTIVITY

ST- & SpT- phosphoforms of KaiC

The CI ring samples the state of the CII ring by entering an ATPase-dependent excited state. Because the CII ring is unreceptive, the CI ring falls back to the ground state. CI ATPase activity is high. CI ATPase cycling takes place during the day when ATP levels are abundant.



pSpT- & pST- phosphoforms of KaiC

The CII ring is selectively receptive to the excited state of the CI ring. Binding energy between the two rings prevents the CI ring from leaving the excited state. CI ATPase activity is low. This state of KaiC is stable at night when ATP levels are low.

Testable predictions of hypothesis:

1. Isolated CI rings have higher ATPase activity than in the presence of rigid CII rings, like E318Q+E319Q+S431E-CII.

2. Isolated CI rings shuffle subunits, but isolated E78Q+E79Q CI rings do not (or do so very slowly).

3. E78Q+E79Q CI rings have a higher binding energy to ATP than do wt-CI.

4. Under high levels of ATP, isolated CI rings stack onto rigid CII rings, but E78Q+E79Q CI rings do not.

5. Under sufficiently high levels of ADP, E78Q+E79Q CI rings will stack onto rigid CII rings.

REFERENCES

1. Albrecht U (2009) The circadian clock

2. de Mairan J (1729) Observation botanique. *In Histoire De L'Academie Royale Des Sciences* : 35-36.

3. Dodd AN, *et al* (2005) Plant circadian clocks increase photosynthesis, growth, survival, and competitive advantage. *Science* 309(5734): 630-633.

4. Ouyang Y, Andersson CR, Kondo T, Golden SS & Johnson CH (1998) Resonating circadian clocks enhance fitness in cyanobacteria. *Proc Natl Acad Sci U S A* 95(15): 8660-8664.

5. Orozco-Solis R & Sassone-Corsi P (2014) Circadian clock: Linking epigenetics to aging. *Curr Opin Genet Dev* 26C: 66-72.

6. Bass J & Takahashi JS (2010) Circadian integration of metabolism and energetics. *Science* 330(6009): 1349-1354.

7. Masri S, Cervantes M & Sassone-Corsi P (2013) The circadian clock and cell cycle: Interconnected biological circuits. *Curr Opin Cell Biol* 25(6): 730-734.

8. Fisher SP, Foster RG & Peirson SN (2013) The circadian control of sleep. *Handb Exp Pharmacol* (217):157-83. doi(217): 157-183.

9. Musiek ES, *et al* (2013) Circadian clock proteins regulate neuronal redox homeostasis and neurodegeneration. *J Clin Invest* 123(12): 5389-5400.

10. Sahar S & Sassone-Corsi P (2009) Metabolism and cancer: The circadian clock connection. *Nat Rev Cancer* 9(12): 886-896.

11. Dunlap JC (1999) Molecular basis for circadian clocks. *Cell* 96: 271-290.

12. Vernberg, F. John,, Vernberg,Winona B.,, (1983) *Behavior and ecology,* (Academic Press, New York),

13. Reyes BA, Pendergast JS & Yamazaki S (2008) Mammalian peripheral circadian oscillators are temperature compensated. *J Biol Rhythms* 23(1): 95-98.

14. Sweeney BM & Hastings JW (1960) Effects of temperature upon diurnal rhythms. *Cold Spring Harb Symp Quant Biol* 25: 87-104.

15. Bunning E (1967) *The Physiological Clock : Circadian Rhythms and Biological Chronometry,* (E.U.P.,

16. Lee C, Etchegaray JP, Cagampang FR, Loudon AS & Reppert SM (2001) Posttranslational mechanisms regulate the mammalian circadian clock. *Cell* 107: 855-867.

17. Lipton JO, *et al* (2015) The circadian protein BMAL1 regulates translation in response to S6K1-mediated phosphorylation. *Cell* 161(5): 1138-1151.

18. Nishiwaki T, *et al* (2004) Role of KaiC phosphorylation in the circadian clock system of synechococcus elongatus PCC 7942. *Proc Natl Acad Sci U S A* 101(38): 13927-13932.

19. Hirayama J, *et al* (2007) CLOCK-mediated acetylation of BMAL1 controls circadian function. *Nature* 450(7172): 1086-1090.

20. Chiu JC, Vanselow JT, Kramer A & Edery I (2008) The phospho-occupancy of an atypical SLIMB-binding site on PERIOD that is phosphorylated by DOUBLETIME controls the pace of the clock. *Genes Dev* 22(13): 1758-1772.

21. Wijnen H & Young MW (2006) Interplay of circadian clocks and metabolic rhythms. *Annu Rev Genet* 40(1): 409-448.

22. Nakajima M, *et al* (2005) Reconstitution of circadian oscillation of cyanobacterial KaiC phosphorylation in vitro. *Science* 308(5720): 414-415.

23. Johnson CH & Egli M (2014) Metabolic compensation and circadian resilience in prokaryotic cyanobacteria. *Annu Rev Biochem* 83: 221-247.

24. Holtman CK, *et al* (2005) High-throughput functional analysis of the synechococcus elongatus PCC 7942 genome. *DNA Res* 12(2): 103-115.

25. Hwang I, Chen H & Sheen J (2002) Two-component signal transduction pathways in arabidopsis. *Plant Physiol* 129(2): 500-515.

26. Williams SB, Vakonakis I, Golden SS & LiWang A (2002) Structure and function from the circadian clock protein KaiA of synechococcus elongatus: A potential clock input mechanism. *Proc Natl Acad Sci USA* 99: 15357-15362.

27. Grobbelaar N, Huang TC, Lin HY & Chow TJ (1986) Dinitrogen-fixing endogenous rhythm in synechococcus RF-1. *FEMS Microbiol Lett* 37(2): 173-177.

28. KIPPERT F (1987) Endocytobiotic coordination, intracellular calcium signaling, and the origin of endogenous rhythms. *Ann N Y Acad Sci* 503(1): 476-495.

29. Mitsui A, *et al* (1986) Strategy by which nitrogen-fixing unicellular cyanobacteria grow photoautotrophically. *Nature* 323(6090): 720-722.

30. Ishiura M, *et al* (1998) Expression of a gene cluster kaiABC as a circadian feedback process in cyanobacteria. *Science* 281(5382): 1519-1523.

31. Markson JS, Piechura JR, Puszynska AM & O'Shea EK (2013) Circadian control of global gene expression by the cyanobacterial master regulator RpaA. *Cell* 155(6): 1396-1408.

32. Hanaoka M, *et al* (2012) RpaB, another response regulator operating circadian clock-dependent transcriptional regulation in synechococcus elongatus PCC 7942. *J Biol Chem* 287(31): 26321-26327.

33. Iwasaki H, *et al* (2000) A kaiC-interacting sensory histidine kinase, SasA, necessary to sustain robust circadian oscillation in cyanobacteria. *Cell* 101(2): 223-233.

34. Taniguchi Y, Takai N, Katayama M, Kondo T & Oyama T (2010) Three major output pathways from the KaiABC-based oscillator cooperate to generate robust circadian kaiBC expression in cyanobacteria. *Proc Natl Acad Sci U S A* 107(7): 3263-3268.

35. Schmitz O, Katayama M, Williams SB, Kondo T & Golden SS (2000) CikA, a bacteriophytochrome that resets the cyanobacterial circadian clock. *Science* 289(5480): 765-768.

36. Taniguchi Y, *et al* (2007) labA: A novel gene required for negative feedback regulation of the cyanobacterial circadian clock protein KaiC. *Genes Dev* 21(1): 60-70.

37. Katayama M, Kondo T, Xiong J & Golden SS (2003) ldpA encodes an iron-sulfur protein involved in light-dependent modulation of the circadian period in the cyanobacterium synechococcus elongatus PCC 7942. *J Bacteriol* 185(4): 1415-1422.

38. Ivleva NB, Bramlett MR, Lindahl PA & Golden SS (2005) LdpA: A component of the circadian clock senses redox state of the cell. *EMBO J* 24(6): 1202-1210.

39. Mackey SR, *et al* (2008) Proteins found in a CikA interaction assay link the circadian clock, metabolism, and cell division in synechococcus elongatus. *J Bacteriol* 190(10): 3738-3746.

40. Gutu A & O'Shea EK (2013) Two antagonistic clock-regulated histidine kinases time the activation of circadian gene expression. *Mol Cell* 50(2): 288-294.

41. Kitayama Y, Iwasaki H, Nishiwaki T & Kondo T (2003) KaiB functions as an attenuator of KaiC phosphorylation in the cyanobacterial circadian clock system. *EMBO J* 22(9): 2127-2134.

42. Iwasaki H, Taniguchi Y, Ishiura M & Kondo T (1999) Physical interactions among circadian clock proteins KaiA, KaiB, and KaiC in cyanobacteria. *EMBO J* 18: 1137-1145.

43. Qin X, *et al* (2010) Intermolecular associations determine the dynamics of the circadian KaiABC oscillator. *Proc Natl Acad Sci U S A* 107(33): 14805-14810.

44. Takai N, *et al* (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(32): 12109-12114.

45. Shultzaberger RK, Boyd JS, Katsuki T, Golden SS & Greenspan RJ (2014) Single mutations in sasA enable a simpler DeltacikA gene network architecture with equivalent circadian properties. *Proc Natl Acad Sci U S A* 111(47): E5069-75.

46. Espinosa J, *et al* (2015) Cross-talk and regulatory interactions between the essential response regulator RpaB and cyanobacterial circadian clock output. *Proc Natl Acad Sci U S A* 112(7): 2198-2203.

47. Pattanayek R, *et al* (2004) Visualizing a circadian clock protein: Crystal structure of KaiC and functional insights. *Mol Cell* 15(3): 375-388.

48. Chang YG, Tseng R, Kuo NW & LiWang A (2012) Rhythmic ring-ring stacking drives the circadian oscillator clockwise. *Proc Natl Acad Sci U S A* 109(42): 16847-16851.

49. Vakonakis I & LiWang AC (2004) Structure of the C-terminal domain of the clock protein KaiA in complex with a KaiC-derived peptide: Implications for KaiC regulation. *Proc Natl Acad Sci U S A* 101(30): 10925-10930.

50. Kim Y, Dong G, Carruthers CW, Golden SS & LiWang A (2008) The day/night switch in KaiC, a central oscillator component of the circadian clock of cyanobacteria. *Proc Natl Acad Sci U S A* 105(35): 12825-12830.

51. Pattanayek R & 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 Ã… resolution. *Biochemistry (N Y)* 54(30): 4575-4578.

52. Chang Y, Kuo N, Tseng R & LiWang A (2011) Flexibility of the C-terminal, or CII, ring of KaiC governs the rhythm of the circadian clock of cyanobacteria. *Proc Natl Acad Sci USA* 108(35): 14431-14436.

53. Mutoh R, Nishimura A, Yasui S, Onai K & Ishiura M (2013) The ATP-mediated regulation of KaiB-KaiC interaction in the cyanobacterial circadian clock. *PLoS One* 8(11): e80200.

54. Rust MJ, Markson JS, Lane WS, Fisher DS & O'Shea EK (2007) Ordered phosphorylation governs oscillation of a three-protein circadian clock. *Science* 318(5851): 809-812.

55. Nishiwaki T, *et al* (2007) A sequential program of dual phosphorylation of KaiC as a basis for circadian rhythm in cyanobacteria. *EMBO J* 26: 4029-4037.

56. Snijder J, *et al* (2014) Insight into cyanobacterial circadian timing from structural details of the KaiB-KaiC interaction. *Proc Natl Acad Sci U S A* 111(4): 1379-1384.

57. Villarreal SA, *et al* (2013) CryoEM and molecular dynamics of the circadian KaiB–KaiC complex indicates that KaiB monomers interact with KaiC and block ATP binding clefts. *J Mol Biol* 425(18): 3311-3324.

58. Pattanayek R, Yadagirl KK, Ohi MD & Egli M (2013) Nature of KaiB-KaiC binding in the cyanobacterial circadian oscillator. *Cell Cycle* 12: 810-817.

59. Chang YG, *et al* (2015) Circadian rhythms. A protein fold switch joins the circadian oscillator to clock output in cyanobacteria. *Science* 349(6245): 324-328.

60. Tseng R, *et al* (2014) Cooperative KaiA-KaiB-KaiC interactions affect KaiB/SasA competition in the circadian clock of cyanobacteria. *J Mol Biol* 426(2): 389-402.

61. Phong C, Markson JS, Wilhoite CM & Rust MJ (2013) Robust and tunable circadian rhythms from differentially sensitive catalytic domains. *Proc Natl Acad Sci U S A* 110(3): 1124-1129.

62. Murakami R, *et al* (2012) The roles of the dimeric and tetrameric structures of the clock protein KaiB in the generation of circadian oscillations in cyanobacteria. *J Biol Chem* 287: 29506-29515.

63. Mori T, *et al* (2007) Elucidating the ticking of an in vitro circadian clockwork. *PLoS Biol* 5(4): e93.

64. Pattanayek R, *et al* (2011) Combined SAXS/EM based models of the S. elongatus post-translational circadian oscillator and its interactions with the output his-kinase SasA. *PLoS ONE* 6(8): e23697.

65. Pattanayek R, *et al* (2008) Structural model of the circadian clock KaiB–KaiC complex and mechanism for modulation of KaiC phosphorylation. *EMBO J* 27: 1767-1778.

66. Krell T, *et al* (2010) Bacterial sensor kinases: Diversity in the recognition of environmental signals. *Annu Rev Microbiol* 64(1): 539-559.

67. Valencia S. J, *et al* (2012) Phase-dependent generation and transmission of time information by the KaiABC circadian clock oscillator through SasA-KaiC interaction in cyanobacteria. *Genes to Cells* 17(5): 398-419.

68. Vakonakis I, Klewer DA, Williams SB, Golden SS & LiWang AC (2004) Structure of the N-terminal domain of the circadian clock-associated histidine kinase SasA. *J Mol Biol* 342(1): 9-17.

69. Zhang X, Dong G & Golden SS (2006) The pseudo-receiver domain of CikA regulates the cyanobacterial circadian input pathway. *Mol Microbiol* 60(3): 658-668.

70. Cohen SE, *et al* (2014) Dynamic localization of the cyanobacterial circadian clock proteins. *Curr Biol* 24(16): 1836-1844.

71. Mutsuda M, Michel KP, Zhang X, Montgomery BL & Golden SS (2003) Biochemical properties of CikA, an unusual phytochrome-like histidine protein kinase that resets the circadian clock in synechococcus elongatus PCC 7942. *J Biol Chem* 278(21): 19102-19110.

72. Gao T, Zhang X, Ivleva NB, Golden SS & LiWang A (2007) NMR structure of the pseudo-receiver domain of CikA. *Protein Sci* 16(3): 465-475.

73. Dunlap, Jay C., Loros, Jennifer J., DeCoursey,Patricia J., (2004) *Chronobiology : biological timekeeping,* (Sinauer Associates, Sunderland, Mass.),

74. Jeyaraj D*, et al* (2012) Circadian rhythms govern cardiac repolarization and arrhythmogenesis. *Nature* 483(7387): 96-99.

75. Woelfle MA, Ouyang Y, Phanvijhitsiri K & Johnson CH (2004) The adaptive value of circadian clocks: An experimental assessment in cyanobacteria. *Curr Biol* 14(16): 1481-1486.

76. Oliver PL*, et al* (2012) Disrupted circadian rhythms in a mouse model of schizophrenia. *Curr Biol* 22(4): 314-319.

77. Goodspeed D, Chehab EW, Min-Venditti A, Braam J & Covington MF (2012) Arabidopsis synchronizes jasmonate-mediated defense with insect circadian behavior. *Proc Natl Acad Sci U S A* 109(12): 4674-4677.

78. Bass J (2012) Circadian topology of metabolism. *Nature* 491(7424): 348-356.

79. Yoshida T, Murayama Y, Ito H, Kageyama H & Kondo T (2009) Nonparametric entrainment of the in vitro circadian phosphorylation rhythm of cyanobacterial KaiC by temperature cycle. *Proceedings of the National Academy of Sciences* 106(5): 1648-1653.

80. Goda K, Ito H, Kondo T & Oyama T (2012) Fluorescence correlation spectroscopy to monitor kai protein-based circadian oscillations in real time. *Journal of Biological Chemistry* 287(5): 3241-3248.

81. Leipe DD, Aravind L, Grishin NV & Koonin EV (2000) The bacterial replicative helicase DnaB evolved from a RecA duplication. *Genome Res* 10(1): 5-16.

82. Pattanayek R*, et al* (2009) Structures of KaiC circadian clock mutant proteins: A new phosphorylation site at T426 and mechanisms of kinase, ATPase and phosphatase. *PLoS One* 4(11): e7529.

83. Pattanayek R*, et al* (2006) Analysis of KaiA–KaiC protein interactions in the cyanobacterial circadian clock using hybrid structural methods. *EMBO J* 25: 2017-2028.

84. Nishiwaki T & Kondo T (2012) The circadian autodephosphorylation of cyanobacterial clock protein KaiC occurs via the formation of ATP as an intermediate. *Journal of Biological Chemistry* 287: 18030-18035.

85. Egli M*, et al* (2012) Dephosphorylation of the core clock protein KaiC in the cyanobacterial KaiABC circadian oscillator proceeds via an ATP synthase mechanism. *Biochemistry (N Y)* 51(8): 1547-1558.

86. Kim Y, Vinyard DJ, Ananyev GM, Dismukes GC & Golden SS (2012) Oxidized quinones signal onset of darkness directly to the cyanobacterial circadian oscillator. *Proc Natl Acad Sci USA* 109(44): 17765-17769.

87. Ivleva NB, Gao T, LiWang AC & Golden SS (2006) Quinone sensing by the circadian input kinase of the cyanobacterial circadian clock. *Proc Natl Acad Sci U S A* 103(46): 17468-17473.

88. Wood TL, *et al* (2010) The KaiA protein of the cyanobacterial circadian oscillator is modulated by a redox-active cofactor. *Proc Natl Acad Sci U S A* 107(13): 5804-5809.

89. Pattanayek R, Sidiqi SK & Egli M (2012) Crystal structure of the redox-active cofactor dibromothymoquinone bound to circadian clock protein KaiA and structural basis for dibromothymoquinone's ability to prevent stimulation of KaiC phosphorylation by KaiA. *Biochemistry (N Y)* 51(41): 8050-8052.

90. Rust MJ, Golden SS & O'Shea EK (2011) Light-driven changes in energy metabolism directly entrain the cyanobacterial circadian oscillator. *Science* 331(6014): 220-223.

91. Liu Y, *et al* (1995) Circadian orchestration of gene expression in cyanobacteria. *Genes Dev* 9(12): 1469-1478.

92. Ito H, *et al* (2009) Cyanobacterial daily life with kai-based circadian and diurnal genome-wide transcriptional control in synechococcus elongatus. *Proceedings of the National Academy of Sciences* 106(33): 14168-14173.

93. Vijayan V, Zuzow R & O'Shea EK (2009) Oscillations in supercoiling drive circadian gene expression in cyanobacteria. *Proc Natl Acad Sci USA* 106: 22564-22568.

94. Nakahira Y, *et al* (2004) Global gene repression by KaiC as a master process of prokaryotic circadian system. *Proc Natl Acad Sci USA* 101(3): 881-885.

95. Dong G, *et al* (2010) Elevated ATPase activity of KaiC applies a circadian checkpoint on cell division in synechococcus elongatus. *Cell* 140(4): 529-539.

96. Yang Q, Pando BF, Dong G, Golden SS & van Oudenaarden A (2010) Circadian gating of the cell cycle revealed in single cyanobacterial cells. *Science* 327(5972): 1522-1526.

97. Smith RM & Williams SB (2006) Circadian rhythms in gene transcription imparted by chromosome compaction in the cyanobacterium synechococcus elongatus. *Proc Natl Acad Sci USA* 103(22): 8564-8569.

98. Iwasaki H, Nishiwaki T, Kitayama Y, Nakajima M & Kondo T (2002) KaiA-stimulated KaiC phosphorylation in circadian timing loops in cyanobacteria. *Proc Natl Acad Sci U S A* 99(24): 15788-15793.

99. Pollard TD (2010) A guide to simple and informative binding assays. *Mol Biol Cell* 21(23): 4061-4067.

100. Delaglio F, *et al* (1995) NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* 6: 277-293.

101. Mutoh R, *et al* (2010) Direct interaction between KaiA and KaiB revealed by a site-directed spin labeling electron spin resonance analysis. *Genes Cells* 15(3): 269-280.

102. Golden SS, Ishiura M, Johnson CH & Kondo T (1997) Cyanobacterial circadian rhythms. *Annu Rev Plant Physiol Plant Mol Biol* 48: 327-354.

103. Garces RG, Wu N, Gillon W & Pai EF (2004) Anabaena circadian clock proteins KaiA and KaiB reveal a potential common binding site to their partner KaiC. *EMBO J* 23(8): 1688-1698.

104. Uzumaki T, *et al* (2004) Crystal structure of the C-terminal clock-oscillator domain of the cyanobacterial KaiA protein. *Nat Struct Mol Biol* 11(7): 623-631.

105. Sprangers R & Kay LE (2007) Quantitative dynamics and binding studies of the 20S proteasome by NMR. *Nature* 445: 618-622.

106. Hitomi K, Oyama T, Han S, Arvai AS & Getzoff ED (2005) Tetrameric architecture of the circadian clock protein KaiB: A novel interface for intermolecular interactions and its impact on the circadian rhythm. *J Biol Chem* 280(19): 19127-19135.

107. Iwase R, *et al* (2005) Functionally important substructures of circadian clock protein KaiB in a unique tetramer complex. *J Biol Chem* 280(52): 43141-43149.

108. Ye S, Vakonakis I, Ioerger TR, LiWang AC & Sacchettini JC (2004) Crystal structure of circadian clock protein KaiA from synechococcus elongatus. *J Biol Chem* 279(19): 20511-20518.

109. Vakonakis I, *et al* (2004) NMR structure of the KaiC-interacting C-terminal domain of KaiA, a circadian clock protein: Implications for the KaiA-KaiC interaction. *Proc Natl Acad Sci USA* 101: 1479-1484.

110. Hayashi F, *et al* (2004) Stoichiometric interactions between cyanobacterial clock proteins KaiA and KaiC. *Biochem Biophys Res Commun* 316: 195-202.

111. Ma L & Ranganathan R (2012) Quantifying the rhythm of KaiB-C interaction for In Vitro cyanobacterial circadian clock. *PLoS ONE* 7(8): e42581.

112. van Zon JS, Lubensky DK, Altena PRH & ten Wolde PR (2007) An allosteric model of circadian KaiC phosphorylation. *Proc Natl Acad Sci USA* 104(18): 7420-7425.

113. Murayama Y, *et al* (2011) Tracking and visualizing the circadian ticking of the cyanobacterial clock protein KaiC in solution. *EMBO J* 30(1): 68-78.

114. Terauchi K, *et al* (2007) ATPase activity of KaiC determines the basic timing for circadian clock of cyanobacteria. *Proc Natl Acad Sci USA* 104: 16377-16381.

115. Kageyama H, *et al* (2006) Cyanobacterial circadian pacemaker: Kai protein complex dynamics in the KaiC phosphorylation cycle in vitro. *Mol Cell* 23(2): 161-171.

116. Nakajima M, Ito H & Kondo T (2010) In vitro regulation of circadian phosphorylation rhythm of cyanobacterial clock protein KaiC by KaiA and KaiB. *FEBS Lett* 584(5): 898-902.

117. Read, Randy J., Sussman,Joel L., (2007) Evolving methods for macromolecular crystallography the structural path to the understanding of the mechanism of action of CBRN agents

118. Winn MD, *et al* (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* 67(Pt 4): 235-242.

119. Evans P (2006) Scaling and assessment of data quality. *Acta Crystallogr D Biol Crystallogr* 62(Pt 1): 72-82.

120. Evans PR (2011) An introduction to data reduction: Space-group determination, scaling and intensity statistics. *Acta Crystallogr D Biol Crystallogr* 67(Pt 4): 282-292.

121. Otwinowski Z. & Minor W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *METHODS ENZYMOL* 276: 307-326.

122. McCoy AJ, *et al* (2007) Phaser crystallographic software. *J Appl Crystallogr* 40(Pt 4): 658-674.

123. Adams PD, *et al* (2010) PHENIX: A comprehensive python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2): 213-221.

124. Emsley P, Lohkamp B, Scott WG & Cowtan K (2010) Features and development of coot. *Acta Crystallogr D Biol Crystallogr* 66(Pt 4): 486-501.

125. Clerico EM, Cassone VM & Golden SS (2009) Stability and lability of circadian period of gene expression in the cyanobacterium synechococcus elongatus. *Microbiology* 155(2): 635-641.

126. Clerico EM, Ditty JL & Golden SS (2007) in *Methods in Molecular Biology*, ed Rosato E (Humana Press, New Jersey), pp 155-172.

127. Mackey SR, Ditty JL, Clerico EM & Golden SS (2007) Detection of rhythmic bioluminescence from luciferase reporters in cyanobacteria. *Methods Mol Biol* 362: 115-129.

128. Garrett DS, Powers R, Gronenborn AM & Clore GM (1991) A common sense approach to peak peaking in two-, three-, and four-dimensional spectra using automatic computer analysis of contour diagrams. *J.Magn.Res*. 95: 214-220.

129. Zhang L, *et al* (2006) Solution structure of the complex between poxvirus-encoded CC chemokine inhibitor vCCI and human MIP-1ß. *Proc Natl Acad Sci USA* 103(38): 13985-13990.

130. Shen Y, Delaglio F, Cornilescu G & Bax A (2009) TALOS+: A hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J Biomol NMR* 44(4): 213-223.

131. Schwieters CD, Kuszewski JJ & Marius Clore G (2006) Using Xplor–NIH for NMR molecular structure determination. *Prog Nucl Magn Reson Spectrosc* 48(1): 47-62.

132. Tian Y, Schwieters CD, Opella SJ & Marassi FM (2014) A practical implicit solvent potential for NMR structure calculation. *J Magn Reson* 243: 54-64.

133. Chou J, Li S & Bax A (2000) Study of conformational rearrangement and refinement of structural homology models by the use of heteronuclear dipolar couplings. *J Biomol NMR* 18(3): 217-227.

134. Bhattacharya A, Tejero R & Montelione GT (2007) Evaluating protein structures determined by structural genomics consortia. *Proteins* 66(4): 778-795.

135. Chen VB, *et al* (2010) MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* 66(Pt 1): 12-21.

136. Whitton, Brian A.,Potts, Malcolm., (2000) The ecology of cyanobacteria their diversity in time and space

137. Jungblut AD, Lovejoy C & Vincent WF (2010) Global distribution of cyanobacterial ecotypes in the cold biosphere. *ISME J* 4(2): 191-202.

138. Iida T, *et al* (2014) Importance of the monomer-dimer-tetramer interconversion of the clock protein KaiB in the generation of circadian oscillations in cyanobacteria. *Genes Cells* : n/a-n/a.

139. Abe J, *et al* (2015) Atomic-scale origins of slowness in the cyanobacterial circadian clock. *Science* 349(6245): 312-316.

140. Krissinel E & Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* 372(3): 774-797.

141. Partch CL, Green CB & Takahashi JS (2014) Molecular architecture of the mammalian circadian clock. *Trends Cell Biol* 24(2): 90-99.

APPENDIX A

PROTOCOLS AND SCRIPTS FOR PREPARING NOE FOR XPLOR-NIH FORMAT

README_APPENDIX_A.txt contains the protocol instructions for preparing NOE for XPLOR-NIH format. All of the scripts included in Appendix A are provided as teaching material from advanced structure refinement workshop at the National Magnetic Resonance Facility at Madison (NMRFAM).

Beginning of README_APPENDIX_A.txt

*** Xipp and Pipp are fully compatible with XPLOR-NIH, no atom name translation needed, just NOE calibration (general, not specific to Pipp):
(example adapted from: duo:/home/noi/work/pixj/pg/nmr/c13noesy/xxplor)

- input file: n15noesy.PCK (Xipp/Pipp NOE file)
- input file: c13noesy.PCK (Xipp/Pipp NOE file)

1. Look for minimum and maximum cross-peak (NON-DIAGONAL!) NOE intensities from n15noesy.PCK/c13noe.PCK and put them in 3dhhc_methyl_n15.awk/3dhhc_methyl_c13.awk:

-> sort -g -k 6 c13noesy.PCK > c13noe.PCK.sorted
-> sort -g -k 6 n15noesy.PCK > n15noe.PCK.sorted
-> edit c13noe.PCK.sorted
-> edit n15noe.PCK.sorted

2. Edit 3dhhc_methyl_n15.awk (for N15 NOE) or 3dhhc_methyl_c13.awk (for C13 NOE) scripts to use these extrema intensities (from noe.PCK.sorted):

MAX_INTENSITY = 1.8e08 (example)
MIN_INTENSITY = 2.07e06 (example)

3. In 3dhhc_methyl_n15.awk/3dhhc_methyl_c13.awk, you may need to adjust:

CUTOFF_DIST = 5.0 (4.5 to 5)
VDW_DIST = 1.8 (1.8 to 2.3)

such that:
- for alpha helices d [HA(i) - HN(i+3)] ~ 3.4 A (for N15 noe)
- if no helices, d [orthoH - metaH] ~ 2.5 A in aromatics etc.

- for C13, use proline proton distances or use a homologous structure in PDB as referece to calibrate similar distance.

The NOE calibration formula in 3dhhc_methyl_n15.awk/3dhhc_methyl_c13.awk is:

absi = VDW_DIST + (-1.0 + absi) * K
K = (CUTOFF_DIST - VDW_DIST)/(-1.0 + (MAX_INTENSITY/MIN_INTENSITY)^1/3)

-can change 1/3 dependence between 1/6 to 1. 1/3 is default.

4. Run the script:

-> 3dhhc_methyl_n15.awk null n15noesy.PCK > n15noe-1-orig
-> 3dhhc_methyl_c13.awk null c13noesy.PCK > c13noe-1-orig

-The purpose of 3dhhc_methyl_n15.awk/3dhhc_methyl_c13.awk will:

==flip sign for aliased (negative) peaks and calibrates NOEs:

absi = VDW_DIST + [-1.0 + (MAX_INTENSITY/NOE_INT)^1/3] * K

==eliminate diagonal peaks, weird labels, unassigned peaks

if(res1==res2 && ass1==ass2) {next}

==reorder constraints to have the smallest residue ID first (needed for XPLOR)!!!

==generate NOE constraints with 40% tolerance (spin diffusion + exp. error)

res1, ass1, res2, ass2, absi, 0.40*absi, 0.40*absi

5. Check range of calibrated NOE restraints:

-> sort -g -k 14 n15noe-1-orig > n15noe-1-orig.sorted
-> sort -g -k 14 c13noe-1-orig > c13noe-1-orig.sorted

->edit n15noe-1-orig.sorted
->edit c13noe-1-orig.sorted

6. Add protein sequence to the top of n15noe-1/c13noe-1 file:
-> cp n15noe-1-orig n15noe-1
-> cp c13noe-1-orig c13noe-1

->edit n15noe-1
->edit c13noe-1

==example of protein sequence format:
```
DATA SEQUENCE MQIFVKTLTG KTITLEVEPS DTIENVKAKI QDKEGIPPDQ QRLIFAGKQL
DATA SEQUENCE EDGRTLSDYN IQKESTLHLV LRLRGG
```
(http://spin.niddk.nih.gov/bax/nmrserver/dc/dObsA.tab)

7. change most assignments as non-stereospecific :

-> nonstereo_2.tcl n15noe-1 1 > n15noe-2
-> nonstereo_2.tcl c13noe-1 1 > c13noe-2

==for example, this script will take:
assign (resid 147 and name HA) (resid 150 and name HB1) 4.2 1.7 1.7

==and replaced with:
assign (resid 147 and name HA) (resid 150 and name HB1|HB2) 4.2 1.7 1.7

8. Eliminate duplicate NOEs:

-> sort -n -k 3,3 -k9,9 n15noe-2 > n15noe-3
-> sort -n -k 3,3 -k9,9 c13noe-2 > c13noe-3

-> select.tcl n15noe-3 > n15noe-4
-> select.tcl c13noe-3 > c13noe-4

9. Eliminate symmetric ambiguous NOEs:

-> cut.tcl n15noe-4 > n15noe-5
-> cut.tcl c13noe-4 > c13noe-5

==For example, this NOE will be removed (Gln-140):
assign (resid 140 and name HG1|HG2) (resid 140 and name HG2|HG1) 3.1 1.3 1.3

10. Transform 'l' in 'or' to generate the final distance restraint file to be used with XPLOR square NOE potential:

-> pipe2or.awk n15noe-5 > n15noe-square.tbl
-> pipe2or.awk c13noe-5 > c13noe-square.tbl

11. Open the lower distance limit for soft NOE potential (except for single proton pairs) to generate the final distance restraint file for XPLOR soft NOE potential):

-> noe-open.tcl n15noe-square.tbl > n15noe-soft.tbl
-> noe-open.tcl c13noe-square.tbl > c13noe-soft.tbl

== Use SOFT noe potential in XPLOR to obtain initial fold (in fold.py)
== Use SQUARE noe potential in XPLOR for final refinement (in refine.py)

12. Run noestat_v2.sh to obtain NOE statistic, this script will prompt for input.
-> noestat_v2.sh

End of README_APPENDIX_A.txt

Beginning of 3dhhc_methyl_n15.awk

```
#!/usr/bin/awk -f
################################################################################
#########
# convert 3D 15N, 13C NOE PCK file to XPLOR
# usage: 3dhhn.awk null 3dnoe.PCK
################################################################################
#########

BEGIN {
 skip = 1
 deleted = 0
 cutoff  = 0

#example: MAX_INTENSITY = 1.03e07
MAX_INTENSITY = 1.80e08
MIN_INTENSITY = 2.07e06

CUTOFF_DIST = 5.00
VDW_DIST = 2.00

#K = 0.34
K = (CUTOFF_DIST - VDW_DIST)/(-1.0 + (MAX_INTENSITY/MIN_INTENSITY)^0.166667)
#print K
}

################################################################################
#########

################################################################################
#########
# Read in exclude defs definitions
################################################################################
#########
FILENAME == ARGV[1] {
        if  ($1 == "#" || NF == 0) next
        exclude[$1] = 1
        next
}
################################################################################
#########
# skip over "header"
################################################################################
#########
FILENAME == ARGV[2] && skip {
            while ($2 != "PkID")
                    {
                     getline
                     }
            getline
            skip = 0
```

```
}

#############################################################################
#########
function mk_ass(ass) {
 gsub(/[,;]/,".",ass)
 ns = split(ass, aa, ".")
 ASS = sprintf( "%s ", aa[3])
 return substr(aa[1],2,length(aa[1]))
}

function mk_ass2(ass) {
 gsub(/[,;]/,".",ass)
 ns = split(ass, aa, ".")
 ASS = sprintf( "%s ", aa[2])
 return substr(aa[1],2,length(aa[1]))
}
#############################################################################
#########
function pr_output( swap, res1, ass1, res2, ass2, dist, neg, pos) {
 if (swap)
          printf "assign (resid %-3s and name %-4s) (resid %-3s and name %-4s) ",res2, ass2,
res1, ass1

 else
          printf "assign (resid %-3s and name %-4s) (resid %-3s and name %-4s) ",res1, ass1,
res2, ass2

# output constraints
 printf "%4.1f %4.1f %4.1f\n", dist, neg, pos

}
#############################################################################
#########
{
# screen for unwanted peak ids

 if ($1 in exclude)
       {if (deleted == 0)
                  print $0 > "noe.deleted"
            else
                  print $0 >> "noe.deleted"
            deleted++
        next
       }

#skip unassigned
 if (match($7,  /[ACDEFGHIKLMNPQRSTVWY][0-9]*\./) == 0 || \
    match($8,  /[ACDEFGHIKLMNPQRSTVWY][0-9]*\./) == 0 )
       {next}

 res1 = mk_ass($7)
```

```
 ass1 = ASS
 res2 = mk_ass2($8)
 ass2 = ASS

# printf ">>> res1=%s, ass1=%s, res2=%s, ass2=%s\n",res1,ass1,res2,ass2

# always put smallest residue id first, or "lower" proton identifier first for
# intra-residue
 swap = (int(res1) > int(res2)) || (int(res1) == int(res2) && ass1 > ass2)

# calc constraint limits and print output

 absi = ($6 < 0.0) ? -1.0*$6 : $6
 absi = MAX_INTENSITY / absi
 absi = (absi)^0.166667
 absi = VDW_DIST + (-1.0 + absi) * K

if(res1==res2 && ass1==ass2) {next}


 if (absi > CUTOFF_DIST)
          {if (cutoff == 0)
                   {print $0 > "noe.cutoff"}
           else
                   {print $0 >> "noe.cutoff"}
           cutoff++
          }
 else if (absi <= CUTOFF_DIST && absi >= WM)
   {{lower=0.4*absi;upper=0.4*absi}
   if (match($0,  /#/) != 0)
      {upper=0.4*absi+0.5}
      {pr_output(swap, res1, ass1, res2, ass2, absi, lower, upper )}}

 else if (absi < WM && absi >= MS)
   {{lower=0.4*absi;upper=0.4*absi}
   if (match($0,  /#/) != 0)
      {upper=0.4*absi+0.5}
      {pr_output(swap, res1, ass1, res2, ass2, absi, lower, upper )}}

 else if (absi < MS && absi >= 2.2)
   {{lower=0.4*absi;upper=0.4*absi}
   if (match($0,  /#/) != 0)
      {upper=0.4*absi+0.5}
      {pr_output(swap, res1, ass1, res2, ass2, absi, lower, upper )}}

 else if (absi < 2.2)
   {{lower=0.4*absi;upper=0.4*absi}
   if (match($0,  /#/) != 0)
      {upper=0.4*absi+0.5}
      {pr_output(swap, res1, ass1, res2, ass2, absi, lower, upper )}}
}
```

End of 3dhhc_methyl_n15.awk

Beginning of 3dhhc_methyl_c13.awk

```
#!/usr/bin/awk -f
################################################################################
#########
# convert 3D 15N, 13C NOE PCK file to XPLOR
# usage: 3dhhn.awk null 3dnoe.PCK
################################################################################
#########

BEGIN {
 skip = 1
 deleted = 0
 cutoff  = 0

#example, MAX_INTENSITY = 1.03e07
MAX_INTENSITY = 9.87e07
MIN_INTENSITY = 2.02e06

CUTOFF_DIST = 5.00
VDW_DIST = 1.80

#K = 0.34
K = (CUTOFF_DIST - VDW_DIST)/(-1.0 + (MAX_INTENSITY/MIN_INTENSITY)^1)
#print K
}

################################################################################
#########

################################################################################
#########
# Read in exclude defs definitions
################################################################################
#########
FILENAME == ARGV[1] {
        if  ($1 == "#" || NF == 0) next
        exclude[$1] = 1
        next
}
################################################################################
#########
# skip over "header"
################################################################################
#########
FILENAME == ARGV[2] && skip {
            while ($2 != "PkID")
                    {
                     getline
                     }
```

```
            getline
            skip = 0
}

##############################################################################
#########
function mk_ass(ass) {
 gsub(/[,;]/,".",ass)
 ns = split(ass, aa, ".")
 ASS = sprintf( "%s ", aa[3])
 return substr(aa[1],2,length(aa[1]))
}

function mk_ass2(ass) {
 gsub(/[,;]/,".",ass)
 ns = split(ass, aa, ".")
 ASS = sprintf( "%s ", aa[2])
 return substr(aa[1],2,length(aa[1]))
}
##############################################################################
#########
function pr_output( swap, res1, ass1, res2, ass2, dist, neg, pos) {
 if (swap)
           printf "assign (resid %-3s and name %-4s) (resid %-3s and name %-4s) ",res2, ass2,
res1, ass1

 else
           printf "assign (resid %-3s and name %-4s) (resid %-3s and name %-4s) ",res1, ass1,
res2, ass2

# output constraints
 printf "%4.1f %4.1f %4.1f\n", dist, neg, pos

}
##############################################################################
#########
{
# screen for unwanted peak ids

 if ($1 in exclude)
       {if (deleted == 0)
                   print $0 > "noe.deleted"
            else
                   print $0 >> "noe.deleted"
            deleted++
        next
       }

#skip unassigned
 if (match($7,  /[ACDEFGHIKLMNPQRSTVWY][0-9]*\./) == 0 || \
     match($8,  /[ACDEFGHIKLMNPQRSTVWY][0-9]*\./) == 0 )
       {next}
```

```
 res1 = mk_ass($7)
 ass1 = ASS
 res2 = mk_ass2($8)
 ass2 = ASS

# printf ">>> res1=%s, ass1=%s, res2=%s, ass2=%s\n",res1,ass1,res2,ass2

# always put smallest residue id first, or "lower" proton identifier first for
# intra-residue
 swap = (int(res1) > int(res2)) || (int(res1) == int(res2) && ass1 > ass2)

# calc constraint limits and print output

 absi = ($6 < 0.0) ? -1.0*$6 : $6
 absi = MAX_INTENSITY / absi
 absi = (absi)^1
 absi = VDW_DIST + (-1.0 + absi) * K

if(res1==res2 && ass1==ass2) {next}


 if (absi > CUTOFF_DIST)
          {if (cutoff == 0)
                   {print $0 > "noe.cutoff"}
           else
                   {print $0 >> "noe.cutoff"}
           cutoff++
          }
 else if (absi <= CUTOFF_DIST && absi >= WM)
   {


{
   {lower=0.4*absi;upper=0.4*absi}

   #if ($0 ~ /#/) c++
   c = 0
   #print NF; NF: number of fields
   for (i=1; i<=NF; i++)
       if ($i ~ /#/)
           c++
   #print c

   if (c == 1)

      {upper=0.4*absi+0.5}

   else if (c >= 2)

      {upper=0.4*absi+1.0}
```

```
        {pr_output(swap, res1, ass1, res2, ass2, absi, lower, upper )}

    }
}

 else if (absi < WM && absi >= MS)
    {


{
    {lower=0.4*absi;upper=0.4*absi}

    #if ($0 ~ /#/) c++
    c = 0
    #print NF; NF: number of fields
    for (i=1; i<=NF; i++)
        if ($i ~ /#/)
            c++
    #print c

    if (c == 1)

        {upper=0.4*absi+0.5}

    else if (c >= 2)

        {upper=0.4*absi+1.0}

        {pr_output(swap, res1, ass1, res2, ass2, absi, lower, upper )}

    }
    }

 else if (absi < MS && absi >= 2.2)
     {


{
    {lower=0.4*absi;upper=0.4*absi}

    #if ($0 ~ /#/) c++
    c = 0
    #print NF; NF: number of fields
    for (i=1; i<=NF; i++)
        if ($i ~ /#/)
            c++
    #print c

    if (c == 1)

        {upper=0.4*absi+0.5}
```

```
    else if (c >= 2)

       {upper=0.4*absi+1.0}

       {pr_output(swap, res1, ass1, res2, ass2, absi, lower, upper )}

    }
 }
 else if (absi < 2.2)
    {


{
    {lower=0.4*absi;upper=0.4*absi}

    #if ($0 ~ /#/) c++
    c = 0
    #print NF; NF: number of fields
    for (i=1; i<=NF; i++)
        if ($i ~ /#/)
            c++
    #print c

    if (c == 1)

       {upper=0.4*absi+0.5}

    else if (c >= 2)

       {upper=0.4*absi+1.0}

       {pr_output(swap, res1, ass1, res2, ass2, absi, lower, upper )}

    }
    }
}
```

End of 3dhhc_methyl_c13.awk

Beginning of nonstereo_2.tcl

```
#!/bin/sh
# The next line restarts using nmrWish \
exec nmrWish "$0" -- "$@"
set auto_path    "[split $env(TCLPATH) :] $auto_path"

# converts PIPP assignments to non-stereospecific assignments
# GC 2006

if {$argc != 2} \
  {
   puts stderr "\nInsert the 1-letter protein sequence (e.g. from stapp.par) at the beginning of the
NOE input file!
                and specify startResID\n"
   puts stderr "usage: nonstereo.tcl file.in startResID\n"
   puts stderr "examples:\n nonstereo.tcl noe.in -10\n"
   puts stderr "or:\n nonstereo.tcl noe.in 15\n"
   puts stderr "or (no offset):\n nonstereo.tcl noe.in 1 \n"
   exit
  }

set inFile [lindex $argv 0]
set startResID [lindex $argv 1]
initResNames

#
# Get sequence from DATA SEQUENCE lines of the input file:

set tab       [gdbRead table -in $inFile]
set seqList   [getSeqList $tab firstRes]
#puts "seqList=$seqList"

set nRes      [llength $seqList]
set firstRes  [expr $firstRes + $startResID -1]
set lastRes   [expr $firstRes + $nRes - 1]

#puts "\nFile $inFile"
#puts " $nRes residues."
#puts " First Residue is $firstRes"
#puts " Last  Residue is $lastRes\n"

for {set resID $firstRes} {$resID <= $lastRes} {incr resID} \
  {
   set curRes          [expr $resID - $firstRes]
   set resName1($resID) [lindex $seqList $curRes]
   set resName3($resID) [getResName3 $resName1($resID)]

#    puts -nonewline "$resName3($resID) "
  }

set inFileID [open $inFile r]
```

```
while {[gets $inFileID line] > 0} \
  {
   set flag [lindex $line 0]

   if {($flag != "assign") && ($flag != "!assign")} \
     {
      continue
     }

   set d0 [lrange $line 0 4]
   set d1 [lrange $line 6 10]
   set d2 [lrange $line 12 end]

   set resID1 [lindex $line 2]
   set resID2 [lindex $line 8]

   set n1 [lindex $line 5]
   set n2 [lindex $line 11]

#puts "resName3($resID1)=$resName3($resID1) resName3($resID2)=$resName3($resID2)"

   switch $n1 \
    {
     "HA1" {set n1 "HA1|HA2"}
     "HA2" {set n1 "HA2|HA1"}
     "HB1" {set n1 "HB1|HB2"}
     "HB2" {set n1 "HB2|HB1"}
     "HG1" {set n1 "HG1|HG2"}
     "HG2" {set n1 "HG2|HG1"}
     "HD1" {
           if {$resName3($resID1) == "HIS" || $resName3($resID1) == "TRP"} \
             {
              set n1 "HD1"
             } \
           else \
             {
              set n1 "HD1|HD2"
             }
          }
     "HD2" {
           if {$resName3($resID1) == "ASP" || $resName3($resID1) == "HIS"} \
             {
              set n1 "HD2"
             } \
           else \
             {
              set n1 "HD2|HD1"
             }
          }
     "HE1" {
           if {$resName3($resID1) == "HIS" || $resName3($resID1) == "TRP"} \
```

```
            {
             set n1 "HE1"
            } \
          else \
            {
             set n1 "HE1|HE2"
            }
         }
    "HE2" {
          if {$resName3($resID1) == "GLU" || $resName3($resID1) == "HIS" ||
$resName3($resID1) == "TRP"} \
            {
             set n1 "HE2"
            } \
          else \
            {
             set n1 "HE2|HE1"
            }
         }
    "HG11" {set n1 "HG11|HG12"}
    "HG12" {set n1 "HG12|HG11"}
    "HG1#" {set n1 "HG1#|HG2#"}
    "HG2#" {
          if {$resName3($resID1) == "ILE"} \
            {
             set n1 "HG2#"
            } \
          elseif {$resName3($resID1) == "THR"} \
            {
             set n1 "HG2#"
            } \
          else \
            {
             set n1 "HG2#|HG1#"
            }
         }
    "HD1#" {
          if {$resName3($resID1) == "ILE"} \
            {
             set n1 "HD1#"
            } \
          else \
            {
             set n1 "HD1#|HD2#"
            }
         }
    "HD2#" {set n1 "HD2#|HD1#"}
    "HD21" {
          if {$resName3($resID1) == "ASN"} \
            {
             set n1 "HD21|HD22"
            }
```

```
        }
   "HD22" {
        if {$resName3($resID1) == "ASN"} \
          {
           set n1 "HD22|HD21"
          }
        }
   "HE21" {
        if {$resName3($resID1) == "GLN"} \
          {
           set n1 "HE21|HE22"
          }
        }
   "HE22" {
        if {$resName3($resID1) == "GLN"} \
          {
           set n1 "HE22|HE21"
          }
        }
  }
switch $n2 \
  {
   "HA1" {set n2 "HA1|HA2"}
   "HA2" {set n2 "HA2|HA1"}
   "HB1" {set n2 "HB1|HB2"}
   "HB2" {set n2 "HB2|HB1"}
   "HG1" {set n2 "HG1|HG2"}
   "HG2" {set n2 "HG2|HG1"}
   "HD1" {
        if {$resName3($resID2) == "HIS" || $resName3($resID2) == "TRP"} \
          {
           set n2 "HD1"
          } \
        else \
          {
           set n2 "HD1|HD2"
          }
        }
   "HD2" {
        if {$resName3($resID2) == "ASP" || $resName3($resID2) == "HIS"} \
          {
           set n2 "HD2"
          } \
        else \
          {
           set n2 "HD2|HD1"
          }
        }
   "HE1" {
        if {$resName3($resID2) == "HIS" || $resName3($resID2) == "TRP"} \
          {
           set n2 "HE1"
```

```
            } \
          else \
            {
             set n2 "HE1|HE2"
            }
         }
    "HE2" {
          if {$resName3($resID2) == "GLU" || $resName3($resID2) == "HIS" ||
$resName3($resID2) == "TRP"} \
            {
             set n2 "HE2"
            } \
          else \
            {
             set n2 "HE2|HE1"
            }
         }
    "HG11" {set n2 "HG11|HG12"}
    "HG12" {set n2 "HG12|HG11"}
    "HG1#" {set n2 "HG1#|HG2#"}
    "HG2#" {
          if {$resName3($resID2) == "ILE"} \
            {
             set n2 "HG2#"
            } \
          elseif {$resName3($resID2) == "THR"} \
            {
             set n2 "HG2#"
            } \
          else \
            {
             set n2 "HG2#|HG1#"
            }
         }
    "HD1#" {
          if {$resName3($resID2) == "ILE"} \
            {
             set n2 "HD1#"
            } \
          else \
            {
             set n2 "HD1#|HD2#"
            }
         }
    "HD2#" {set n2 "HD2#|HD1#"}
    "HD21" {
          if {$resName3($resID2) == "ASN"} \
            {
             set n2 "HD21|HD22"
            }
         }
    "HD22" {
```

```
            if {$resName3($resID2) == "ASN"} \
              {
               set n2 "HD22|HD21"
              }
            }
       "HE21" {
            if {$resName3($resID2) == "GLN"} \
              {
               set n2 "HE21|HE22"
              }
            }
       "HE22" {
            if {$resName3($resID2) == "GLN"} \
              {
               set n2 "HE22|HE21"
              }
            }
      }

    puts "$d0 $n1 $d1 $n2 $d2"
   }

exit
```

End of nonstereo_2.tcl

Beginning of select.tcl

```
#!/bin/sh
# The next line restarts using nmrWish \
exec nmrWish "$0" -- "$@"
set auto_path    "[split $env(TCLPATH) :] $auto_path"

if {$argc != 1} \
  {
   puts stderr "error usage: select.tcl file.in"
   exit
  }

set inFile [lindex $argv 0]
set inFileID [open $inFile r]

set list ""

while {[gets $inFileID line] > 0} \
  {
   set index [string last ")" $line]
   set sub [string range $line 0 $index]

   if {[lsearch $list $sub] == -1} \
     {
      lappend list $sub
      puts $line
     }

#puts "sub=$sub"
#puts "list=$list"

  }

exit
```

Eng of select.tcl

Beginning of cut.tcl

```
#!/bin/sh
# The next line restarts using nmrWish \
exec nmrWish "$0" -- "$@"
set auto_path    "[split $env(TCLPATH) :] $auto_path"

# eliminates symmetric ambiguous NOEs for protons or methyls (e.g. Val):
#cut.tcl c13noe-4 > c13noe-5
#e.g. these NOE will be removed (Gln-140):
#assign (resid 140 and name HG1|HG2 ) (resid 140 and name HG2|HG1 ) 3.1 1.3 1.3
#(Ile-138):
#assign (resid 138 and name HD1#|HG2# ) (resid 138 and name HG2#|HD1# ) 2.9 1.2 1.2

if {$argc != 1} \
  {
   puts stderr "error usage: cut.tcl file.in"
   exit
  }

set inFile [lindex $argv 0]
set inFileID [open $inFile r]

while {[gets $inFileID line] > 0} \
  {
   set test1 [lindex $line 5]

   if {[string first "|" $test1] == -1} \
     {
      puts $line
      continue
     }

   set res1 [lindex $line 2]
   set at1 [lindex [split $test1 |] 0]
   set at2 [lindex [split $test1 |] 1]

#puts "at1=$at1 at1=$at1"

   set carbon1 [string index $at1 1]
   set carbon2 [string index $at2 1]

   set test2 [lindex $line 11]

   if {[string first "|" $test2] == -1} \
     {
      puts $line
      continue
     }

   set res2 [lindex $line 8]
```

```
# this original version ignores the "bound to a common carbon" requirement!
    if {($res2 == $res1) && (( [lsearch [split $test2 l] $at1] != -1 ) ll\
                          ( [lsearch [split $test2 l] $at2] != -1 ))} \
       {
        continue
       }

#but NOE won't be removed (Ile-138):
#assign (resid 138 and name HD1#IHG2# ) (resid 138 and name HG2#IHD1# ) 2.9 1.2 1.2

# check if the "bound to a common carbon" requirement is needed!
#     if {($res2 == $res1) && ($carbon1 == $carbon2) && \
#         (( [lsearch [split $test2 l] $at1] != -1 ) ll\
#         (  [lsearch [split $test2 l] $at2] != -1 ))} \
#        {
#         continue
#        }

    puts $line
   }

exit
```

End of cut.tcl

Beginning of pipe2or.awk

```
#!/usr/bin/awk -f
###############################################################################
#########
# convert | into and or for xplor
###############################################################################
#########

BEGIN {
a=b=0
}

FILENAME==ARGV[1] {
a++
resnumber1[a]=$3
resnumber2[a]=$9
atom2[a]=$12
dis1[a]=$14
dis2[a]=$15
dis3[a]=$16
atom1[a]=$6
char1[a]=substr($6,4,1)
char2[a]=substr($6,5,1)
char3[a]=substr($12,4,1)
char4[a]=substr($12,5,1)
}
END {
for (i=1; i<=a; i++) {
     if ((char1[i]=="|" || char2[i]=="|") && (char3[i]!="|" && char4[i]!="|")) {
        z=split(atom1[i],splitatom,"|")
        printf"assign (resid %3d and (name %s or name %s  )) (resid %3d and name %s  )  %2.1f
%2.1f  %2.1f\n", resnumber1[i],  splitatom[1], splitatom[2], resnumber2[i], atom2[i], dis1[i], dis2[i],
dis3[i]
                                }
     if ((char3[i]=="|" || char4[i]=="|") && (char1[i]!="|" && char2[i]!="|")) {
        z=split(atom2[i],splitatom2,"|")
        printf"assign (resid %3d and name %s  ) (resid %3d and (name %s or name %s  )) %2.1f
%2.1f  %2.1f\n", resnumber1[i], atom1[i], resnumber2[i], splitatom2[1], splitatom2[2], dis1[i],
dis2[i], dis3[i]
                                }
     if ((char3[i]=="|" || char4[i]=="|") && (char1[i]=="|" || char2[i]=="|")) {
        z=split(atom2[i],splitatom2,"|")
        y=split(atom1[i],splitatom,"|")
        printf"assign (resid %3d and (name %s or name %s  )) (resid %3d and (name %s or name
%s  )) %2.1f  %2.1f  %2.1f\n", resnumber1[i], splitatom[1], splitatom[2], resnumber2[i],
splitatom2[1], splitatom2[2], dis1[i], dis2[i], dis3[i]
                                                        }
     if (char1[i]!="|" && char2[i]!="|" && char3[i]!="|" && char4[i]!="|") {
        printf"assign (resid %3d and name %s  ) (resid %3d and name %s  )  %2.1f  %2.1f
%2.1f\n", resnumber1[i], atom1[i], resnumber2[i], atom2[i], dis1[i], dis2[i], dis3[i]
                                }
```

```
        }
}
```

End of pipe2or.awk

Beginning of noe-open.tcl

```
#!/bin/sh
# The next line restarts using nmrWish \
exec nmrWish "$0" -- "$@"
set auto_path    "[split $env(TCLPATH) :] $auto_path"

if {$argc != 1} \
  {
   puts stderr "error usage: noe_open.tcl file.in"
   exit
  }

set inFile [lindex $argv 0]
set inFileID [open $inFile r]

while {[gets $inFileID line] > 0} \
  {
   set pos1_resid [expr [lsearch $line "(resid"] + 1]

   set substring [lrange $line $pos1_resid end]

   set pos2_resid [expr [lsearch $substring "(resid"] + $pos1_resid + 1]

   set i [lindex $line $pos1_resid]
   set j [lindex $line $pos2_resid]

   set string [join $line 0]

   set test1 [string first or $string]
   set test2 [string first # $string]
   set test3 [string first * $string]

   set test [expr $test1+$test2+$test3]

   if {($test != -3) || ($i == $j)} \
     {
      set  n [expr [llength $line] - 2]
      set nd [expr [llength $line] - 3]
           set d [lindex $line $nd]
      set newline [lreplace $line $n $n $d]
      puts $newline
     } \
   else\
     {
      puts $line
     }
  }

exit
```

End of noe-open.tcl

Beginning of noestat_v2.sh

```
#!/bin/sh

# NOE constraint counter
# this program generates a list of NOEs per residue
# and counts the intra/sequential/medium/long range NOEs
#
# Kuen-Phon Wu, v0.3 done on 2006/June/1
#              v0.4  make interactive mode, ask input, output file name (06/01/06)

# initialize, ask file name
read -p "what's the name of your NOE file (xplor-nih format)? " file_name
read -p "output file name for NOE count? " _output

# generate approxiate format file to count
cat $file_name | sed -e 's/residue/resid/g' > _temp_x0
cat _temp_x0 |sed -e 's/[()]/ /g' |sed -e 's/and/ /g' | sed -e 's/resid/ /g' |sed -e 's/or name //g' |sed -e
's/[A-Z][0-9]/ /g' |sed -e 's/#/ /g' |sed -e 's/[A-Z]/ /g'> _temp_x1
cat _temp_x1 | grep ^assign | awk '{print $2,$4,$6+$8}' | sort -nk 1 > _temp_xx

# find the NOE counts, the initial res # and the end res #
i=`wc -l _temp_xx | awk '{print $1}'`
res_bg=`head -1 _temp_xx | awk '{print $1}'`
res_end=`tail -1 _temp_xx | awk '{print $1}'`

# output the information
echo total NOE constraints: $i, from residue $res_bg to residue $res_end
keep_bg=$res_bg
res_end_1=`expr $res_end + 1`

while [ $res_bg -lt $res_end_1 ]
do
 cat _temp_xx |grep ^$res_bg' ' > _temp_$res_bg
 echo residue $res_bg has `wc -l _temp_$res_bg | awk '{print $1}'` NOEs >> $_output
 res_bg=`expr $res_bg + 1`
 done

echo " "
echo "NOE analysis done! output saved as $_output"
echo " "
echo "NOW counting NOEs, may take 10 sec to 1 min..."

# set variables for intra,inter-short/medium/long range NOEs
cp _temp_xx _temp_yy
intra=0
sequential=0
inter_2=0
inter_3=0
inter_4=0
```

```
large5=0
j=0

while [ $j -lt $i ]
 do
 data=`head -1 _temp_yy | awk '{print $1-$2}' | sed -e 's/-//g'`
  if [ $data -eq 0 ];
     then intra=`expr $intra + 1`
  elif [ $data -eq 1 ];
    then sequential=`expr $sequential + 1`
  elif [ $data -eq 2 ];
    then inter_2=`expr $inter_2 + 1`
  elif [ $data -eq 3 ];
    then  inter_3=`expr $inter_3 + 1`
  elif [ $data -eq 4 ];
    then inter_4=`expr $inter_4 + 1`
  else
      large5=`expr $large5 + 1`
 fi
 sed '1,1d' _temp_yy > _temp_z
 mv _temp_z _temp_yy
 j=`expr $j + 1`
 done

echo "================================================================="
echo total NOEs in $file_name residue $keep_bg to residue $res_end: $i
echo "================================================================="
echo "intra_residual (j-i = 0): " $intra
echo "sequential  (j-i = 1):    " $sequential
echo "medium range (j-i = 2 ):  " $inter_2
echo "medium range (j-i = 3 ):  " $inter_3
echo "medium range (j-i = 4 ):  " $inter_4
echo "long range (j-i >= 5 ):   " $large5
rm _temp_*
```

End of noestat_v2.sh

APPENDIX B

PROTOCOLS AND SCRIPTS FOR RUNNING XPLOR-NIH STRUCTURAL CALCULATION

README_APPENDIX_B.txt contains the protocol instructions for run XPLOR-NIH structure calculation python script. All of the scripts included in Appendix B are modified from XPLOR-NIH eginput folder example scripts, except seq2pdb.py is provided by NMRFAM.

Beginning of README_APPENDIX_B.txt

1. Generate protein.seq text file with 1 letter code protein amino acid to be calculated. Example, MAPLRKTAVLKLYVAGNTPNSVRALKTLANILEKEFK

2. Convert protein.seq to protein_extend.pdb, the .pdb file is the very initial input for Xplor-NIH.

-> pyXplor seq2pdb.py

3. Run the initial folding script following Tian et al (132) .

-> pyXplor -smp 24 fold.py

4. Run the refine script after initial folding similar to protocols in Tian et al. except no filtering top10 structures

-> pyXplor -smp 24 refine.py

5. Run rdc refine script by combining the approaches of Tian et al. and Chou et al. together. Chou et al. uses 2-step refinement to incorrperate RDC data.

-> pyXplor -smp 24 refine_eefx.py

6. Run second step of refinement by Chou et al (133) , except dihedral constant (CDIH) is kept at 200, not 50, also no selecting top structure for next round of calculation.

-> pyXplor -smp 24 refine_eefx_slow.py
End of README_APPENDIX_B.txt

Beginning of seq2pdb.py

```
xplor.requireVersion("2.12")

xplor.parseArguments() # check for typos on the command-line

#=================== Load Topology and Parameters
==============================
import protocol
protocol.initTopology('protein')
protocol.initParams('protein')


#========================== Import a Sequence
================================
import psfGen
### Sequence file containing complete amino acid sequence (1-letter or 3-letter)
seq = open("protein.seq").read()

### Can provide the amino acid sequence here instead of in a file
#seq="""
#MET GLN ILE CYS PHE VAL LYS THR LEU THR GLY LYS THR ILE THR LEU CYS
#"""
#======================== Create the PSF File
================================
psfGen.seqToPSF(seq, seqType='protein', startResid=1, \
            disulfide_bonds=[], singleChar=True)

### For 1-letter amino acid codes change singleChar to True
### For 3-letter amino acid codes change singleChar to False
### For disulfide bonds (actual bond), create resid number tuple: (1,78)
### Disulfide bonds can be represented by distance restraint instead by changing
### disulfide_bonds to disulfide_bridges.

#=================== Generate Extended Structure
==============================
import random
from atomAction import SetProperty
### This assigns X, Y, and Z coordinates to each atom. The Y and Z coordinates
### are random (but small enough (within a range of -0.5 to 0.5) to allow
### bonded atoms to form their bonds)
### and the X coordinate is the atom number divided by 10. This will result in
### an extended configuration along the X axis.
from vec3 import Vec3
for atom in AtomSel("all"):
   atom.setPos( Vec3(float(atom.index())/10,
                random.uniform(-0.5,0.5),
                random.uniform(-0.5,0.5)) )
   pass
### Performs a minimization to remove bond, angle, improper, and van der
### Waals violations.
protocol.fixupCovalentGeom(useVDW=1,maxIters=100)
```

```
### Automatically generates extended structure
#protocol.genExtendedStructure()

#====================== Output PSF and PDB Files
==============================
xplor.command("write psf output=PrepOutput/protein.psf end")
protocol.writePDB("PrepOutput/protein_ext.pdb")
```

End of seq2pdb.py

Beginning of fold.py

```
import protocol

opts,args = xplor.parseArguments(["quick:0"])
quick=True if opts and opts[0][0] == "quick" else False

# Total number of structures.
numberOfStructures=1 if quick else 100

print numberOfStructures

# Base name for output pdb files.
outfilename = './fold_result/SCRIPT_STRUCTURE.pdb'

# Initialize random seed.
protocol.initRandomSeed(3421)

# topology and parameters specific to EEFx
#
protocol.topology['protein']  ="protein_eef.top"
protocol.parameters['protein']="protein_eef.par"
protocol.initParams("protein")

# Generate PSF - using EEFx topology
import psfGen
psfGen.pdbToPSF('protein_ext.pdb')

# Generate extended structure.
protocol.genExtendedStructure()

# Generate cis prolines
import psfGen
psfGen.cisPeptide(62)
psfGen.cisPeptide(69)
psfGen.cisPeptide(71)

#list to hold potential terms in structure calculation
from potList import PotList
potList = PotList()
crossTerms = PotList('cross')

#compute RMSD to reference structure
#  backbone and heavy atom RMSDs will be printed in the output
#  structure files
#
#from posDiffPotTools import create_PosDiffPot
#refRMSD = create_PosDiffPot("refRMSD",
                      # "(name CA or name C or name N) and resid 1:230",
                      # pdbFile='1EZA.pdb',
                      # cmpSel="not name H*")
```

```
# lists to hold force constant settings for various stages of
# structure calculation
from simulationTools import MultRamp, StaticRamp
hiTempParams = []
# Settings for annealing stage.
rampedParams = []

# Set up NOE potential.
from noePotTools import create_NOEPot

noe = create_NOEPot('noe', 'n15c13noe-soft.tbl')

potList.append(noe)
rampedParams.append(MultRamp(2.0,30.0, "noe.setScale( VALUE )"))

# Set up dihedral angle restraints
torsionFile='dihedral.tbl'
protocol.initDihedrals(torsionFile)
from xplorPot import XplorPot
potList.append(XplorPot('CDIH'))
#hiTempParams.append( StaticRamp("potList['CDIH'].setScale(10)") )
rampedParams.append( StaticRamp("potList['CDIH'].setScale(200)") )

# Statistical torsion angle potential
#
from torsionDBPotTools import create_TorsionDBPot
torsionDBPot = create_TorsionDBPot('tDB')
potList.append( torsionDBPot )
rampedParams.append( MultRamp(.002,2,"torsionDBPot.setScale(VALUE)") )


#covalent terms
from xplorPot import XplorPot
for term in ('BOND', 'ANGL', 'IMPR'):
   potList.append( XplorPot(term) )
   pass

# Set threshold for terms in potList to allow violation analysis.
potList['ANGL'].setThreshold(5.0)   # default is 2.0
potList['IMPR'].setThreshold(5.0)   # default is 2.0
# Use default values for the rest (bond: 0.05, cdih: 5.0, noe: 0.5).

rampedParams.append(MultRamp(0.4, 1.0, "potList['ANGL'].setScale(VALUE)"))
rampedParams.append(MultRamp(0.1, 1.0, "potList['IMPR'].setScale(VALUE)"))

#nonbonded term - EEFx
from eefxPotTools import create_EEFxPot, param_LK
eefxpot=create_EEFxPot("eefxpot","not name H*",paramSet=param_LK)
eefxpot.setVerbose(False)
potList.append(eefxpot)
rampedParams.append(MultRamp(0.1,1.0,"eefxpot.setScale(VALUE)"))
```

```
rampedParams.append(StaticRamp("potList['VDW'].setScale(0)"))

#use repel at high temp
potList.append( XplorPot("VDW") )
hiTempParams.append(StaticRamp("""protocol.initNBond(cutnb=100,
                                        rcon=0.04,
                                        tolerance=45,
                                        nbxmod=5,
                                        repel=1.2,
                                                        onlyCA=True)""""))
hiTempParams.append(StaticRamp("potList['VDW'].setScale(1)"))
hiTempParams.append(StaticRamp("eefxpot.setScale(0)"))

# Set up IVM object(s).

# IVM object for torsion-angle dynamics.
from ivm import IVM
dyn = IVM()
protocol.torsionTopology(dyn)

# IVM object for final Cartesian minimization.
minc = IVM()
protocol.cartesianTopology(minc)



temp_ini = 3500.0
temp_fin = 25.0


# Give atoms uniform weights, except for the anisotropy axis.
protocol.massSetup()


def calcOneStructure(loopInfo):
   """Calculate a structure.

   """
   # Generate an initial structure by randomizing torsion angles.
   from monteCarlo import randomizeTorsions
   randomizeTorsions(dyn)
   protocol.fixupCovalentGeom(maxIters=100, useVDW=1)

   #
   # High Temperature Dynamics Stage.
   #
   # Set torsion angles from restraints.
   # (They start satisfied, allowing the shortening of high temp dynamics.)
   from torsionTools import setTorsionsFromTable
   setTorsionsFromTable(torsionFile)

   # Initialize parameters for high temp dynamics.
```

```
from simulationTools import InitialParams
InitialParams(rampedParams) # overides some in rampedParams
InitialParams(hiTempParams) # overides some in rampedParams

#
# Torsion angle minimization.
#
protocol.initMinimize(dyn,
                potList=potList,
                numSteps=50,
                printInterval=10)
dyn.run()


protocol.initDynamics(dyn,
                potList=potList,
                bathTemp=temp_ini,
                initVelocities=1,
                finalTime=15,  # run for finalTsime or
                        # numSteps * 0.001, whichever is less
                numSteps=50 if quick else 15000,
                printInterval=100)


dyn.setETolerance(temp_ini/100) # used to det. stepsize. default: temp/1000
dyn.run()

#dynamics with all-atom repel potential
protocol.initDynamics(dyn,
                potList=potList,
                bathTemp=temp_ini,
                finalTime=40,  # run for finalTime or
                        # numSteps * 0.001, whichever is less
                numSteps=400 if quick else 800,
                printInterval=10)

protocol.initNBond()

dyn.setETolerance(temp_ini/100) # used to det. stepsize. default: temp/1000
dyn.run()



InitialParams(rampedParams)

#dynamics with EEFx potential
protocol.initDynamics(dyn,
                potList=potList,
                bathTemp=temp_ini,
                initVelocities=1,
                finalTime=30,  # run for finalTime or
                        # numSteps * 0.001, whichever is less
```

```
                numSteps=3000 if quick else 30000,
                printInterval=100)

dyn.setETolerance(temp_ini/100) # used to det. stepsize. default: temp/1000
dyn.run()


#
# Simulated Annealing Stage.
#
# Initialize parameters for annealing.
InitialParams(rampedParams)

# Set up IVM object for annealing.
protocol.initDynamics(dyn,
                potList=potList,
                finalTime=0.2,  # run for finalTime or
                        # numSteps * 0.001, whichever is less
                numSteps=20 if quick else 201,
                printInterval=100)

# Set up cooling loop and run.
from simulationTools import AnnealIVM
AnnealIVM(initTemp=temp_ini,
        finalTemp=temp_fin,
        tempStep=100 if quick else 12.5,
        ivm=dyn,
        rampedParams=rampedParams).run()

#
# Torsion angle minimization.
#
protocol.initMinimize(dyn,
                potList=potList,
                numSteps=500,
                printInterval=50)
if not quick: minc.run()

#
# Final Cartesian minimization.
#
protocol.initMinimize(minc,
                potList=potList,
                numSteps=500,
                dEPred=10)
if not quick: minc.run()


# Write structure.
```

```
from simulationTools import FinalParams
from simulationTools import StructureLoop
StructureLoop(numStructures=numberOfStructures,
        structLoopAction=calcOneStructure,
        doWriteStructures=True,
        averagePotList=potList,
        pdbTemplate=outfilename,
        genViolationStats=True,
        #averageCrossTerms=refRMSD,
        averageTopFraction=0.20, #report only on best 20% of structs
        averageContext=FinalParams(rampedParams),
        averageFilename="./fold_result/SCRIPT_ave.sa",   #generate regularized ave structure
        averageRegularize=True,
        averageRefineSteps=50
        ).run()
```

End of fold.py

Beginning of refine.py

```
import protocol

opts,args = xplor.parseArguments(["quick:0"])
quick=True if opts and opts[0][0] == "quick" else False

# Total number of structures - 100 structures from 10 starting structures
numberOfStructures=1 if quick else 100

# inTemplate= "fold_result/fold_##.pdb.best"


#inTemplate='fold' + '_*.pdb.best'

inTemplate="fold_result/fold_*.pdb"

# Base name for output pdb files.
outfilename = 'refine_result/SCRIPT_STRUCTURE.pdb'

# Initialize random seed.
protocol.initRandomSeed(3421)

protocol.parameters['protein']="protein_eef.par"
protocol.topology['protein']  ="protein_eef.top"
protocol.initParams("protein")

# load PDB to generate PSF
import glob
protocol.loadPDB( glob.glob(inTemplate)[0],
                deleteUnknownAtoms=True )

# Generate cis prolines
import psfGen
psfGen.cisPeptide(62)
psfGen.cisPeptide(69)
psfGen.cisPeptide(71)


#list to hold potential terms in structure calculation
from potList import PotList
potList = PotList()
crossTerms = PotList('cross')

#compute RMSD to reference structure
#  backbone and heavy atom RMSDs will be printed in the output
#  structure files
# #
# from posDiffPotTools import create_PosDiffPot
# refRMSD = create_PosDiffPot("refRMSD",
#                         "(name CA or name C or name N) and resid 1:230",
#                         pdbFile='1EZA.pdb',
```

```
#                       cmpSel="not name H*")


# lists to hold force constant settings for various stages of
# structure calculation
from simulationTools import MultRamp, StaticRamp
hiTempParams = []
# Settings for annealing stage.
rampedParams = []

# Set up NOE potential.
from noePotTools import create_NOEPot

noe = create_NOEPot('noe', 'n15c13noe-square.tbl')



potList.append(noe)
rampedParams.append(MultRamp(2.0,30.0, "noe.setScale( VALUE )"))

# Set up dihedral angles.
torsionFile='dihedral.tbl'
protocol.initDihedrals(torsionFile)
from xplorPot import XplorPot
potList.append(XplorPot('CDIH'))
hiTempParams.append( StaticRamp("potList['CDIH'].setScale(10)") )
rampedParams.append( StaticRamp("potList['CDIH'].setScale(200)") )

# Statistical torsion angle potential
#
from torsionDBPotTools import create_TorsionDBPot
torsionDBPot = create_TorsionDBPot('tDB')
potList.append( torsionDBPot )
rampedParams.append( MultRamp(.002,2,"torsionDBPot.setScale(VALUE)") )

#covalent terms
from xplorPot import XplorPot
for term in ('BOND', 'ANGL', 'IMPR'):
   potList.append( XplorPot(term) )
   pass

# Set threshold for terms in potList to allow violation analysis.
potList['ANGL'].setThreshold(5.0)   # default is 2.0
potList['IMPR'].setThreshold(5.0)   # default is 2.0
# Use default values for the rest (bond: 0.05, cdih: 5.0, noe: 0.5).

rampedParams.append(MultRamp(0.4, 1.0, "potList['ANGL'].setScale(VALUE)"))
rampedParams.append(MultRamp(0.1, 1.0, "potList['IMPR'].setScale(VALUE)"))

#nonbonded term
from eefxPotTools import create_EEFxPot, param_LK
eefxpot=create_EEFxPot("eefxpot","not name H*",paramSet=param_LK)
```

```
eefxpot.setVerbose(False)
potList.append(eefxpot)
rampedParams.append(MultRamp(0.1,1.0,"eefxpot.setScale(VALUE)"))

##use repel at high temp
#rampedParams.append(StaticRamp("potList['VDW'].setScale(0)"))
#potList.append( XplorPot("VDW") )
#hiTempParams.append(StaticRamp("""protocol.initNBond(cutnb=100,
#                                    rcon=0.04,
#                                    tolerance=0.5,
#                                    nbxmod=5,
#                                    repel=0.9,)"""))
#hiTempParams.append(StaticRamp("xplor.command('param nbonds rcon 1 end end')"))
#rampedParams.append(StaticRamp("potList['VDW'].setScale(1)"))
#hiTempParams.append(StaticRamp("eefxpot.setScale(0)"))


# Set up IVM object(s).

# IVM object for torsion-angle dynamics.
from ivm import IVM
dyn = IVM()
protocol.torsionTopology(dyn)

# IVM object for final Cartesian minimization.
minc = IVM()
protocol.cartesianTopology(minc)



temp_ini = 3000.0
temp_fin = 25.0


# Give atoms uniform weights, except for the anisotropy axis.
protocol.massSetup()


def calcOneStructure(loopInfo):
   """Calculate a structure.

   """
   # initial structure specified by inTemplate.
#    protocol.fixupCovalentGeom(maxIters=100, useVDW=1)


   # Initialize parameters for high temp dynamics.
   from simulationTools import InitialParams
   InitialParams(rampedParams) # overides some in rampedParams
   InitialParams(hiTempParams) # overides some in rampedParams

   #
```

```
# High Temperature Dynamics Stage.
#


protocol.initDynamics(dyn,
                potList=potList,
                bathTemp=temp_ini,
                initVelocities=True,
                finalTime=15,  # run for finalTime or
                        # numSteps * 0.001, whichever is less
                numSteps=50 if quick else 15000,
                printInterval=100)


dyn.setETolerance(temp_ini/100) # used to det. stepsize. default: temp/1000
dyn.run()


#
# Simulated Annealing Stage.
#
# Initialize parameters for annealing.
InitialParams(rampedParams)

# Set up IVM object for annealing.
protocol.initDynamics(dyn,
                potList=potList,
                finalTime=0.63,  # run for finalTime or
                        # numSteps * 0.001, whichever is less
                numSteps=2 if quick else 631,
                printInterval=100)

# Set up cooling loop and run.
from simulationTools import AnnealIVM
AnnealIVM(initTemp=temp_ini,
        finalTemp=temp_fin,
        tempStep=1000 if quick else 12.5,
        ivm=dyn,
        rampedParams=rampedParams).run()

#
# Torsion angle minimization.
#
protocol.initMinimize(dyn,
                potList=potList,
                numSteps=500,
                printInterval=50)
if not quick: minc.run()

#
# Final Cartesian minimization.
#
```

```
    protocol.initMinimize(minc,
                          potList=potList,
                          numSteps=500,
                          dEPred=10)
    if not quick: minc.run()

    # Write structure.
    return



from simulationTools import FinalParams, StructureLoop
StructureLoop(#numStructures=numberOfStructures,
              structLoopAction=calcOneStructure,
              pdbFilesIn=inTemplate,
              doWriteStructures=True,
              averagePotList=potList,
              pdbTemplate=outfilename,
              genViolationStats=True,
              # averageCrossTerms=refRMSD,
              averageTopFraction=0.20, #report only on best 20% of structs
              averageContext=FinalParams(rampedParams),
              averageFilename="refine_result/SCRIPT_ave.sa",   #generate regularized ave structure
              averageRegularize=True,
              averageRefineSteps=50
              ).run()
```

End of refine.py

Beginning of refine_eefx.py

```
xplor.requireVersion("2.18")

#
# slow cooling protocol in torsion angle space for protein G. Uses
# NOE, RDC, J-coupling restraints.
#
# this version refines from a reasonable model structure.
#
# CDS 2005/05/10
#

#inputStructures='refine' + '_*.pdb.best'
inputStructures="./refine_result/refine_*.pdb"
outfilename = 'refine_eefx_result/SCRIPT_STRUCTURE.pdb'


opts,args = xplor.parseArguments(["quick:0"])
quick=True if opts and opts[0][0] == "quick" else False

# Total number of structures - 100 structures from 10 starting structures
numberOfStructures=1 if quick else 100


# protocol module has many high-level helper functions.
#
import protocol
protocol.initRandomSeed(3421)   #explicitly set random seed

protocol.parameters['protein']="protein_eef.par"
protocol.topology['protein']  ="protein_eef.top"
protocol.initParams("protein")

#
# annealing settings
#

command = xplor.command

###############################################################################

# ## or read an existing model ##
#
# protocol.loadPDB("./refine_result/refine_ave.pdb")
# # xplor.simulation.deleteAtoms("not known")

###############################################################################

## load PDB to generate PSF  ##
import glob
protocol.loadPDB( glob.glob(inputStructures)[0],
```

```
            deleteUnknownAtoms=True )


# Generate cis prolines
import psfGen
psfGen.cisPeptide(62)
psfGen.cisPeptide(69)
psfGen.cisPeptide(71)


# a PotList contains a list of potential terms. This is used to specify which
# terms are active during refinement.


from potList import PotList
potList = PotList()

# parameters to ramp up during the simulated annealing protocol
#
from simulationTools import MultRamp, StaticRamp, InitialParams

rampedParams=[]
highTempParams=[]

# compare atomic Cartesian rmsd with a reference structure
#  backbone and heavy atom RMSDs will be printed in the output
#  structure files
#
# from posDiffPotTools import create_PosDiffPot
# refRMSD = create_PosDiffPot("refRMSD","name CA or name C or name N",
#                         pdbFile='g_xray.pdb',
#                         cmpSel="not name H*")

# orientation Tensor - used with the dipolar coupling term
#  one for each medium
#   For each medium, specify a name, and initial values of Da, Rh.
#
from varTensorTools import create_VarTensor
media={}
#                    medium  Da   rhombicity
for (medium,Da,Rh) in [ ('b',   13.241, 0.158)]:
   oTensor = create_VarTensor(medium)
   oTensor.setDa(Da)
   oTensor.setRh(Rh)
   media[medium] = oTensor
   pass

   ############################################################
# for (medium,Da,Rh) in [ ('a',   13.241, 0.158),
#                     ('b',   -9.9, 0.23),
#                     ('c',   -9.9, 0.23)]:
#    oTensor = create_VarTensor(medium)
```

```
#     oTensor.setDa(Da)
#     oTensor.setRh(Rh)
#     media[medium] = oTensor
#     pass
   ############################################################

# dipolar coupling restraints for protein amide NH.
#
# collect all RDCs in the rdcs PotList
#
# RDC scaling. Three possible contributions.
#   1) gamma_A * gamma_B / r_AB^3 prefactor. So that the same Da can be used
#      for different expts. in the same medium. Sometimes the data is
#      prescaled so that this is not needed. scale_toNH() is used for this.
#      Note that if the expt. data has been prescaled, the values for rdc rmsd
#      reported in the output will relative to the scaled values- not the expt.
#      values.
#   2) expt. error scaling. Used here. A scale factor equal to 1/err^2
#      (relative to that for NH) is used.
#   3) sometimes the reciprocal of the Da^2 is used if there is a large
#      spread in Da values. Not used here.
#
############################################################
# this is for steric alignment media
#
# from sardcPotTools import create_SARDCPot  #, scale_toNH (using this has error)
# rdcs = PotList('rdc')
# for (medium,expt,file,                scale) in \
#     [('b','NHN' ,'DNHN_TB_gel.tbl' ,1),
#      ('b','HACA' ,'DHACA_TB_gel.tbl' ,2.077),
#      ('b','CAC' ,'DCAC_TB_gel.tbl' ,0.198)
#      ]:
#     rdc = create_SARDCPot("%s_%s"%(medium,expt),file,media[medium])
#
#     #1) scale prefactor relative to NH
#     #   see python/rdcPotTools.py for exact calculation
#     # scale_toNH(rdc) - not needed for these datasets -
#     #                       but non-NH reported rmsd values will be wrong.
#
#     #3) Da rescaling factor (separate multiplicative factor)
#     # scale *= ( 1. / rdc.oTensor.Da(0) )**2
#     rdc.setScale(scale)
#     rdc.setShowAllRestraints(1) #all restraints are printed during analysis
#     rdc.setThreshold(2.0)       # in Hz
#     rdcs.append(rdc)
#     pass
# potList.append(rdcs)
# #rampedParams.append( MultRamp(0.05,5.0, "rdcs.setScale( VALUE )") )
# if RDCs are overfitted, reduce weights
# rampedParams.append( MultRamp(0.0005,0.5, "rdcs.setScale( VALUE )") )

############################################################
```

```
############################################################
#This is for other RDC medium

from rdcPotTools import create_RDCPot, scale_toNH
rdcs = PotList('rdc')
for (medium,expt,file,                scale) in \
   [('b','NHN' ,'DN-HN_TB_gel.tbl' ,1),
   ('b','HACA' ,'DH-CA_TB_gel_complete_v2.tbl' ,0.315),
   ('b','CACO' ,'DCA-C_TB_gel.tbl' ,0.943)
   ]:
   rdc = create_RDCPot("%s_%s"%(medium,expt),file,media[medium])
   scale_toNH(rdc)

    #1) scale prefactor relative to NH
    #   see python/rdcPotTools.py for exact calculation
    # scale_toNH(rdc) - not needed for these datasets -
    #                   but non-NH reported rmsd values will be wrong.

    #3) Da rescaling factor (separate multiplicative factor)
    # scale = inverse of experimental error relative to HN exp
    #       to calculate exp error, read bax: dipolar coupling in macromolecular structure
determination

   rdc.setScale(scale)
   rdc.setShowAllRestraints(1) #all restraints are printed during analysis
   rdc.setThreshold(3.0)       # in Hz
   rdcs.append(rdc)
   pass
potList.append(rdcs)
rampedParams.append( MultRamp(0.0005,0.5, "rdcs.setScale( VALUE )") )
############################################################


# calc. initial tensor orientation
# and setup tensor calculation during simulated annealing
#
from varTensorTools import calcTensorOrientation, calcTensor
for medium in media.keys():
   calcTensorOrientation(media[medium])
   rampedParams.append( StaticRamp("calcTensor(media['%s'])" % medium) )
   pass

# set up NOE potential
noe=PotList('noe')
potList.append(noe)
from noePotTools import create_NOEPot
for (name,scale,file) in [('noe',1,"n15c13noe-square.tbl"),
                    #add entries for additional tables
                    ]:
   pot = create_NOEPot(name,file)
```

```
   pot.setPotType("hard") # if you think there may be bad NOEs
   pot.setScale(scale)
   noe.append(pot)
rampedParams.append( MultRamp(2,30, "noe.setScale( VALUE )") )


################################################################
# # set up J coupling - with Karplus coefficients
# from jCoupPotTools import create_JCoupPot
# jCoup = create_JCoupPot("jcoup","jna_coup.tbl",
#                      A=6.98,B=-1.38,C=1.72,phase=-60.0)
# potList.append(jCoup)
################################################################

# Set up dihedral angles
from xplorPot import XplorPot
protocol.initDihedrals("dihedral.tbl",
                  #useDefaults=False  # by default, symmetric sidechain
                                      # restraints are included
                  )
potList.append( XplorPot('CDIH') )
highTempParams.append( StaticRamp("potList['CDIH'].setScale(300)") )
rampedParams.append( StaticRamp("potList['CDIH'].setScale(300)") )

# set custom values of threshold values for violation calculation
#
potList['CDIH'].setThreshold(5.0) #5 degrees is the default value, though


################################################################
#############roger take out for now##############
#############gyration volume term

#
# from gyrPotTools import create_GyrPot
# gyr = create_GyrPot("Vgyr",
#                  "resid 8:97") # selection should exclude disordered tails
# potList.append(gyr)
# rampedParams.append( MultRamp(.002,1,"gyr.setScale(VALUE)") )
################################################################

################################################################
# # hbda - distance/angle bb hbond term
# #
# protocol.initHBDA('hbda.tbl')
# potList.append( XplorPot('HBDA') )
################################################################

# Statistical torsion angle potential
#
from torsionDBPotTools import create_TorsionDBPot
torsionDBPot = create_TorsionDBPot('tDB')
```

```
potList.append( torsionDBPot )
#rampedParams.append( MultRamp(2,2,"torsionDBPot.setScale(VALUE)") )
rampedParams.append( MultRamp(.002,2,"torsionDBPot.setScale(VALUE)") )

#nonbonded term
from eefxPotTools import create_EEFxPot, param_LK
#eefxpot=create_EEFxPot("eefxpot","not name H* and resid 7:95",paramSet=param_LK)
eefxpot=create_EEFxPot("eefxpot","not name H*",paramSet=param_LK)
eefxpot.setVerbose(False)
potList.append(eefxpot)
rampedParams.append(MultRamp(0.1,1.0,"eefxpot.setScale(VALUE)"))

#################################################################
#############roger take out for now##############

# rampedParams.append(StaticRamp("potList['VDW'].setScale(0)"))
#
# #use repel at high temp
# potList.append( XplorPot("VDW") )
# highTempParams.append(StaticRamp("""protocol.initNBond(cutnb=100,
#                                     rcon=0.04,
#                                     tolerance=0.5,
#                                     nbxmod=5,
#                                     repel=0.9,)"""))
# highTempParams.append(StaticRamp("xplor.command('param nbonds rcon 1 end end')"))
# rampedParams.append(StaticRamp("potList['VDW'].setScale(1)"))
# highTempParams.append(StaticRamp("eefxpot.setScale(0)"))
#################################################################


potList.append( XplorPot("BOND") )
potList.append( XplorPot("ANGL") )
potList['ANGL'].setThreshold(5.0)
rampedParams.append( MultRamp(0.4,1,"potList['ANGL'].setScale(VALUE)") )
potList.append( XplorPot("IMPR") )
potList['IMPR'].setThreshold(5.0)
rampedParams.append( MultRamp(0.1,1,"potList['IMPR'].setScale(VALUE)") )



# Give atoms uniform weights, except for the anisotropy axis
#
protocol.massSetup()


# IVM setup
#   the IVM is used for performing dynamics and minimization in torsion-angle
#   space, and in Cartesian space.
#
from ivm import IVM
dyn = IVM()
```

```
###############################################################################
# initially minimize in Cartesian space with only the covalent constraints.
#   Note that bonds, angles and many impropers can't change with the
#   internal torsion-angle dynamics
#   breaks bonds topologically - doesn't change force field
#
#dyn.potList().add( XplorPot("BOND") )
#dyn.potList().add( XplorPot("ANGL") )
#dyn.potList().add( XplorPot("IMPR") )
#
#dyn.breakAllBondsIn("not resname ANI")
#import varTensorTools
#for m in media.values():
#    m.setFreedom("fix")               #fix tensor parameters
#    varTensorTools.topologySetup(dyn,m) #setup tensor topology
#
#protocol.initMinimize(dyn,numSteps=1000)
#dyn.run()
###############################################################################


# reset ivm topology for torsion-angle dynamics
#

dyn.reset()

for m in media.values():
#    m.setFreedom("fixDa, fixRh")        #fix tensor Rh, Da, vary orientation
   m.setFreedom("varyDa, varyRh")      #vary tensor Rh, Da, vary orientation
protocol.torsionTopology(dyn)

# minc used for final cartesian minimization
#
minc = IVM()
protocol.initMinimize(minc)

for m in media.values():
   m.setFreedom("varyDa, varyRh")    #allow all tensor parameters float here
   pass
protocol.cartesianTopology(minc)



# object which performs simulated annealing
#
from simulationTools import AnnealIVM
init_t  = 200     # Need high temp and slow annealing to converge
cool = AnnealIVM(initTemp=init_t,
                 finalTemp=20,
                 tempStep=10,
                 ivm=dyn,
                 rampedParams=rampedParams)
```

```
def accept(potList):
    """
    return True if current structure meets acceptance criteria
    """
    if potList['noe'].violations()>5:
        return False
    if potList['rdc'].rms()>5.0: #this might be tightened some
        return False
    if potList['CDIH'].violations()>5:
        return False
    if potList['BOND'].violations()>5:
        return False
    if potList['ANGL'].violations()>5:
        return False
    if potList['IMPR'].violations()>5:
        return False

    return True

def calcOneStructure(loopInfo):
    """ this function calculates a single structure, performs analysis on the
    structure, and then writes out a pdb file, with remarks.
    """

    # Initialize parameters for high temp dynamics.
    from simulationTools import InitialParams
    InitialParams(rampedParams) # overides some in rampedParams
    InitialParams(highTempParams) # overides some in rampedParams

    #
    # High Temperature Dynamics Stage.
    #
    protocol.initDynamics(dyn,
                          potList=potList, # potential terms to use
                          bathTemp=init_t,
                          initVelocities=True,
                          finalTime=15,  # run for finalTime or
                                  # numSteps * 0.001, whichever is less
                          numSteps=50 if quick else 15000,
                          printInterval=100)

    dyn.setETolerance( init_t/10 )  #used to det. stepsize. default: t/1000
    dyn.run()

    # initialize parameters for cooling loop
    InitialParams( rampedParams )


    # initialize integrator for simulated annealing
    #
    protocol.initDynamics(dyn,
```

```
                potList=potList,
                finalTime=0.63,  # run for finalTime or
                        # numSteps * 0.001, whichever is less
                numSteps=2 if quick else 631,
                printInterval=100)

    # perform simulated annealing
    #
    cool.run()


    #
    # Torsion angle minimization.
    #
    protocol.initMinimize(dyn,
                potList=potList,
                numSteps=500,
                printInterval=50)
    dyn.run()

    #
    # Final Cartesian minimization.
    #
    protocol.initMinimize(minc,
                potList=potList,
                numSteps=500,
                dEPred=10)
    minc.run()

    #do analysis and write structure when this routine is finished.
    pass

from simulationTools import StructureLoop, FinalParams
StructureLoop(#numStructures=numberOfStructures,
        structLoopAction=calcOneStructure,
        pdbFilesIn=inputStructures,
        doWriteStructures=True,  #analyze and write coords after calc
        genViolationStats=True,
        averagePotList=potList,
        averageSortPots=[potList['BOND'],potList['ANGL'],potList['IMPR'],
                noe,rdcs,potList['CDIH']],
        # averageCrossTerms=refRMSD,
#         averageTopFraction=0.5, #report only on best 50% of structs
        averageTopFraction=0.2, #report only on best 20% of structs
#         averageAccept=accept,   #only use structures which pass accept()
        averageContext=FinalParams(rampedParams),
        pdbTemplate=outfilename,
        averageFilename="refine_eefx_result/SCRIPT_ave.sa",    #generate regularized ave
structure
        averageRegularize=True,
        averageRefineSteps=50,
        averageFitSel="name CA",
```

```
    averageCompSel="not resname ANI and not name H*"
    ).run()
```

End of refine_eefx.py

Beginning of refine_eefx_slow.py

```
xplor.requireVersion("2.18")

#
# slow cooling protocol in torsion angle space for protein G. Uses
# NOE, RDC, J-coupling restraints.
#
# this version refines from a reasonable model structure.
#
# CDS 2005/05/10
#

#inputStructures='refine_eefx' + '_*.pdb.best'
inputStructures="./refine_eefx_result/refine_eefx_*.pdb"

outfilename = 'refine_eefx_slow_result/SCRIPT_STRUCTURE.pdb'


opts,args = xplor.parseArguments(["quick:0"])
quick=True if opts and opts[0][0] == "quick" else False

# Total number of structures - 100 structures from 10 starting structures
numberOfStructures=1 if quick else 100


# protocol module has many high-level helper functions.
#
import protocol
protocol.initRandomSeed(3421)   #explicitly set random seed

protocol.parameters['protein']="protein_eef.par"
protocol.topology['protein']  ="protein_eef.top"
protocol.initParams("protein")

#
# annealing settings
#

command = xplor.command

###############################################################################

# ## or read an existing model ##
#
# protocol.loadPDB("./refine_result/refine_ave.pdb")
# # xplor.simulation.deleteAtoms("not known")

###############################################################################

## load PDB to generate PSF  ##
import glob
```

```
protocol.loadPDB( glob.glob(inputStructures)[0],
                deleteUnknownAtoms=True )


# Generate cis prolines
import psfGen
psfGen.cisPeptide(62)
psfGen.cisPeptide(69)
psfGen.cisPeptide(71)


# a PotList contains a list of potential terms. This is used to specify which
# terms are active during refinement.


from potList import PotList
potList = PotList()

# parameters to ramp up during the simulated annealing protocol
#
from simulationTools import MultRamp, StaticRamp, InitialParams

rampedParams=[]
highTempParams=[]

# compare atomic Cartesian rmsd with a reference structure
#  backbone and heavy atom RMSDs will be printed in the output
#  structure files
#
# from posDiffPotTools import create_PosDiffPot
# refRMSD = create_PosDiffPot("refRMSD","name CA or name C or name N",
#                         pdbFile='g_xray.pdb',
#                         cmpSel="not name H*")

# orientation Tensor - used with the dipolar coupling term
#  one for each medium
#   For each medium, specify a name, and initial values of Da, Rh.
#
from varTensorTools import create_VarTensor
media={}
#                     medium  Da   rhombicity
for (medium,Da,Rh) in [ ('b',   13.241, 0.158)]:
   oTensor = create_VarTensor(medium)
   oTensor.setDa(Da)
   oTensor.setRh(Rh)
   media[medium] = oTensor
   pass

   ############################################################
# for (medium,Da,Rh) in [ ('a',   13.241, 0.158),
#                     ('b',   -9.9, 0.23),
#                     ('c',   -9.9, 0.23)]:
```

```
#     oTensor = create_VarTensor(medium)
#     oTensor.setDa(Da)
#     oTensor.setRh(Rh)
#     media[medium] = oTensor
#     pass
   ############################################################

# dipolar coupling restraints for protein amide NH.
#
# collect all RDCs in the rdcs PotList
#
# RDC scaling. Three possible contributions.
#   1) gamma_A * gamma_B / r_AB^3 prefactor. So that the same Da can be used
#      for different expts. in the same medium. Sometimes the data is
#      prescaled so that this is not needed. scale_toNH() is used for this.
#      Note that if the expt. data has been prescaled, the values for rdc rmsd
#      reported in the output will relative to the scaled values- not the expt.
#      values.
#   2) expt. error scaling. Used here. A scale factor equal to 1/err^2
#      (relative to that for NH) is used.
#   3) sometimes the reciprocal of the Da^2 is used if there is a large
#      spread in Da values. Not used here.
#
############################################################
# this is for steric alignment media
#
# from sardcPotTools import create_SARDCPot  #, scale_toNH (using this has error)
# rdcs = PotList('rdc')
# for (medium,expt,file,                scale) in \
#     [('b','NHN' ,'DNHN_TB_gel.tbl' ,1),
#      ('b','HACA' ,'DHACA_TB_gel.tbl' ,2.077),
#      ('b','CAC' ,'DCAC_TB_gel.tbl' ,0.198)
#      ]:
#     rdc = create_SARDCPot("%s_%s"%(medium,expt),file,media[medium])
#
#     #1) scale prefactor relative to NH
#     #   see python/rdcPotTools.py for exact calculation
#     # scale_toNH(rdc) - not needed for these datasets -
#     #                       but non-NH reported rmsd values will be wrong.
#
#     #3) Da rescaling factor (separate multiplicative factor)
#     # scale *= ( 1. / rdc.oTensor.Da(0) )**2
#     rdc.setScale(scale)
#     rdc.setShowAllRestraints(1) #all restraints are printed during analysis
#     rdc.setThreshold(2.0)       # in Hz
#     rdcs.append(rdc)
#     pass
# potList.append(rdcs)
# #rampedParams.append( MultRamp(0.05,5.0, "rdcs.setScale( VALUE )") )
# if RDCs are overfitted, reduce weights
# rampedParams.append( MultRamp(0.0005,0.5, "rdcs.setScale( VALUE )") )
```

```
#############################################################


#############################################################
#This is for other RDC medium

from rdcPotTools import create_RDCPot, scale_toNH
rdcs = PotList('rdc')
for (medium,expt,file,                scale) in \
   [('b','NHN' ,'DN-HN_TB_gel.tbl' ,1),
   ('b','HACA' ,'DH-CA_TB_gel_complete_v2.tbl' ,0.315),
   ('b','CACO' ,'DCA-C_TB_gel.tbl' ,0.943)
   ]:
   rdc = create_RDCPot("%s_%s"%(medium,expt),file,media[medium])
   scale_toNH(rdc)

    #1) scale prefactor relative to NH
    #   see python/rdcPotTools.py for exact calculation
    # scale_toNH(rdc) - not needed for these datasets -
    #                   but non-NH reported rmsd values will be wrong.

    #3) Da rescaling factor (separate multiplicative factor)
    # scale = inverse of experimental error relative to HN exp
    #       to calculate exp error, read bax: dipolar coupling in macromolecular structure
determination

   rdc.setScale(scale)
   rdc.setShowAllRestraints(1) #all restraints are printed during analysis
   rdc.setThreshold(3.0)       # in Hz
   rdcs.append(rdc)
   pass
potList.append(rdcs)
rampedParams.append( MultRamp(0.2,0.2, "rdcs.setScale( VALUE )") )
#############################################################


# calc. initial tensor orientation
# and setup tensor calculation during simulated annealing
#
from varTensorTools import calcTensorOrientation, calcTensor
for medium in media.keys():
   calcTensorOrientation(media[medium])
   rampedParams.append( StaticRamp("calcTensor(media['%s'])" % medium) )
   pass

# set up NOE potential
noe=PotList('noe')
potList.append(noe)
from noePotTools import create_NOEPot
for (name,scale,file) in [('noe',1,"n15c13noe-square.tbl"),
                    #add entries for additional tables
                    ]:
```

```
    pot = create_NOEPot(name,file)
    pot.setPotType("hard") # if you think there may be bad NOEs
    pot.setScale(scale)
    noe.append(pot)
rampedParams.append( MultRamp(30,30, "noe.setScale( VALUE )") )


################################################################
# # set up J coupling - with Karplus coefficients
# from jCoupPotTools import create_JCoupPot
# jCoup = create_JCoupPot("jcoup","jna_coup.tbl",
#                         A=6.98,B=-1.38,C=1.72,phase=-60.0)
# potList.append(jCoup)
################################################################

# Set up dihedral angles
from xplorPot import XplorPot
protocol.initDihedrals("dihedral.tbl",
                       #useDefaults=False  # by default, symmetric sidechain
                                           # restraints are included
                       )
potList.append( XplorPot('CDIH') )
highTempParams.append( MultRamp(300,200,"potList['CDIH'].setScale(VALUE)") )
rampedParams.append( MultRamp(300,200,"potList['CDIH'].setScale(VALUE)") )

#highTempParams.append( StaticRamp("potList['CDIH'].setScale(300)") )
#rampedParams.append( StaticRamp("potList['CDIH'].setScale(300)") )


# set custom values of threshold values for violation calculation
#
potList['CDIH'].setThreshold(5.0) #5 degrees is the default value, though


################################################################
#############roger take out for now#############
#############gyration volume term

#
# from gyrPotTools import create_GyrPot
# gyr = create_GyrPot("Vgyr",
#                     "resid 8:97") # selection should exclude disordered tails
# potList.append(gyr)
# rampedParams.append( MultRamp(.002,1,"gyr.setScale(VALUE)") )
################################################################

################################################################
# # hbda - distance/angle bb hbond term
# #
# protocol.initHBDA('hbda.tbl')
# potList.append( XplorPot('HBDA') )
################################################################
```

```
# Statistical torsion angle potential
#
from torsionDBPotTools import create_TorsionDBPot
torsionDBPot = create_TorsionDBPot('tDB')
potList.append( torsionDBPot )
rampedParams.append( MultRamp(2,2,"torsionDBPot.setScale(VALUE)") )


#nonbonded term
from eefxPotTools import create_EEFxPot, param_LK
#eefxpot=create_EEFxPot("eefxpot","not name H* and resid 7:95",paramSet=param_LK)
eefxpot=create_EEFxPot("eefxpot","not name H*",paramSet=param_LK)
eefxpot.setVerbose(False)
potList.append(eefxpot)
rampedParams.append(MultRamp(1.0,1.0,"eefxpot.setScale(VALUE)"))

#################################################################
#############roger take out for now##############

# rampedParams.append(StaticRamp("potList['VDW'].setScale(0)"))
#
# #use repel at high temp
# potList.append( XplorPot("VDW") )
# highTempParams.append(StaticRamp("""protocol.initNBond(cutnb=100,
#                                     rcon=0.04,
#                                     tolerance=0.5,
#                                     nbxmod=5,
#                                     repel=0.9,)"""))
# highTempParams.append(StaticRamp("xplor.command('param nbonds rcon 1 end end')"))
# rampedParams.append(StaticRamp("potList['VDW'].setScale(1)"))
# highTempParams.append(StaticRamp("eefxpot.setScale(0)"))
#################################################################


potList.append( XplorPot("BOND") )
potList.append( XplorPot("ANGL") )
potList['ANGL'].setThreshold(5.0)
rampedParams.append( MultRamp(1,1,"potList['ANGL'].setScale(VALUE)") )
potList.append( XplorPot("IMPR") )
potList['IMPR'].setThreshold(5.0)
rampedParams.append( MultRamp(1,1,"potList['IMPR'].setScale(VALUE)") )



# Give atoms uniform weights, except for the anisotropy axis
#
protocol.massSetup()


# IVM setup
#   the IVM is used for performing dynamics and minimization in torsion-angle
```

```
#   space, and in Cartesian space.
#
from ivm import IVM
dyn = IVM()

################################################################################
# initially minimize in Cartesian space with only the covalent constraints.
#   Note that bonds, angles and many impropers can't change with the
#   internal torsion-angle dynamics
#   breaks bonds topologically - doesn't change force field
#
#dyn.potList().add( XplorPot("BOND") )
#dyn.potList().add( XplorPot("ANGL") )
#dyn.potList().add( XplorPot("IMPR") )
#
#dyn.breakAllBondsIn("not resname ANI")
#import varTensorTools
#for m in media.values():
#    m.setFreedom("fix")               #fix tensor parameters
#    varTensorTools.topologySetup(dyn,m) #setup tensor topology
#
#protocol.initMinimize(dyn,numSteps=1000)
#dyn.run()
################################################################################


# reset ivm topology for torsion-angle dynamics
#

dyn.reset()

for m in media.values():
#    m.setFreedom("fixDa, fixRh")       #fix tensor Rh, Da, vary orientation
   m.setFreedom("varyDa, varyRh")      #vary tensor Rh, Da, vary orientation
protocol.torsionTopology(dyn)

# minc used for final cartesian minimization
#
minc = IVM()
protocol.initMinimize(minc)

for m in media.values():
   m.setFreedom("varyDa, varyRh")    #allow all tensor parameters float here
   pass
protocol.cartesianTopology(minc)



# object which performs simulated annealing
#
from simulationTools import AnnealIVM
init_t  = 20     # Need high temp and slow annealing to converge
```

```
cool = AnnealIVM(initTemp=init_t,
                 finalTemp=1,
                 tempStep=1,
                 ivm=dyn,
                 rampedParams=rampedParams)

def accept(potList):
    """
    return True if current structure meets acceptance criteria
    """
    if potList['noe'].violations()>5:
        return False
    if potList['rdc'].rms()>5.0: #this might be tightened some
        return False
    if potList['CDIH'].violations()>5:
        return False
    if potList['BOND'].violations()>5:
        return False
    if potList['ANGL'].violations()>5:
        return False
    if potList['IMPR'].violations()>5:
        return False

    return True

def calcOneStructure(loopInfo):
    """ this function calculates a single structure, performs analysis on the
    structure, and then writes out a pdb file, with remarks.
    """

    # Initialize parameters for high temp dynamics.
    from simulationTools import InitialParams
    InitialParams(rampedParams) # overides some in rampedParams
    InitialParams(highTempParams) # overides some in rampedParams

    #
    # High Temperature Dynamics Stage.
    #
    protocol.initDynamics(dyn,
                          potList=potList, # potential terms to use
                          bathTemp=init_t,
                          initVelocities=True,
                          finalTime=15,  # run for finalTime or
                                  # numSteps * 0.001, whichever is less
                          numSteps=50 if quick else 15000,
                          printInterval=100)

    dyn.setETolerance( init_t/1 )  #used to det. stepsize. default: t/1000
    dyn.run()

    # initialize parameters for cooling loop
    InitialParams( rampedParams )
```

```
    # initialize integrator for simulated annealing
    #
    protocol.initDynamics(dyn,
                          potList=potList,
                          finalTime=0.63,  # run for finalTime or
                                           # numSteps * 0.001, whichever is less
                          numSteps=2 if quick else 631,
                          printInterval=100)

    # perform simulated annealing
    #
    cool.run()


    #
    # Torsion angle minimization.
    #
    protocol.initMinimize(dyn,
                          potList=potList,
                          numSteps=500,
                          printInterval=50)
    dyn.run()

    #
    # Final Cartesian minimization.
    #
    protocol.initMinimize(minc,
                          potList=potList,
                          numSteps=500,
                          dEPred=10)
    minc.run()

    #do analysis and write structure when this routine is finished.
    pass

from simulationTools import StructureLoop, FinalParams
StructureLoop(#numStructures=numberOfStructures,
              structLoopAction=calcOneStructure,
              pdbFilesIn=inputStructures,
              doWriteStructures=True,  #analyze and write coords after calc
              genViolationStats=True,
              averagePotList=potList,
              averageSortPots=[potList['BOND'],potList['ANGL'],potList['IMPR'],
                               noe,rdcs,potList['CDIH']],
              # averageCrossTerms=refRMSD,
#             averageTopFraction=0.5, #report only on best 50% of structs
#             averageTopFraction=0.2, #report only on best 20% of structs
              averageTopNum=20,
#             averageAccept=accept,   #only use structures which pass accept()
              averageContext=FinalParams(rampedParams),
```

```
        pdbTemplate=outfilename,
        averageFilename="refine_eefx_slow_result/SCRIPT_ave.sa",    #generate regularized
ave structure
        averageRegularize=True,
        averageRefineSteps=50,
        averageFitSel="name CA",
        averageCompSel="not resname ANI and not name H*"
        ).run()
```

End of refine_eefx_slow.py