A CYANOBACTERIAL CIRCADIAN TIMING MECHANISM

J.L. Ditty, S.B. Williams, and S.S. Golden

1 Department of Biology, University of St. Thomas, St. Paul, Minnesota 55105; email: jlditty@stthomas.edu; 2 Department of Biology, University of Utah, Salt Lake City, Utah 84112; email: williams@biology.utah.edu; 3 Department of Biology, Texas A&M University, College Station, Texas 77843; email: sgolden@tamu.edu

Key Words ATPase, biological clock, receiver domain, kai, signal transduction

Abstract Cyanobacteria such as Synechococcus elongatus PCC 7942 exhibit 24-h rhythms of gene expression that are controlled by an endogenous circadian clock that is mechanistically distinct from those described for diverse eukaryotes. Genetic and biochemical experiments over the past decade have identified key components of the circadian oscillator, input pathways that synchronize the clock with the daily environment, and output pathways that relay temporal information to downstream genes. The mechanism of the cyanobacterial circadian clock that is emerging is based principally on the assembly and disassembly of a large complex at whose heart are the proteins KaiA, KaiB, and KaiC. Signal transduction pathways that feed into and out of the clock employ protein domains that are similar to those in two-component regulatory systems of bacteria.

CONTENTS

INTRODUCTION .................................................... 514
THE CYANOBACTERIA ........................................... 514
THE CIRCADIAN CLOCK ........................................ 517
A CYANOBACTERIAL CIRCADIAN CLOCK ................... 518
THE SYNECHOCoccus ELONGATUS PCC 7942 CIRCADIAN CLOCK ................................................................. 520
The Model System: S. elongatus PCC 7942 ....................... 520
The S. elongatus PCC 7942 Clock Proteins: KaiA, KaiB, KaiC .......... 521
TRANSCRIPTIONAL CONTROL IN THE S. ELONGATUS TIMING MECHANISM ............................................. 524
KAI OSCILLATOR INPUT: CIKA, LDPA, AND PEX .............. 526
KAI OSCILLATOR OUTPUT: SASA, CPMA, AND GROUP TWO SIGMA FACTORS ................................................. 528

*These authors contributed equally.
INTRODUCTION

A consequence of Earth’s rotation about its axis once every 24 h is that the majority of organisms that populate her are subject to daily fluctuations in light and temperature. As a result, inhabitants of this planet, from bacteria to humans, have optimized their existence by evolving mechanisms to adjust to and anticipate daily changes in the environment (38, 130). Chronobiologists have developed various model systems to investigate circadian (24-h rhythmic) phenomena; this review focuses on the circadian mechanism that has evolved within the cyanobacteria, focusing on the model organism *Synechococcus elongatus* PCC 7942. A brief review of the cyanobacteria as a group and the broader implications of clock gene orthologs among these fascinating and widespread bacteria are discussed, as well as the present state of knowledge concerning the prokaryotic biochemical mechanism of timekeeping. In its revelation, the timing mechanism of the cyanobacterial circadian clock is unfolding as an intricate relationship between protein-protein interactions, protein modifications, and novel twists to traditional bacterial signal transduction pathways. Many players involved in every aspect of the *S. elongatus* circadian clock—from environmental input, to temporal output, and the oscillator itself—have been identified, but proposed models raise many questions that are still unresolved.

THE CYANOBACTERIA

Cyanobacteria are a fascinating and exceptionally diverse group of photoautotrophic prokaryotes. Their genetic lineage is evidently among the oldest on Earth, as fossilized cyanobacterium-like organisms are present in 3500 Ma-old conglomeratic Apex chert (17, 145, 146). Oxygenic photosynthesis originated in those antiquated cyanobacteria, and that distinctive metabolic activity was paramount to the creation of our present day oxygen-enriched atmosphere (80). There is remarkable genetic diversity among extant cyanobacteria as exemplified by comparing the mol% G + C content of genomes. For example, *Nostoc* species strain PCC 7524 has 39% G + C, *Synechococcus elongatus* (PCC 6301/7942) has 55% G + C, and *Cyanobium* species strain PCC 6707 has a genome with nearly 70 mol% G + C (158). The morphological diversity among cyanobacteria is also intriguing. Numerous species, including those within the *Aphanacapsa, Chroococcus, Merismopedia, Synechocystis*, and *Synechococcus* genera, grow as ovoid- or rod-shaped unicells ranging in diameter from 0.4 to 40 µm (173). Some of these unicellular species can live as single cells but also may remain, after cell division,
in tightly grouped cell aggregates (127, 129). Unicellular species likely regulate this lifestyle choice based upon prevailing environmental conditions (129). These cell aggregates often appear highly organized, perhaps reflecting an underlying social order (53, 127). Other cyanobacterial species, such as those from the genera *Anabaena*, *Lyngbya*, *Scytonema*, *Stigonema*, *Tolypothrix*, and *Trichodesmium*, are long, thin, multicellular filaments commonly surrounded by a mucilaginous sheath. They are typically about 10 µm in diameter and can be several hundred micrometers long (55, 149). Many filamentous species can form differentiated cells including hormagonia (motile fragment of a cyanobacterial filament), akinetes (resting cyanobacterial spores), and terminally differentiated cells called heterocysts, which develop under nitrogen-limited conditions and essentially function as anaerobic chambers for nitrogen fixation (46, 49, 180).

As would be expected from their long evolutionary history, genetic diversity, and morphological malleability, cyanobacteria are found in nearly every habitat that sunlight penetrates. Amazingly, a cave-dwelling *Gloeocapsa* species survives under light intensities as low as 1 lux (≈0.02 µmol photon m$^{-2}$ s$^{-1}$) (27). Thermophilic cyanobacterial species have been isolated from geothermal hot springs throughout the world and have maximum growth temperatures ranging from 50° to 74°C (1, 45, 113, 114, 121). Mesophilic species are ubiquitous and have been isolated from most dry land ecosystems, including karst and travertine regions. They also flourish in benthic, limnetic, lotic, and pelagic fresh- and saltwater habitats (6, 21, 28, 42, 99, 123, 124, 126, 136, 143). Aquatic species, including *Oscillatoria agardhii*, *Aphanizomenon flos-aquae*, and *Microcystis aeruginosa*, find “a place in the sun” by regulating their buoyancy and relative position in the water column via production of proteinaceous gas vesicles (7, 30, 31, 171). *Aphanothece halophytica*, *Dactylococcopsis salina*, *Microcoleus chthonoplastes*, and *Spirulina major*, among many other cyanobacterial species, are halotolerant if not true halophilic organisms. These or closely related species have been isolated from practically every known hypersaline environment (18, 32, 41). Psychrophilic species, such as *Nodularia harveyana*, *Phormidium frigidum*, and *Rivularia minutula*, are typically the predominant life forms in their low-temperature surroundings. Their seemingly inhospitable habitats include the tundra, ice shelves, glacial moraines, and polar desert soils of both the Arctic and the Antarctic regions (132, 148, 172). Both cold and hot desert cyanobacteria have an uncommon and astonishing ability to withstand multiple rounds of desiccation and subsequent rehydration (13, 131). As a general survival strategy in harsh environments, cyanobacteria have evolved to resourcefully create their own macroscopic, insular environments; these include such complex ecosystems as coastal tidal, hot spring and hypersaline microbial mats, the environmentally essential (and extremely sensitive) cryptobiotic desert crusts, and even fresh- and saltwater blooms (58, 115, 166).

The range of both growth rates and metabolic activities in cyanobacteria are also noteworthy. Typical freshwater *Synechococcus* species have doubling times of several hours, whereas cyanobacterial populations in the cold, oligotrophic,
The dry deserts of Antarctica may have doubling times of nearly 10,000 years (43, 116). Carbon dating in these polar regions supports one implication of this slow growth rate estimate by showing that living cyanobacterial cells can be over 1000 years old (16). Clearly, many interesting survival strategies have evolved in the cyanobacteria, and not all of them are passive. *Cylindrospermopsis raciborskii*, *Hapalosiphon fontinalis*, *Hormothamnion enteromorphoides*, *Umezakia natans*, and most of the aforementioned genera make a variety of cyanotoxins as secondary metabolites (8, 60, 61, 72, 90, 181). These toxic metabolites are species-specific alkaloids, macrolides, and short linear or cyclic peptides that can be cytotoxic, hepatotoxic, or even neurotoxic to many organisms including mammals (35).

The primary metabolic activities such as photosynthesis, carbon and nitrogen fixation, and de novo vitamin and cofactor biosynthesis that have allowed cyanobacteria to inhabit practically every environment have also made them common participants in symbiotic associations. Species of the *Calothrix*, *Cylindrospermum*, *Fischerella*, and *Nostoc* genera form endophytic, epiphytic, and true symbiotic relationships with numerous plants, fungi, sponges, and protists (26, 34, 54, 59, 76, 137, 162, 176). It is generally accepted that modern plastids (including the land plant chloroplasts) evolved from a free-living cyanobacterium after its sequestration by a primordial eukaryotic-like cell (23, 98). The success of this particular 1 to 2 billion-year-old endosymbiotic event was evidently unique, as all extant plastids are considered monophyletic (34). Primary plastids, those directly descended from that first cyanobiont, still reside in rhodophyte, chlorophyte, and glaucocystophyte algae (34). Evolutionary relationships among cyanobacteria and these plastids remain an intriguing and not well-resolved area of study (22, 23, 98, 107, 119, 134, 153, 163, 164).

Given their exceptional variety of form, faculty, and function—a spectrum that has only been hinted at in this short introduction—what biological properties help define the cyanobacteria? Cyanobacteria have an intracytoplasmic, thylakoid membrane used to house their photosynthesis machinery (77). If exception makes the rule, then consider *Gloeobacter violaceous* PCC 7421, whose photosystems are located within the cytoplasmic membrane (138). They also harbor both photosystem I and photosystem II and, thus, use water as a reductant during oxygenic photosynthesis (44, 183, 184). Cyanobacteria absorb light energy for photosynthesis by synthesizing and utilizing the chlorophyll *a* molecule, phycobiliproteins, and accessory phycobilin pigments like phycoerythrin, allophycocyanin, and phycocyanin (19, 103). High concentrations of these latter two pigments often make the organisms appear greenish-blue, leading to their previous designation as “blue-green algae” and current moniker of cyanobacteria. However, not all are blue-green in color, and the broadly distributed prochlorophyte species use chlorophylls *a* and *b* as antenna pigments and do not elaborate phycobilin antennae (83, 128). In addition to the accepted photosynthetic traits, we speculate here that the presence of circadian clock genes, and presumably a functional circadian clock, are also general cyanobacterial characteristics (96).
THE CIRCADIAN CLOCK

The circadian clock is an endogenous cellular mechanism that allows organisms to temporally regulate gene expression and thereby regulate complex biological processes as a function of time (38, 40, 130). Temporal organization of cellular functions allows organisms to capitalize on the environmental condition at any given time of day. General conceptual models for circadian systems involve three basic elements. The circadian oscillator is the entity directly responsible for keeping time on an approximately 24-h scale. Input pathways are the conduits for entrainment, a process that uses the daily reception of environmental stimuli to reset the circadian clock to local time every day. Output pathways couple timing information from the oscillator to downstream genes or other targets and allow the expression of subsequent circadian behaviors (Figure 1).

Three main characteristics define rhythms of behavior as a circadian process, as distinct from oscillations that are driven by environmental or specific metabolic cycles (Figure 1B). The first intrinsic characteristic is that circadian rhythms persist under constant environmental conditions, i.e., in the absence of a light/dark (LD), temperature, or humidity cycle, demonstrating an endogenous source for the rhythmicity. The persistent oscillation in the artificial constant condition is designated the free-run, and the period (peak-to-peak or trough-to-trough duration) of the free-running rhythm is characteristic of the timing mechanism of an individual organism’s intrinsic clock. Second, circadian systems must be sensitive to an environmental stimulus to achieve entrainment with the sidereal day. The periods of free-running circadian rhythms are rarely 24 h precisely in length; daily entrainment is imperative so as not to fall out of phase with the light and dark cycles of an exactly 24-h day. A final criterion that distinguishes a circadian process is that the timing mechanism maintains its intrinsic periodicity at different constant temperatures within the organism’s physiological range, a feature known as temperature compensation. This is characterized by the observation that at different ambient temperatures the rate of the timing mechanism varies only slightly, in contrast to simple biochemical reactions whose rates are greatly affected by changes in temperature (38, 40, 130).

These general characteristics of circadian rhythms are useful for chronobiologists in interpreting rhythmic observations. However, as rules, they are oversimplifications. Intrinsic periods vary depending on environmental conditions such as light intensity, as formalized in Aschoff’s Rule (38). In general, diurnal organisms have shorter free-running periods as light intensity increases, and the converse holds for nocturnal species (Figure 1C). Temperature compensation is often misinterpreted as temperature insensitivity. Circadian clocks are indeed sensitive to temperature via their input pathways. A temperature cycle can be as strong an entrainment cue as an LD cycle in some organisms (93). However, the biochemical mechanism of the oscillator accommodates changes in temperature, such that gross changes in period do not occur as would be observed for the rate of an enzymatic reaction over a similar range of degrees. Some buffering occurs to keep the clock...
within an entrainable range no matter which climate on earth the organism finds itself inhabiting.

A CYANOBACTERIAL CIRCADIAN CLOCK

Physiological studies regarding the temporal separation of cyanobacterial nitrogen fixation and oxygenic photosynthesis were among the first to suggest that cyanobacteria have endogenous timing mechanisms (50). Many filamentous (multicellular) diazotrophic cyanobacteria, such as *Anabaena* sp. strain PCC 7120, develop microaerobic heterocysts to spatially separate oxygen-sensitive nitrogen fixation from oxygen-evolving photosynthetic metabolism (79, 175). The discovery of temporal separation arose from the question of how unicellular species, lacking the option of differentiation, could balance these seemingly incompatible processes. During the mid-1980s, a body of evidence from many groups showed that cyanobacteria time the onset of nitrogen fixation; however, the reference point for the time schedule was ascribed to various cellular functions by different groups (50, 51, 56, 106, 150). Physical evidence of temporal metabolic regulation in bacteria confronted a long-standing belief that genuine circadian rhythms were a phenomenon restricted to eukaryotic organisms. Eukaryotic models for circadian rhythmicity were based upon intercellular communication and nuclear sequestration of key clock components. Furthermore, it was expected that an organism that could grow and divide at a rate faster than 24 h would not be able to maintain a circadian regulatory regimen (40, 84). This particular supposition was later shown to be unfounded (87, 107a).

The first convincing demonstrations of a prokaryotic circadian clock were published in 1989–1990. Sweeney & Borgese reported a temperature-compensated 24-h rhythm of cell division in *Synechococcus* sp. strain WH 7803 (155). Huang’s group showed that a rhythm in nitrogen fixation by *Synechococcus* sp. strain RF-1 (PCC 8801) fulfilled all three criteria for circadian control as outlined for eukaryotic circadian rhythms (69). Subsequently, they demonstrated circadian rhythms in amino acid uptake by strain RF-1 as well (24). Sherman’s group showed the alternation of nitrogen fixation and photosynthesis visually in a *Cyanothece* species, in which storage granules from the products of each process are visible by electron microscopy (144a). Determining the underlying genetic and molecular basis for cyanobacterial circadian rhythms required their demonstration in a cyanobacterial strain to which modern, molecular biological tools could be applied, and for which an assay suitable for repeated sampling and high throughput could be developed. *Synechococcus elongatus* PCC 7942 ultimately met those requirements (71, 88, 89).

Although somewhat alluring, prevailing notions that the circadian clock and its overt temporal rhythms evolved in cyanobacteria to separate oxygenic photosynthesis from oxygen-sensitive nitrogen fixation are, at best, not the complete story. *S. elongatus*, the unicellular, photoautotrophic model system for prokaryotic circadian biology, does not fix nitrogen (66). The marine, filamentous, and
non-heterocyst-forming genus *Trichodesmium* evidently uses both spatial and temporal separation of these metabolic activities (11). It appears that *Trichodesmium* balances its rates of aerobic respiration and photosynthetic oxygen production so that no net oxygen is produced during part of the photoperiod, which allows for nitrogen fixation.

None of the strategies described above for the separation of oxygen evolution and nitrogen fixation has an obvious requirement for circadian timekeeping. Nonetheless, the circadian clock provides a selective growth advantage in *S. elongatus* (125). Strains whose circadian oscillators generate different period lengths in gene-expression rhythms were put into direct competition with one another under differing LD cycles. Those strains with endogenous circadian period lengths that most closely matched the periodicity of the externally applied LD cycle prevailed in the competition. Remarkably, when two strains were first mixed in nearly equal proportions and then allowed to grow together, the time-matched strain completely “eliminated” its competitor in about seven generations (125). Note that the competing strains had indistinguishable growth rates when grown individually, and both could entrain to the periodicity of the applied LD cycle (125).

Results from these competition experiments suggest that there is a preset, temporal pace to clock control of cellular metabolism. Adjustments to that pace, to match perceived environmental changes, have an associated energetic and fitness cost. In a 15-h light/15-h dark cycle, an organism with a 30-h endogenous period outcompetes one with a 24-h endogenous period because, theoretically, it does not have to make costly, energetically demanding adjustments to its endogenous rhythmic metabolic progressions. Mathematical models based upon the above competition data have ruled out nutrient uptake kinetics as providing the competitive advantage (141). Another model is consistent with a poisoning mechanism, whereby the winning competitor secretes a growth inhibitor at a time when the losing competitor, because of its metabolic state, is not immune to it (52). Experimental data to support or exclude either model have not been collected.

Even though the outright incompatibility of nitrogen fixation and oxygen-evolving photosynthesis is not a sufficient justification for circadian rhythms, additional data suggest that metabolic separation of disparate processes is advantageous. The expression of *purF*, a gene that encodes another oxygen-sensitive enzyme, is phased differently from that of the bulk of genes, including the photosynthesis genes, in *S. elongatus* PCC 7942 (94). The same is true for *opcA*, whose product is important for the oxidative pentose phosphate pathway, the chief nonphotosynthetic means of reductant generation in this organism (105). Loss of *opcA* disproportionately affects reductant levels at dawn relative to dusk.

Our understanding of how the circadian clock controls cellular metabolism, and how extra- and intracellular environments impact the clock, is still very limited. The prokaryotic-clock research community has taken the prototypical reductive approach to determine the composition of the timing system and discern how it might function. Models based on empirical observations and testable hypotheses are emerging for the mechanism that underlies cyanobacterial timekeeping.
THE SYNECHOCOCCUS ELONGATUS PCC 7942 CIRCADIAN CLOCK

The Model System: S. elongatus PCC 7942

An authentic circadian clock has been confirmed in the cyanobacterium Synechococcus elongatus PCC 7942 (71, 88, 89). The identification of components of the circadian system in S. elongatus was made possible by the amazing malleability of the organism to molecular genetic techniques (3, 20). The S. elongatus clock is comprised of the products of at least three genes, kaiA, kaiB, and kaiC (Figure 2A), named after the Japanese word kaiten for a cycle or “turning of the heavens.” The Kai proteins are intrinsic to the machinery for generating and maintaining rhythms of gene expression that have a free-running period of approximately 24–25 h depending on incident light intensity. Point mutations in any kai gene can result in strains that have a variety of period phenotypes, and deletion of any kai gene results in an arrhythmic strain, while having no apparent effect upon growth rate (71, 117). The kai locus is expressed from two promoters—one upstream of kaiA (monocistronic message) and one upstream of kaiB (dicistronic kaiBC message)—that drive transcription in the same circadian phase in wild-type cells, with peak expression at dusk (71). Circadian rhythms under the control of the kai oscillator are easily tracked in S. elongatus by using a bioluminescent reporter generated from a fusion between any S. elongatus promoter and the luxAB genes of Vibrio harveyi or the luc gene from the firefly, which encode nonhomologous luciferase enzymes (2, 3, 50). Expression of either reporter gene results in light production as a function of promoter activity, which can be monitored by a photomultiplier tube or sensitive camera.

The pervasiveness of Kai circadian clock control in S. elongatus is unprecedented in any other circadian model system. Based on random insertion of promoterless luxAB into the S. elongatus genome, at least 800 genes of an estimated 2700 in the genome are expressed rhythmically in this organism (95). In fact, constitutive expression has never been detected for any promoter by luciferase monitoring in wild-type S. elongatus. Differences in amplitude, waveform, overall expression, and relative phasing of bioluminescence are promoter-specific, so that gene expression patterns can be categorized. The most pervasive, Class 1, rhythms have their circadian peak near subjective dusk, whereas Class 2 rhythms are antiphase, peaking at dawn (95) (Figure 2B). There are probably other minor classes, although the characterized genes fall into one of these two general patterns based on peak time. Mutations that differentially affect some Class 1 genes suggest that more than one output pathway connects various members of this large category to the clock (82, 112, 165) (Figure 1A).

The global transcriptional control that is revealed by bioluminescent reporting does not necessarily translate into global cycling of mRNA levels (95; N. Lebedeva & S.S. Golden, unpublished data). Under constant light conditions, many S. elongatus mRNAs have longer half-lives than the unstable reporter messages and
luciferase enzyme activity, so that troughs rise and rhythms of transcription are masked in mRNA accumulation levels. A similar phenomenon was described previously in the plant *Arabidopsis thaliana* (104). Thus, we cannot conclude that metabolism is globally circadian in *S. elongatus*, although there may be circadian components to most or all processes. Amino acid uptake, which is rhythmic in an LD cycle, does not continue to oscillate in LL in *S. elongatus*, although it does in *Synechococcus* sp. strain RF-1 (24). In the real world, however, organisms are always in a day/night cycle, and intrinsic circadian processes are intertwined with the “driven” processes that respond more directly to environmental stimuli.

The *S. elongatus* PCC 7942 Clock Proteins: KaiA, KaiB, KaiC

The *kai* locus was originally identified in a screen to complement a long period (44-h) mutant of *S. elongatus* PCC 7942 with a wild-type genomic library. The complementing locus restored wild-type rhythms, or otherwise modified the circadian phenotype, when introduced into a wide array of period and arrhythmic mutants (71, 89). The *kaiA* gene is 855 base pairs in length, with GTG as start codon, and encodes a 284 amino acid and 32.6-kD protein. The *kaiB* (309 bp) and *kaiC* genes (1560 bp) encode 11.4-kD and 58-kD proteins, respectively (Figure 3A). Primary motif searches of the Kai proteins revealed that KaiC contains two sets of putative P-loop ATP binding motifs (Walker’s A motif), two imperfect Walker’s B motifs, and two putative catalytic carboxylate glutamyl residues that are present in other ATP-binding proteins. In addition, KaiC carries two putative DXXG motifs, which are conserved in GTPase proteins. There is a significant similarity between the first and second halves of KaiC, with each half (CI and CII) containing one of the aforementioned ATP motifs, suggesting that KaiC is comprised of tandem duplicated domains (74). Neither KaiA nor KaiB has any recognizable motifs based upon primary protein sequence (71).

Further biochemical characterization revealed that KaiC binds ATP in vitro, with lower affinity, GTP (108, 118). Site-directed mutations in the Walker’s A motif within the CI domain of KaiC (K52H) resulted in a recombinant protein with markedly reduced ATP-binding activity in vitro. The reciprocal mutation in the CII domain (K294H) had no effect on nucleotide binding, suggesting distinct functions for the two halves of the KaiC protein (118). The molar ratio of ATP per KaiC protein was calculated at 22.6 +/− 0.8 pmol of ATP per 12.5 pmol of KaiC protein, suggesting that each KaiC protein binds two ATP molecules (64). In vivo, the K52H mutant was arrhythmic, and the K294H mutant had an extremely long period (70 h) with reduced amplitude (118). KaiC autophosphorylates when incubated with [γ-32p]ATP in vitro. Chemical stability assays (108, 118) and 2-D thin-layer chromatography (73) indicate that phosphorylation occurs at serine and/or threonine residues. The in vitro ATP binding and autophosphorylation data, correlated with Walker box mutant phenotypes in vivo, suggest that the ATP binding ability of KaiC, and hence autophosphorylation, is integral to the timekeeping capabilities of the Kai oscillator (108, 118).
The KaiA and KaiB proteins themselves have no autokinase activity (118); however, they play a significant role in the rate of KaiC autophosphorylation (73, 174). The addition of KaiA protein to KaiC in vitro increases the rate of KaiC autophosphorylation by approximately 2.5-fold (Figure 3B). Addition of KaiB alone to KaiC has no effect; however, when KaiA and KaiB are concomitantly added to KaiC, the rate of KaiC autophosphorylation is reduced to approximately half that of the KaiA-KaiC protein combination (174). Sequence alignments of KaiA proteins from various cyanobacteria show that two forms of KaiA proteins exist: a “long” KaiA (approximately 300 aa) similar to KaiA of *S. elongatus*, and a “short” form that aligns with the C-terminal segment of the *S. elongatus* protein (Figure 4A). The existence of short KaiA proteins in some species suggested that two independent domains exist in KaiA, which was confirmed by limited proteolysis of recombinant *S. elongatus* KaiA purified from *Escherichia coli*. The carboxy-terminal domain (residues 180–284) of KaiA, the portion shared among diverse cyanobacteria, stimulates KaiC autophosphorylation as observed for full-length KaiA, whereas the amino-terminal domain does not (174). A potential role for the N-terminal domain in input to the clock is discussed later.

The data define at least one of the functions of KaiA and KaiB: to act as opposing modulators of KaiC autophosphorylation (73, 174). KaiA and KaiB also influence the dephosphorylation rate of KaiC (179). Fully phosphorylated KaiC, when incubated alone in vitro, dephosphorylates to 50% after a 10-h incubation time. When incubated with KaiA, KaiC retains its phosphorylated state; however, when incubated with KaiA and KaiB, KaiC dephosphorylation occurs, albeit at a reduced rate relative to KaiC alone. KaiB has no effect on KaiC dephosphorylation in the absence of KaiA (179). Furthermore, KaiA protein is essential for the in vivo phosphorylation of KaiC; a phosphorylated KaiC band is not detectable in *kaiA* null, but remains present in *kaiB* null, mutant cell extracts (73, 85, 179).

As expected from their biochemical interactions, the Kai proteins physically interact in vitro and in vivo (74, 179a). Homotypic interaction between KaiB monomers can be detected in vivo by bioluminescence resonance energy transfer (179a). In yeast two-hybrid assays and in vitro affinity resin assays, each Kai protein forms either homotypic or heterotypic interactions. The KaiA/KaiB interaction is quite weak; however, the addition of KaiC enhances this interaction, which suggests that the Kai proteins might exist in heteromultimeric complexes within the cell (74). Interestingly, the in vitro interaction of E103K-KaiA, a 33-h period mutant of KaiA, with KaiB is also enhanced, and not further strengthened by the addition of KaiC. Each domain of KaiC, CI and CII, is also capable of interacting with the other Kai proteins. Although both domains are imperative for Kai oscillator function, it was hypothesized that each domain has its own functional property, because CI enhances the KaiA/KaiB interaction (even better than full-length KaiC), whereas the CII domain does not (74) (Figure 5A).

To identify KaiA binding regions of KaiC, six segments (60 to 110 aa in length) of KaiC were generated and tested in yeast two-hybrid and in vitro assays for interactive activity with full-length KaiA. Two KaiA-interacting domains of KaiC were
identified, from amino acyl residues 206–263 and 418–519 (designated C_{KABD1} and C_{KABD2}, respectively) (159) (Figure 3A). Known kaiC mutations that affect oscillator period fall within these regions. These same mutations, when introduced within the recombinant C_{KABD} fragments, reduced the interactive ability with KaiA protein. Reciprocally, period mutants of KaiA had altered binding with C_{KABD} fragments (159).

Most importantly, the Kai proteins interact in vivo as shown by an affinity chromatography method (74). Three kai loci were constructed that contained an allele of kaiA, kaiB, or kaiC that encodes an amino-terminal His6-tagged variant; these were then expressed in a kai null strain. Complementation of P_{kaiBC::luxAB} (promoter for kaiBC transcriptionally fused to luxAB) rhythms in the null background demonstrated that each His6-Kai protein is functional. Affinity purification of each His6-Kai from cell extracts, and subsequent immunoblot analysis, showed heterotypic interactions among all Kai proteins in S. elongatus cell extracts (74). Genetically, interactions between KaiA and KaiC are supported by the identification of a kaiC allele that suppresses a kaiA long-period allele phenotype (73).

More insights into the mechanism of the Kai oscillator arise from observing kai RNA and protein dynamics over a 24-h timescale. Entrained cyanobacterial extracts that were sampled in constant light (LL) for 48 h showed that kaiA and kaiBC mRNA levels cycle with peak accumulation at circadian time (CT) 9 to 12 and 33 to 36 h, or subjective dusk (CT0 = subjective dawn, CT12 = subjective dusk) (71, 75). Accordingly, levels of KaiB and KaiC proteins also cycle in LL, with peak expression at CT 15, following peak levels of kaiBC mRNA. However, KaiA protein level remains relatively high at all phases even though kaiA mRNA is rhythmic, demonstrating that a rhythm in mRNA abundance does not necessarily predicate rhythmic protein production and/or activity (178). Artificially inducing a pulse of KaiC protein during the trough of KaiC accumulation induces a large phase shift in P_{psbAI::luxAB} reporter activity, demonstrating that timing of KaiC accumulation is integral to Kai oscillator function (178).

The autokinase activity of KaiC is also rhythmic. It was noted that KaiC appears on immunoblots as a doublet band in LL cell extracts, with the upper, phosphorylated band predominant around CT 12–16 (early subjective night) and nearly absent around CT 0–4. Lambda phosphatase treatment (which removes phosphoryl groups from phosphorylated serine, threonine, tyrosine, and histidine residues) of the same cycling samples abolished the upper, phosphorylated KaiC band (73).

Timing of Kai interactions and cellular localization provide a window into clock complex dynamics in vivo (78, 85). Size exclusion gel filtration chromatography and subsequent immunoblots of a time series of cyanobacterial extracts showed that the Kai proteins are present in large complexes that change in stoichiometry over the circadian cycle. In the subjective day, KaiC is present in complexes that range from 350 to 440 kD; during the subjective night (CT16 and 22), KaiC-containing complexes increase in size to approximately 500 kD or larger. Coimmunoprecipitation assays performed after gel filtration of the large KaiC complexes determined that these large evening complexes contain KaiA and KaiB proteins (78). Investigation
into the subcellular localization of Kai proteins demonstrated that, whereas KaiA and KaiC are localized within the cytosol, KaiB protein appears to be localized to a particulate fraction, presumably the cell membrane, during the subjective day, and is released into the cytoplasm around CT 20, 4 h later than the total KaiB accumulation peak (85). Whether or how KaiB is sequestered to the cell membrane is not known; however, the presumptive membrane interaction would probably be peripheral, as KaiB has no identifiable hydrophobic transmembrane-spanning regions, and the interaction is sensitive to high salt and alkaline conditions (85).

KaiC also forms a higher-order complex with itself (Figure 5B). Sedimentation velocity analysis and gel filtration chromatography of purified recombinant KaiC from *S. elongatus* or *Synechococcus lividus* P2 (a thermophilic cyanobacterial isolate) showed that KaiC has a mass consistent with a hexameric form that is dependent upon the presence of ATP (109). Electron microscopy of KaiC reveals a hexagonal ring-like structure of KaiC, with an estimated diameter of the ring at 10 nm (109). The CI domain of KaiC also forms ring structures, so it is hypothesized that full-length KaiC forms a hexamer of “dumb-bell”-shaped monomers (109). The hexameric structure of KaiC was independently determined using recombinant KaiC protein from *Thermosynechococcus elongatus* BP-1. Mutants of the *T. elongatus* KaiC Walker’s A motif inhibited hexamer formation, and addition of a non-hydrolyzable ATP analog induced ring formation, suggesting that ATP binding, but not hydrolysis, is essential for KaiC hexamerization (64). The hexameric formation of KaiC is reversible, as dialysis of ATP away from KaiC hexamers shifts the protein to the monomeric form, and addition of ATP back to KaiC monomers reassembles the hexamer (64, 109).

These data suggest that self-interactions of KaiC protein, interaction of KaiC with KaiA and KaiB, and the subsequent phosphorylation state of KaiC protein, are central to the timekeeping ability of the Kai oscillator. The cyclic timing of KaiC production and phosphorylation, and the role KaiC plays as a scaffold for production of higher-order KaiC structures and the KaiA/KaiB/KaiC oscillator complex, all point to KaiC as the driving force for circadian rhythms in cyanobacteria.

**TRANSCRIPTIONAL CONTROL IN THE S. ELONGATUS TIMING MECHANISM**

At the center of models for eukaryotic circadian clock mechanisms are delayed negative feedback loops that regulate circadian timing, and subsequent circadian outputs. The negative feedback loops involve positive effector proteins that stimulate core clock gene expression. These core clock components then negatively regulate their own expression, resulting in an oscillatory pattern of gene expression on an approximately 24-h timescale (62, 182). The feedback loops of all eukaryotic model systems investigated thus far involve regulation at the transcription level. Positive elements (CLOCK/BMAL1 in mammals; CLOCK/CYCLE and PAR domain protein 1 in *Drosophila*; and White collar 1/White collar 2 in
Neurospora) activate transcription of core oscillator negative elements (mPER1, mPER2, mPER3, mTIM, mCRY1, and mCRY2 in mammals; PER, TIM, and VRI in Drosophila; and FRQ in Neurospora) that, along with posttranscriptional delays via phosphorylation, degradation, and nuclear transport, negatively inhibit their own transcription (10, 25, 29, 37, 62, 122, 169).

In S. elongatus, kai genes are expressed from two promoter regions, one upstream of kaiA and one upstream of kaiBC; however, no formal analysis of either promoter region has been published. Both messages are expressed with the same circadian phase relationship in wild-type cells, peaking under constant conditions at subjective dusk (71). Analysis of kai gene expression patterns has shown that the KaiC protein is required for wild-type levels of expression from the kaiBC promoter; however, ectopic overexpression of kaiC inhibits kaiBC gene expression, suggesting a negative role for KaiC protein in regulating its production. KaiA is required for expression from the kaiBC promoter, and overexpression of kaiA increases the level and abolishes the rhythmic pattern of kaiBC expression, suggesting a positive activator role for KaiA (71). The KaiA enhancement of kaiBC transcription is KaiC dependent, suggesting that there may be a cooperative effort between the KaiA and KaiC proteins for efficient transcription regulation at the kaiBC promoter (73). Despite homology of KaiC with the bacterial RecA/DnaB DNA recombinase and helicase family (92), the Kai proteins themselves have no obvious DNA binding motifs to suggest that the effects of Kai proteins on kai gene transcription are due to direct interactions with kai promoter DNA. No Kai protein has been shown to bind to specific kai promoter sequences; however, KaiC, in its phosphorylated hexameric form, has been shown to bind forked DNA molecules (109). This is consistent with the hypothesis that circadian control of chromosome structure may underlie global rhythmic gene regulation in S. elongatus (108).

Although some aspects of kai regulation are consistent with animal and fungal circadian timing models, there are data showing that direct transcriptional information is not required for circadian timekeeping in cyanobacteria. Mutations in S. elongatus have been identified that alter the phase relationship between kaiA and kaiBC expression rhythms without disrupting overall circadian timing, suggesting that the relative transcriptional activity of expression from at least the kaiA promoter is not important for generating circadian rhythms (82, 112). In addition, expressing kaiC from an E. coli promoter-consensus sequence in a kaiC genetic background restores circadian rhythmicity to gene expression patterns. These data indicate that neither kai-specific promoter regions, nor their resultant specific transcription regulation, are essential for generating circadian rhythms in S. elongatus (179). One caveat to these data is that this synthetic E. coli promoter is also expressed rhythmically in S. elongatus, and is recognized as Class 1 phase, similar to the kai genes. To directly test whether transcription timing at the kai locus is required for circadian oscillations, the natural timing of kaiA was bypassed by expressing it from the S. elongatus purF promoter, such that kaiA expression was delayed 12 h with respect to wild-type expression timing.
[peaking at subjective dawn rather than the normal subjective dusk (105)]. In a kaiA null but otherwise wild-type background, an ectopic kaiA gene restores wild-type circadian rhythms of gene expression, whether it is expressed from its native promoter or from the oppositely phased purF promoter (J.L. Ditty & S.S. Golden, manuscript in preparation). Curiously, in genetic backgrounds that increase the intracellular concentration of KaiB and/or KaiC protein, expressing kaiA from the purF promoter lengthens the free-running period of gene expression rhythms to nearly 33 h, whereas expression of kaiA from the wild-type promoter leaves the period unchanged (J.L. Ditty & S.S. Golden, unpublished results). Therefore, the scheduling of kaiA transcription seems to play a role in the robustness of circadian regulation in S. elongatus, for only when the oscillator is perturbed (e.g., elevated levels of KaiB or KaiC) is the timing of kaiA expression important for circadian rhythm generation. We conclude that transcriptional feedback is a minor reinforcer of the timing cycle, rather than the underlying mechanism.

**KAI OSCILLATOR INPUT: CIKA, LDPA, AND PEX**

A key player in the molecular mechanism for coupling diurnal environmental signals to the circadian clock, i.e., the input pathway, has been identified. The cikA gene (for circadian input kinase) was originally found in a genetic screen for mutants of S. elongatus that had a subtle alteration in light-responsive regulation of the photosystem II gene, psbAII (144). Circadian gene expression assays revealed additional phenotypes. Bioluminescence oscillation patterns in a cikA mutant from either Class 1 (kaiA, kaiB, psbAII) or Class 2 (purF) reporter fusions have a reduced amplitude and a period that is shortened by approximately 2 h; some genes also show unusual phasing. However, the most striking phenotype of the cikA mutant is that its circadian clock is virtually unresponsive to various environmental stimuli. A lack of phase resetting in the cikA mutant background in response to a 5-h dark pulse (144) or a temperature pulse (M. Katayama, H. Iwasaki & T. Kondo, personal communication) demonstrates that CikA is involved in the input pathway to the Kai circadian clock.

The deduced amino acid sequence of the CikA protein reveals three motifs consistent with CikA involvement in clock input: a potential chromophore binding (GAF) motif, a histidine protein kinase (HPK) motif, and a cryptic two-component receiver domain (RR) (Figure 6). A GAF domain in true bacterial or plant phytochromes possesses lyase activity for covalent bilin attachment to the protein. However, the GAF of CikA is atypical because CikA lacks the conserved cysteinyl or histidyl residues expected for covalent bilin binding (70). A histidine-tagged CikA (CikA6His) will ligate phytochromobilin and phycocyanobilin chromophores in vitro, indicating lyase activity even in the absence of conserved bilin-binding residues. However, the adduct is not photoactive. Removing the GAF motif reduces, but does not eliminate, the ability of CikA to form a covalent adduct with phycocyanobilin. Isolation of CikA6His from the cyanobacterium did not
reveal a covalently bound bilin, and expression in recombinant *E. coli* that express phycocyanobilin or biliverdin, which could have detected a photoactive non-covalent holoprotein, was negative. These results suggest that the cryptic bilin binding by CikA in vitro is unlikely to reflect a biliprotein complex in *S. elongatus* (110).

Downstream of the GAF, CikA possesses the conserved histidyl residue (H393) for autophosphorylation in an HPK, and the conserved N-, D/F-, and G-boxes of the ATPase motif (151, 152). CikA has authentic HPK activity by in vitro autophosphorylation assays, and chemical stability is consistent with phosphoryl linkage at a histidyl residue. Changing the H393 codon to encode an alanyl residue blocks autophosphorylation. Removal of the N-terminal portion or GAF of CikA protein drastically reduces autokinase activity, suggesting that these domains modulate CikA autophosphorylation (110).

C-terminal to the HPK of CikA is a motif that is similar to receiver domains of response regulators involved in two-component signal transduction systems (151, 152). However, this motif (RR) of CikA is cryptic in that it does not contain the conserved aspartyl residue that, in a response regulator, would be phosphorylated by a cognate HPK. The potential for transfer of a phosphoryl group from the HPK to the RR was tested using a CikA donor that lacks RR but retains autokinase activity and a recipient H393A-CikA that contains RR but is defective for autokinase activity. No *trans*-phosphorylation occurred. It is most likely that the RR region of CikA is a *pseudo*-receiver, and does not accept a phosphoryl group from the HPK motif of CikA or any other phosphoryl donor. The conformational switch that phosphoryl transfer induces on a bona fide RR may be regulated in CikA by a protein-protein interaction instead (120). Removal of the RR motif from CikA greatly increases autophosphorylation activity of the protein, consistent with its role as a modulator (110). Exactly how CikA functions in Kai-clock input remains to be elucidated. We predict that CikA interacts with other, yet unidentified, *S. elongatus* phytochromes and/or response regulators for environmental information relay. A number of potential CikA partners have been identified by yeast two-hybrid assays (J.-S. Choi, S.R. Canales & S.S. Golden, unpublished).

The *ldpA* gene (light-dependent period) affects input to the clock in a different way than CikA. It was identified from two independent transposon insertions in a screen for mutants that respond differently than wild type to a phase-altering dark pulse (81). Although insertions at this locus were originally identified as phase-response mutants, the actual circadian phenotype of an *ldpA* null is more complex: It has a conditional 1-h shortening of the free-running period (81). As described previously, the bioluminescence from many cyanobacterial reporters adhere to Aschoff’s Rule, such that the period length varies with an inverse relationship to light intensity (81, 112) (Figure 1C). In the *ldpA* mutant background, these typical period variations are abrogated (81). The deduced amino acid sequence of *ldpA* to motif databases suggests that LdpA is a soluble protein that contains two Fe-S centers, likely one 3Fe-4S and the other 4Fe-4S. This would be consistent with
a pathway for transducing a measure of photosynthetic activity to the clock via a redox pathway (81).

The pex gene (period extender) was originally identified as “apparently complementing” a 22-h short period circadian clock mutant (91). Further analysis showed that an ectopic copy of pex acts as a suppressor of the 22-h period by extending the period of circadian rhythms by about 2 h. Overproduction of Pex in either wild-type or various period-mutant backgrounds extends circadian periods and decreases amplitude (91). The pex gene is 447 bp in length, and encodes a 17-kDa protein with no obvious functional motifs. Inactivation of the pex gene results in a slightly short period phenotype, with no apparent growth defect. The disruption of pex on kaiA expression is noteworthy, as loss of pex greatly increases expression of kaiA, indicating that Pex may act as either a direct or indirect repressor of kaiA expression. No specific role in an input pathway can be assigned.

At the end of environmental information relay to the oscillator there must be a mechanism to receive and interpret the information by the Kai complex. Structural information of the KaiA protein has provided insights into this mechanism (174). As described previously, the conserved C-terminal domain of KaiA contacts KaiC and stimulates phosphorylation of the latter protein. In S. elongatus KaiA an N-terminal domain is connected to the C-terminal domain via a flexible linker. NMR structure determination of the first 135 residues of KaiA reveals a fold similar to that of the receiver domains of two-component signal transduction systems (167, 174) (Figure 4B). Despite this structural similarity, it is not likely that the amino-terminal domain of KaiA acts as a canonical receiver, as it lacks conserved aspartyl residues known to be required for phosphoryl transfer activity. Instead, like the RR domain of CikA, N-terminal KaiA may act as a pseudo-receiver as does AmiR, which responds conformationally not to autophosphorylation, but rather to protein-protein interactions (120, 174). We propose that the input pathways impinge on the clock via this domain, which in turn modulates the autophosphorylation of KaiC. However, CikA does not interact directly with full-length KaiA in a yeast two-hybrid assay (J.-S. Choi & S.S. Golden, unpublished data). We predict a signal transduction pathway that connects CikA and KaiA.

KAI OSCILLATOR OUTPUT: SASA, CPMA, AND GROUP TWO SIGMA FACTORS

The conduit by which the Kai clock transfers temporal information to downstream gene expression and subsequent biological activity is as important as the timing mechanism itself. One protein, SasA (Synechococcus adaptive sensor), appears to play a key role in circadian output. SasA was originally identified in 1993 not for its circadian functions, but from a genetic screen seeking cyanobacterial sensor
The sasA gene complements E. coli mutants that lack the histidine protein kinase sensors of prototypical two-component regulatory systems: an envZ mutant defective for signaling needed in the production of the outer membrane protein OmpC, and a phoR/creC mutant that fails to produce alkaline phosphatase (111, 151, 152). The role that SasA plays in cyanobacterial timekeeping was discovered several years later when independent lines of research identified SasA as integral to the circadian mechanism (75). A database search highlights the amino terminus of SasA as having strong resemblance to the full-length KaiB protein, with identity and similarity of 26% and 60%, respectively. Disruption of sasA in S. elongatus reduces PkaiBC::luxAB expression, and drastically reduces amplitude, but does not completely abrogate circadian bioluminescence rhythms. The residual rhythms have a short circadian period. The same effect is seen on kaiBC mRNA and on KaiB and KaiC protein levels. Expression from most other promoters is arrhythmic in a sasA null background, clamped high, low, or midline for different genes.

SasA was independently identified as interacting physically with KaiC via a yeast two-hybrid screen (75). The first 97 residues of SasA (homologous to KaiB) were pulled out as a “prey” with a KaiC “bait.” Subsequent in vitro assays confirmed that full-length SasA, as well as SasA residues 1–97 alone, interacts with itself and with KaiC. Either the CI or CII half of KaiC protein is sufficient for this interaction. In temporally collected cell extracts, SasA protein does not seem to oscillate over the circadian cycle; however, SasA coimmunoprecipitates with the aforementioned subjective evening 500-kD Kai protein complexes, interacting with KaiC at CT16–CT22 (78). Therefore SasA, although not essential for rhythmicity, is in close physical association with the Kai oscillator complex and is important for robust rhythms of transcription from downstream genes.

NMR analysis of the first 105 amino acid residues of SasA indicates a secondary structure of βαβαβαβα, and preliminary structures suggest a thioredoxin-like fold (86). The carboxy terminus of SasA is homologous to histidine protein kinases, and undergoes autophosphorylation in vitro (111). To determine the role of SasA autophosphorylation in circadian timing, the conserved catalytic histidyl residue was substituted with a glutaminyl residue, H162Q. The resultant phenotype was as the sasA null: reduced kaiBC promoter activity and amplitude of rhythmicity, indicating a requirement of SasA autophosphorylation for wild-type rhythms in S. elongatus (75). A cognate response regulator (151) that would receive a phosphoryl signal from SasA has not yet been identified, so the pathway for SasA communication with downstream genes remains elusive.

The circadian clock in the sasA null background entrains normally to 12-h dark and temperature pulses, but cells are impaired when grown in an LD cycle (75). One possible interpretation is that SasA acts on an input pathway, affecting interpretation of light and dark signals. However, disruption of sasA globally affects downstream genes, and the LD defect could arise indirectly from altered expression
of any one of them. Because the sasA LD defect is not shared with kai null strains, it may result from a function of SasA that is unrelated to its role in the clock. The strongest evidence for SasA as an output pathway component is that whereas association with KaiC stimulates SasA autophosphorylation, SasA has no effect on KaiC autophosphorylation (S.B. Williams & S.S. Golden, manuscript in preparation). Thus, information flow seems to be from the Kai complex to SasA, rather than the other way around.

Continuous overproduction of SasA in cyanobacterial cells completely represses the kaiBC promoter; however, transient pulses of SasA affect kaiBC promoter expression differently as a function of the point in the circadian cycle at which SasA is elevated. When expressed during subjective day, SasA production causes a phase delay in kaiBC expression, and during subjective night a phase advance, consistent with role in close association with, but that is not an integral part of, the Kai circadian oscillator (75).

Because the circadian oscillator pervades expression of the entire Synechococcus elongatus genome, the regulatory mechanism that couples timing information to downstream genes may be global in nature and tied to fundamental metabolic processes (95). Some evidence points to an underlying rhythmicity in the basic transcriptional machinery (112, 165). A mutant with an insertion in rpoD2, which encodes a group 2 sigma factor, results in a low-amplitude rhythm phenotype from a subset of cyanobacterial promoters (165). Sigma factors are subunits of prokaryotic RNA polymerase that confer promoter specificity (57), and cyanobacteria are atypical in that they contain multiple, closely related, sigma factors that are not essential for growth (group 2) in addition to the essential housekeeping sigma RpoD1 (157). Further analysis showed that inactivation of any of the four known group 2 sigma factor genes (rpoD2, rpoD3, rpoD4, and sigC), either singly or in pairs, alters circadian expression of the psbAI promoter. The sigC mutation, which consistently lengthens the period of psbAI expression by 2 h, has little or no effect on the period of expression from either the purF (Class 2 reporter) or kaiBC promoter fusions. Interestingly, the kaiBC promoter is affected only by mutations in rpoD2, or pair-wise mutations in rpoD3/rpoD4 and rpoD2/rpoD3, suggesting a relative insulation of the kaiBC promoter from other gene-expression pathways. Because each sigma factor dissimilarly affects transcription from specific cyanobacterial promoters, the working model contends that the cyanobacterial transcription apparatus oscillates in a circadian manner (oscillations in RpoD4 protein level have been demonstrated), and the composition of the transcriptional holoenzyme for individual group 2 sigma factors changes over the circadian cycle (112). No direct link between SasA and the expression and/or activity of any group 2 sigma factor has been established.

The cpmA gene (circadian phase modifier) also seems to play a role in output from the S. elongatus circadian clock. Inactivation of cpmA drastically alters the relative phasing of a subset of cyanobacterial transcriptional reporters. Most interestingly, the cpmA mutation affects the phase angle of PkaiA::luxAB, but has no effect on PkaiBC::luxAB, resulting in a relative phase difference of up to 10 h
between the two kai operons. Despite lack of transcriptional coordination in this background, circadian timing is intact (82).

**DISCUSSION**

**kai Genes and Circadian Clock Evolution**

Subsequent to confirmation that the three kai genes are essential components for the circadian timing mechanism in *S. elongatus* (71), the genomes of six phylogenetically diverse cyanobacteria, *Anabaena* sp. strain PCC 7120, *Nostoc punctiforme*, *Synechococcus* sp. strain WH 8102, *Synechocystis* sp. strain PCC 6803, *Thermosynechococcus elongatus* BP-1, and *Trichodesmium erythraeum*, have been completely sequenced and each contains the genetic basis for a circadian clock: at least one kai locus. To date, functional clock-driven circadian rhythms have been tested, and subsequently demonstrated, only in species of the unicellular genera *Synechococcus* and *Synechocystis* (4, 5, 24, 69, 88, 89, 155). However, 40 diverse cyanobacterial strains from 20 different genera have at least one kai gene (96). In addition to the typical phycobilisome-containing cyanobacteria, kai clock genes are found in two prochlorophyte, chlorophyll b-containing strains, *Prochlorococcus marinus* MED4 and MIT 9313; both have kaiB and kaiC genes, but lack an obvious kaiA ortholog (68). Although circadian rhythms, and their requisite functional circadian oscillator, have not been demonstrated directly in either of these strains, *Prochlorococcus* sp. strain PCC 9511 displays distinct diel patterns of gene expression (67). Diel patterns do not necessarily indicate the existence of a circadian clock, but the absence of the diel rhythms would have suggested absence of a clock.

In *S. elongatus* KaiA plays an essential role in circadian rhythm generation. However, the N-terminal pseudo-receiver domain is missing in many species, and the entire protein is absent from prochlorophyte species that are likely to have circadian rhythms. Lack of ubiquity of KaiA is not really surprising across the vast evolutionary distance of the cyanobacteria, especially given the diverse mechanisms these organisms use for harvesting and interpreting light (12, 14, 20). If the role of KaiA is really as a conduit for information from the input pathways, this function may be filled by completely unrelated proteins in other species. The chief requirement is an ability to receive environmental information and dock with the clock complex, inducing a conformational change.

Molecular phylogenetic analysis of the kai locus suggests that the kaiBC operon is nearly 2 billion years old (39). It seems likely that the common ancestor shared by cyanobacteria and plastids also had prokaryotic circadian kai genes. So, where are the kai gene homologs in the plastid lineage? It is not clear. Plastid genomes from algae such as the glaucocystophyte *Cyanophora paradoxa*, the rhodophyte *Porphyra purpurea*, and the chlorophytes *Chlamydomonas reinhardtii* and *Chlorella vulgaris* are among over two dozen that have been completely sequenced (48, 100, 135, 170); none of these encodes obvious kai orthologs. Perhaps kai genes will
be found as more plastid genomes are sequenced or else they may be genuinely absent in plants, reflecting something fundamental about the prokaryotic circadian clock and endosymbiosis.

Plastid genome evolution has evidently been a process of size reduction, as most plant genes considered cyanobacterial in origin have been incorporated into the nuclear genome (34, 97, 133). Cyanobacterial genomes range from about 1.8 to 9 million base pairs, whereas plastid genomes are much smaller, ranging from around 35–200 kb (33, 68). As a result, in A. thaliana about 18% of nuclear genes appear to be of cyanobacterial origin (98). Although there is no clear correlation between the evolutionary origin of an A. thaliana gene and the cellular localization of its functional product, many cyanobacterial-like nuclear genes encode products that act exclusively in chloroplast photosynthetic chemistry (97, 98). Perhaps in A. thaliana, circadian regulation of incorporated cyanobacterial genes by an existing nuclear-encoded clock (15, 65, 140, 142, 160, 161) was sufficient for circadian control of chloroplast development and photosynthetic activity (2, 9, 47). As plastid genes were relocated to the nuclear genome, likely for regulatory reasons, the plastid clock became redundant, in light of the fact that the plastid clock was not likely amenable to temporal signals from the plant’s tissue-specific circadian oscillator(s).

Another fascinating plant circadian system may shed more light on the prokaryotic (kai) circadian clock mechanism. Organisms from several genera of single-celled eukaryotic green algae, such as Acetabularia, Chlorella, and Gonyaulax, generate circadian rhythms in a variety of physiological functions including photosynthetic metabolism and individual protein abundances (36, 63, 101, 147, 156, 168, 177). Incredibly, the nucleus can be removed from the rest of the Acetabularia mediterranea cell and circadian oscillations in photosynthesis and protein abundance continue, under constant conditions, for many weeks in the enucleated cell (154). When present, the nucleus may impart phase information upon rhythms because, when the nucleus is entrained differentially relative to the rest of the cell and then added back to the enucleated cell, the nucleus affects phasing of cellular rhythms (102, 147, 154).

The underlying timing mechanism in this single-celled eukaryotic alga may resemble that of the prokaryotic Kai clock. Current eukaryotic circadian mechanism models hold that delayed re-entry of transcription regulators into the nucleus is a key temporal component of circadian oscillators (182). In contrast, models that have been proposed to account for most of the circadian phenomena in A. mediterranea include the partitioning of an essential, rhythm-driving protein across an organellar membrane (instead of the nuclear membrane) to monitor time. Theoretically, properties of the membrane, integral membrane transport proteins, or a proposed delayed-release protein, would be altered as a function of time and thereby participate in the A. mediterranea circadian timing mechanism (147, 154). Recent work that suggests KaiB sequestration to the membrane and delayed re-entry into the cytoplasm as a function of the Kai timing mechanism (85) is reminiscent of this 30-year-old hypothesis (147, 154).
A MODEL FOR S. ELONGATUS PCC 7942 TIMEKEEPING

A comprehensive model for the *S. elongatus* Kai data reviewed above is illustrated in Figure 7. At subjective dawn or sunrise, CikA protein, either through its own chromophore or the chromophores of other postulated interactive proteins, recognizes light input information, and relays that information to the Kai clock for synchronization to local time. The daily light (and temperature) information is integrated into the Kai mechanism by the amino-terminal, *pseudo*-receiver domain of KaiA. The information relay is probably via a protein-protein interaction, with an as-yet unidentified protein(s), given the structural similarity of amino-terminal KaiA to bacterial receivers that propagate conformational changes based on information from protein interactions (120, 139). At dawn and during the subjective day, transcription of the *kaiA* and *kaiBC* operons and subsequent translation of KaiA, KaiB, and KaiC monomers occurs. KaiB protein is sequestered to the cell membrane by an as-yet unidentified, membrane-associated or membrane-bound KaiB-interactive protein.

Meanwhile, KaiC protein monomers bind ATP, inducing the hexamerization of KaiC, and the carboxy-terminal domain of KaiA aids in KaiC autophosphorylation. By subjective dusk, KaiC is fully phosphorylated and associated with KaiA, and *kaiBC* transcription has reached its peak level. KaiB protein is still membrane localized. Interactions between the amino terminus of SasA (which has high sequence similarity to full-length KaiB) with KaiC may be occurring at this point, allowing for the autophosphorylation of SasA to occur while KaiB is still localized to the cell membrane. Phosphorylated SasA could then transfer its phosphoryl group to activate its cognate, as-yet unidentified, response regulator (SasR), which transmits timing information to downstream genes. Additionally, the putative SasR protein could react, either directly or indirectly, with the *kaiBC* promoter, to negatively inhibit transcription. At CT20, KaiB is released from the membrane, and association of KaiB with the KaiC/KaiA/SasA complex is initiated, reducing the autophosphorylation of KaiC, and subsequently the phosphorylation of SasA. Introduction of KaiB to the clock complex itself, or the action of inhibiting KaiC phosphorylation, may induce dissociation of the of the Kai clock complex by the subsequent subjective dawn.

FINAL COMMENTS

Exploiting the genetically flexible cyanobacterium *S. elongatus* as a model for the elucidation of a circadian mechanism has been fruitful. In the decade since luciferase reporter rhythms were first reported, core components from input pathways, output pathways, and the oscillator itself have been identified, and the key features of the underlying biochemistry of timekeeping have been established. It is clear that the Kai oscillator is a physical complex: an entity that is built, modified, and re-generated over a 24-h time span. The cyanobacterial timing mechanism
is fundamentally different from models for eukaryotic timing in this respect. In eukaryotic systems, circadian timing is based upon two separately phased, but intertwining, feedback loops that involve partner switching between positive and negative elements (37, 62, 169). In \emph{S. elongatus}, it is the intricate relationship between the Kai proteins and SasA, and timed modifications to these proteins by phosphorylation, that constitute the oscillator. Transcriptional autoregulation occurs, but plays a minor role in the timepiece. As is evident from the proposed model of Kai timing, there are still a great many missing pieces of the clock machinery to be found. A comprehensive functional genomics project is under way (http://www.bio.tamu.edu/synecho/) to assay the effect of inactivating each locus in the \emph{S. elongatus} genome on circadian rhythms to identify all the missing components. Only time will tell.

ACKNOWLEDGMENTS

We gratefully recognize our long-time collaborators in the T. Kondo and C.H. Johnson labs for sharing research findings with us prior to publication. We thank I. Vakonakis, A. LiWang, J. Sun, and A. Holzenburg for providing images used in some of the figures. The research of the authors was supported by postdoctoral fellowships to J.L.D. (NSF PA 99-025) and S.B.W. (NIH GM19644) and grants from the NIH (P01 NS39546 and R01 GM62419) and NSF (MCB-9982852 and 0235292) to S.S.G.

\textit{The Annual Review of Genetics} is online at http://genet.annualreviews.org

LITERATURE CITED

1. Abed RM, Garcia-Pichel F, Hernandez-Marine M. 2002. Polyphasic characterization of benthic, moderately halophilic, moderately thermophilic cyanobacteria with very thin trichomes and the proposal of \emph{Halomicronema excentricum} gen. nov., sp. nov. \textit{Arch. Microbiol.} 177: 361–70
7. Beard SJ, Hayes PK, Pfeifer F, Walsby AE. 2002. The sequence of the major gas vesicle protein, GvpA, influences...


activity and apparent toxicity of microcystins and related compounds. *Toxicon* 40:1017–25


128. Palemik B, Haselkorn R. 1992. Multiple evolutionary origins of prochlorophytes,


Figure 1  General characteristics of the *S. elongatus* PCC 7942 circadian clock and rhythms. A. General diagram of the three conceptual divisions of circadian systems: input pathways, the circadian oscillator, and output pathways. Known *S. elongatus* clock genes are listed below the division in which each is thought to function. B. Characteristics of circadian bioluminescence rhythms as measured from *S. elongatus* reporter strains. On the abscissa negative time designates time during light/dark cycles (L/D) for clock entrainment. Positive values represent circadian time (CT), i.e., time in a constant light (LL) environmental condition. By convention, a circadian hour is 1/24 of a cycle. Alternating black bars (12-h time blocks) represent time spent in darkness during L/D. Alternating hatched bars represent “subjective” dark times during LL conditions. Bioluminescence in counts per second (cps) is indicated on the ordinate axis. Circadian parameters of period (peak-to-peak duration), phase (CT of the peak during a cycle relative to a reference, here ‘lights on’), and amplitude (the magnitude of oscillation from the mean). C. Bioluminescence traces following the *PpsbA1::luxAB* reporter strain AMC539 under high and low light conditions demonstrate Aschoff’s Rule. Closed circle, high light (shorter period); open circle, low light (longer period).
Figure 2  *S. elongatus* PCC 7942 kai genes and mutant phenotypes. A. Schematic of the kai locus. Yellow arrows, promoters for both kaiA and kaiBC operons; hash marks below genes, positions of point mutations known to alter bioluminescence rhythms in vivo. B. Wild-type and kaiA mutant bioluminescence traces from Class 1 and Class 2 reporters. Green symbols and ordinate scale, Class 1; blue symbols and ordinate scale, Class 2. Closed green square, AMC541 (P<sub>kaiB</sub>::luc, wild-type background); open green square, AMC702 (P<sub>kaiB</sub>::luc, kaiA deletion background); closed blue circle, AMC601 (P<sub>purF</sub>::luc, wild-type background); open blue circle, AMC1148 (P<sub>purF</sub>::luc, kaiA deletion background).

Figure 3  Physical and biochemical characteristics of Kai proteins. A. Diagram of KaiA, KaiB, and KaiC primary structures. Asterisk in KaiA, the E103K mutation; wavy black line, a flexible linker between folded (green) domains; yellow bars in the CI and CII portion of KaiC, Walker’s A, or P-loop motifs, required for ATP binding; asterisks within P-loop 1 and 2, K52H and K294H mutations; red bars in both CI and CII, Walker’s B motifs; blue bars in CI, DXXG motifs; orange bars, glutamyl residues E77 and E78 in CI and E318 and E319 in CII. The KaiA-binding regions in CI and CII are indicated by brackets (C<sub>KARD1</sub> and C<sub>KARD2</sub>). B. Effects of Kai proteins on KaiC autophosphorylation. KaiC autophosphorylation in the presence of: no other proteins (open squares); KaiA (closed squares); KaiB (closed circles); KaiA and KaiB (open circles).
A.

*S. elongatus* PCC 7942
*Synechocystis* PCC 6803
*Synechococcus* WH 8102
*Nostoc punctiforme*
*Anabaena* PCC 7120
*S. elongatus* KaiA domains

KaiA135N
KaiA180C

B.

Figure 4  KaiA protein structural information. A. Graphic representation of KaiA homologs from various cyanobacteria. Multicolored amino-terminal domains represent regions of divergence from *S. elongatus* KaiA; carboxy-terminal regions (all green) represent regions of high sequence conservation. *S. elongatus* KaiA domains were determined by limited proteolysis (174). B. NMR structure of the amino terminus of KaiA (residues 1–135) and bona fide receiver domains of NtrC and CheY. An aspartyl residue that is phosphorylated by a cognate HPK is shown in red for NtrC and CheY; and is replaced by asparagine in KaiA135N.

Figure 6  Scheme of CikA protein domain interactions. Phosphorylation of the H box in the kinase domain of CikA is positively regulated by the GAF domain and the amino-terminal region of CikA, and is inhibited by the RR pseudo-receiver.
Figure 5  S. elongatus PCC 7942 clock protein interactions. A. Graphic representation of known interactions between KaiA, KaiB, KaiC, and SasA. Yellow stars labeled “P” represent phosphoryl groups; bound ATP molecules are indicated in yellow. B. Electron micrograph of a 2-dimensional crystal of hexameric KaiC (courtesy of Y. Jingchuan and A. Holzenburg, Microscopy and Imaging Center, Texas A&M University).

Figure 7  Current model for the circadian timekeeping mechanism of S. elongatus PCC 7942. Shaded area designates proposed events that may drive downstream circadian gene expression. See text for full description.
# CONTENTS

<table>
<thead>
<tr>
<th>Topic</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>IRA HERSKOWITZ (1946–2003), The Editors</td>
<td>1</td>
</tr>
<tr>
<td>TRANPOSITION-BASED STRATEGIES FOR MICROBIAL FUNCTIONAL GENOMICS AND PROTEOMICS, Finbarr Hayes</td>
<td>3</td>
</tr>
<tr>
<td>ERROR-PRONE DNA POLYMERASES: WHEN MAKING A MISTAKE IS THE ONLY WAY TO GET AHEAD, Alison J. Rattray and Jeffrey N. Strathern</td>
<td>31</td>
</tr>
<tr>
<td>GENETICS OF HAIR AND SKIN COLOR, Jonathan L. Rees</td>
<td>67</td>
</tr>
<tr>
<td>THIOL-BASED REGULATORY SWITCHES, Mark S.B. Paget and Mark J. Buttner</td>
<td>91</td>
</tr>
<tr>
<td>PSEUDGENES: ARE THEY “JUNK” OR FUNCTIONAL DNA? Evgeniy S. Balakirev and Francisco J. Ayala</td>
<td>123</td>
</tr>
<tr>
<td>UNUSUAL LIFE STYLE OF GIANT CHLORELLA VIRUSES, James L. Van Etten</td>
<td>153</td>
</tr>
<tr>
<td>GENETICS OF LACTASE PERSISTENCE AND LACTOSE INTOLERANCE, Dallas M. Swallow</td>
<td>197</td>
</tr>
<tr>
<td>CELL POLARITY AND THE CYTOSKELETON IN THE CAENORHABDITIS ELEGANS ZYGOTE, Stephan Q. Schneider and Bruce Bowerman</td>
<td>221</td>
</tr>
<tr>
<td>THE SPINDLE ASSEMBLY AND SPINDLE POSITION CHECKPOINTS, Daniel J. Lew and Daniel J. Burke</td>
<td>251</td>
</tr>
<tr>
<td>GENETICS OF AGING IN THE FRUIT FLY, DROSOPHILA MELANOGASTER, Stephen L. Helfand and Blanka Rogina</td>
<td>329</td>
</tr>
<tr>
<td>NATURAL SELECTION AND THE EVOLUTION OF GENOME IMPRINTING, Elena de la Casa-Esperón and Carmen Sapienza</td>
<td>349</td>
</tr>
<tr>
<td>THE NEED FOR WINTER IN THE SWITCH TO FLOWERING, Ian R. Henderson, Chikako Shindo, and Caroline Dean</td>
<td>371</td>
</tr>
<tr>
<td>TRANSMISSION RATIO DISTORTION IN MICE, Mary F. Lyon</td>
<td>393</td>
</tr>
</tbody>
</table>
CONTENTS

STRUCTURE, DIVERSITY, AND EVOLUTION OF PROTEIN TOXINS FROM SPORE-FORMING ENTOMOPATHOGENIC BACTERIA, Ruud A. de Maagd, Alejandra Bravo, Colin Berry, Neil Crickmore, and H. Ernst Schnepf  409

YEAST VACUOLE INHERITANCE AND DYNAMICS, Lois S. Weisman  435

HEPARAN SULFATE CORE PROTEINS IN CELL-CELL SIGNALING, Kenneth L. Kramer and H. Joseph Yost  461

RETROTRANPOSONS PROVIDE AN EVOLUTIONARILY ROBUST NON-TELOMERASE MECHANISM TO MAINTAIN TELOMERES, Mary-Lou Pardue and P.G. DeBaryshe  485

A CYANOBACTERIAL CIRCADIAN TIMING MECHANISM, S.S. Golden, J.L. Ditty, and S.B. Williams  513

REGULATION OF CELL CYCLES IN DROSOPHILA DEVELOPMENT: INTRINSIC AND EXTRINSIC CUES, Laura A. Lee and Terry L. Orr-Weaver  545

RECOGNITION AND RESPONSE IN THE PLANT IMMUNE SYSTEM, Zachary Nimchuk, Thomas Eulgem, Ben F. Holt III, and Jeffrey L. Dangl  579

RECA-DEPENDENT RECOVERY OF ARRESTED DNA REPLICATION FORKS, Justin Courcelle and Phillip C. Hanawalt  611

INDEXES

Subject Index  647
Cumulative Index of Contributing Authors, Volumes 33–37  679
Cumulative Index of Chapter Titles, Volumes 33–37  682

ERRATA

An online log of corrections to Annual Review of Genetics chapters may be found at http://genet.annualreviews.org/errata.shtml