NMR structure of the KaiC-interacting C-terminal domain of KaiA, a circadian clock protein: Implications for KaiA–KaiC interaction

Ioannis Vakonakis*, Jingchun Sun†, Tianfu Wu§, Andreas Holzenburg*‡, Susan S. Golden§, and Andy C. LiWang*§

Departments of *Biochemistry and Biophysics and †Biology and ‡Microscopy and Imaging Center, Texas A&M University, College Station, TX 77843

Edited by Adriaan Bax, National Institutes of Health, Bethesda, MD, and approved December 5, 2003 (received for review August 27, 2003)

KaiA is a two-domain circadian clock protein in cyanobacteria, acting as the positive element in a feedback loop that sustains the oscillation. The structure of the N-terminal domain of KaiA is that of a pseudo-receiver, similar to those of bacterial response regulators, which likely interacts with components of the clock-resetting pathway. The C-terminal domain of KaiA is highly conserved among cyanobacteria and enhances the autokinase activity of KaiC. Here we present the NMR structure of the C-terminal domain of KaiA from the thermophilic cyanobacterium Thermosynecho-
coccus elongatus BP-1. This domain adopts a novel all-α-helical homodimeric structure. Several mutations known to affect the period of the circadian oscillator are shown to be located at an exposed groove near the dimer interface. This NMR structure and a 21-Å-resolution electron microscopy structure of the hexameric KaiC particle allow us to postulate a model of KaiA–KaiC interaction, in which KaiA binds a linker region connecting two globular KaiC domains.

Circadian timekeeping oscillators and their accompanying machinery convey widespread regulatory mechanisms that control the ~24-h temporal pattern of diverse physiological processes, ranging from leaf movement to transcription regulation (1, 2). Circadian oscillators have been identified in virtually all eucaryotic groups, including insects, fungi, plants, and mammals (3), and in cyanobacteria (4) and share mechanistic similarities, such as positive and negative transcription–translation feedback loops (3). The cyanobacterial clock is the simplest and potentially oldest (4) identified thus far, composed of three interacting clock proteins, KaiA, KaiB, and KaiC (4, 6) that form heteromultimeric complexes in vitro of sizes that oscillate in a circadian manner (7). Transcription from practically all promoter regions in Synechococcus elongatus PCC 7942 is under circadian control, and transcription levels oscillate in phase or in antiphase with transcription from the kai locus (4, 8). Indeed, even promoters from other organisms, such as Escherichia coli, become rhythmically expressed in S. elongatus (8).

We previously reported that KaiA enhances the autokinase activity of KaiC in vitro (9). The KaiC phosphorylation state changes in vivo with a circadian pattern (10, 11), and ATP-binding by KaiC promotes formation of a hexameric KaiC particle (12, 13), which is presumed to be the functional form of the protein. Evidence suggests that KaiC binds forked DNA when in the hexameric form (13). KaiC preparations have been subjected to electron microscopic studies by two independent groups resulting in two structures (12, 13). KaiB is known to attenuate the KaiA-enhanced KaiC autokinase activity (9, 10, 14); however, no other information regarding its biochemical function is presently available.

KaiA is a two-domain protein. We previously showed that the C-terminal KaiA domain alone is sufficient in vitro to enhance the autokinase activity of KaiC (9), whereas the N-terminal domain has no effect in that assay. The N-terminal domain of KaiA adopts the fold of a canonical response regulator receiver domain (9, 15), although the primary sequence is dissimilar to that of receivers, and it lacks the conserved aspartyl residue necessary for phosphorylation. Similar to other pseudo-receiver domains (16), KaiA activation most likely involves direct protein–protein interactions of the N-terminal domain (with an as-yet-unidentified protein) that result in functional modulation of the C-terminal effector domain. We postulated that KaiA is a two-domain response regulator that acts as a timing input device to control the phase of the circadian period by modulating the KaiC phosphorylation state (9).

Here we report the solution structure of the ~12-kDa KaiC-interacting C-terminal KaiA domain (Fig. 5, which is published as supporting information on the PNAS web site) enabled us to solve the structure at the physiological temperature of 50°C. We also determined the structure of S. elongatus KaiC by electron microscopy (EM) and single-particle reconstruction to a resolution of 21 Å. The structures raise interesting proposals for the mode of KaiA–KaiC interaction.

Materials and Methods

Protein Preparation. Samples of S. elongatus KaiC and full-length KaiA were prepared as described (9). Expression and purification of ThKaiA180C are described elsewhere (18). NMR samples consisted of ~1.3 mM protein in a buffer containing 20 mM NaCl and 20 mM Na2HPO4, pH 7.07, at 23°C. Heterodimeric (13C/15N) samples of ThKaiA180C were prepared by mixing equimolar amounts of oxidized 13C/15N-enriched and -unenriched protein in 6 M GdHCl/50 mM NaCl/20 mM Tris-HCl, pH 7.0, with the addition of 100 mM β-mercaptoethanol. Complete reduction of the intermolecular disulfide bond was monitored by SDS/PAGE. The protein was reconstituted by slow dilution with a 20-fold excess volume of 50 mM NaCl/20 mM Tris-HCl, pH 7.0, buffer and exchanged to the final NMR buffer and reoxidized. The final protein concentration of the heterodimeric samples was a 1.0 mM ThKaiA180C dimer (0.5 mM heterodimer plus 0.25 mM fully enriched homodimer plus 0.25 mM unenriched homodimer). Protein concentration was determined by absorbance at 280 nm by using a corrected extinction coefficient as suggested by Pace et al. (19).

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: NOE, nuclear Overhauser effect; EM, electron microscopy.

Data Deposition: The solution structure has been deposited in the Protein Data Bank, www.rcsb.org (PDB ID codes 1Q6A and 1Q6B for the average minimized structure and the 25-structure ensemble, respectively).

§To whom correspondence regarding circadian aspects should be addressed. E-mail: golinden@tamu.edu.

© 2004 by The National Academy of Sciences of the USA
EM. N-terminal His-tagged KaiC (0.1 mg/ml) in 300 mM NaCl/40 mM TrisHCl, pH 7.5/1 mM ATP was negatively stained with 1% (wt/vol) uranyl acetate on a glow-discharged carbon grid. Images were recorded in a JEOL 1200 transmission electron microscope operating at 100 kV and at a calibrated magnification of ×38,900. Micrographs were digitized with a Leafscan 45 microdensitometer (Leaf Systems, Southborough, MA) at 20-μm increments corresponding to 5.14 Å per pixel at the specimen level. Single-particle analysis was performed by using the EMAN software package (28). About 1,000 particle images were used for the final reconstruction. The volume of the particle was rendered based on the calculated molecular mass of 344.5 kDa and an average density of 1.37 Da/Å³. The resolution of the 3D reconstruction was estimated to be ~21 Å by Fourier shell correlation (Fig. 6B, which is published as supporting information on the PNAS web site).

Figure Preparation. Figures were prepared by spock (29). The three N-terminal non-KaiA-derived residues of ThKaiA180C are not shown in the figures.

Results

The ThKaiA180C Domain. KaiA is the cyanobacterial clock protein for which the greatest divergence in sequence exists among different cyanobacterial groups (9, 30). This divergence appears primarily in the N-terminal pseudo-receiver domain; the C-terminal domain is highly conserved among all KaiA proteins known to date (Fig. 7, which is published as supporting information on the PNAS web site). For instance, the S. elongatus and T. elongatus N- and C-terminal domains share 30% and 62% identical residues, respectively. Furthermore, there are KaiA proteins that lack the N-terminal domain altogether (Anabaena sp. PCC 7120) or have only a truncated version (Nostoc punctiforme). The highly conserved C-terminal domain does not show significant sequence similarity to any other protein or derived domain from any organism studied. This simultaneous high degree of conservation in cyanobacteria and absence in other organisms suggests that the C-terminal domain of KaiA serves a function specific to this system.

The C-terminal domain of the S. elongatus KaiA was unfavorable for NMR structure determination because of limited aliphatic proton chemical shift dispersion and fast exchange of amide protons. In addition, thermal denaturation experiments showed little cooperativity of unfolding (Fig. 5). In contrast, the C-terminal domain of KaiA from T. elongatus has greater chemical shift dispersion and displays excellent cooperativity (Fig. 5). Thus, we proceeded to solve the structure of the T. elongatus KaiA C-terminal domain, which by sequence identity should be virtually equivalent to that of S. elongatus.

We previously showed that the N-terminal domain of S. elongatus KaiA is monomeric in solution even at high concentrations (9). Full-length S. elongatus KaiA, however, elutes off of a calibrated gel filtration column at a molecular mass higher than the monomeric 32 kDa (data not shown). Analytical ultracentrifugation equilibrium of full-length KaiA showed a single species with molecular mass consistent with a dimer (Fig. 8, which is published as supporting information on the PNAS web site). Additionally, NMR relaxation measurements performed at 50°C on ThKaiA180C yielded average 15N T2 values for ordered residues (15N(−1H) NOE values >0.6; Fig. 9C, which is published as supporting information on the PNAS web site) of ~70 ms and a calculated total correlation time, τc, of ~9.4 ms (Fig. 9B), which are consistent with a molecular species of ~24 kDa, corresponding to the ThKaiA180C homodimer. Only a single set of resonances was observed for each residue, which indicates that ThKaiA180C exists as a symmetric homodimer on the NMR timescale.

Structure Determination. The solution structure of ThKaiA180C was calculated from a total of 2,207 experimental restraints per monomeric subunit (Table 1). There were 50 unambiguous intersubunit distance restraints, most identified through 13C-edited/13C-filtered 3D NOE 1H−1H−1H−1H NOE) distance restraints were obtained by 13C-edited 4D NOE (100-ms mixing time), 15N-edited 3D NOE (100-ms mixing time), and 13C-edited/13C-filtered 3D NOE (21) (120-ms mixing time) spectra. Hydrogen bond restraints were applied based on hydrogen exchange protection data collected by NMR and the existence of expected regular secondary structure NOEs (20). Backbone dynamics measurements (15N T1, 15N T2, and 15N(−1H) NOE) were performed and analyzed as described (22). Spectrometer temperature was calibrated with a methanol sample.

NMR Structure Calculations. ϕ and ψ dihedral angle values were derived from TALOS (23). Only residues for which all 10 predictions lie in the same region of the Ramachandran plot were used. The XPLOR-NIH software package (24) was used for all stages of NMR structure calculations. A distance-geometry simulating annealing protocol in Cartesian space was used to generate an initial set of structures from NOE and dihedral angle data. After inspection, the structures were refined by a simulated slow-cooling process and a final minimization. During refinement, additional restraints were used, including a conformational database potential (25), direct refinement against 13Cα and 13Cβ chemical shifts (26), and a radius of gyration restraint (27). The latter was applied independently to each monomeric subunit for residues of the globular core (inclusive, residues 186–204, 215–243, and 257–275) with a calculated value of 10.94 Å for the radius of gyration. Noncrystallographic symmetry was enforced between the two monomeric subunits throughout the calculation because only one set of resonances was observed. However, to better sample possible conformations of the intermolecular disulfide bond, C272 was excluded from this symmetry term. The family of structures consists of 25 low-energy structures that satisfy all experimental restraints (Table 1, which is published as supporting information on the PNAS web site).
experience significant internal motion as indicated both by the lower number of long-range $^1$H–$^1$H NOEs and heteronuclear $^{15}$N($-^1$H) NOE values $<0.6$. In contrast, the loop connecting $\alpha_2$ with $\alpha_3$ has an average $^{15}$N($-^1$H) NOE value of 0.75 and shows contacts both in the same monomeric subunit and across the dimer interface (see below).

The hydrophobic core of the monomeric subunit is mostly contained in the region where the four helices pack against each other. It is composed of 11 hydrophobic residues whose sidechains are completely buried, three of which are aromatic (Fig. 7). An additional 11 nonpolar residues pack around this core, thereby increasing the total hydrophobic area buried. The hydrogen bonding network is mostly restricted to residues that are close in the primary sequence (e.g., the $i$ to $i+4$ hydrogen bonds of $\alpha$-helical segments). Only four hydrogen bonds were inferred from the structure to occur between residues far apart in the sequence. These hydrogen bonds are between $\alpha_1$ and $\alpha_2$ ($O'$ of L189 and N$^\alpha$ of K221), between $\alpha_1$ and $\alpha_3$ ($O'$ of Y197 and N$^\alpha$ of H236) and between $\alpha_3$ and $\alpha_4$ ($O'$ of Y197 and N of L264, and between O$^\gamma$ of Y204 and the O$^{\delta1}$ of D266). Stability and fold specificity is further enhanced through a number of intra- and interhelical salt-bridges (e.g., between residues R187 and E234 of $\alpha_1$ and $\alpha_3$). Sequence alignment between the C-terminal domains of KaiA from T. elongatus and S. elongatus shows that all buried residues are identical or conservatively substituted (Fig. 7). Similarly, three of the four long-range hydrogen bonds inferred from the structure involve residues conserved between these two organisms. In contrast, many of the favorable electrostatic interactions in T. elongatus KaiA are eliminated in S. elongatus through substitutions of charged residues with uncharged polar ones. This difference could account for the significant reduction in stability and cooperativity of unfolding observed in the C-terminal domain of S. elongatus KaiA (Fig. 5).

**The dimer interface.** The dimer interface of ThKaiA180C (Fig. 2B) involves almost the entire $\alpha_4$, the N-terminal part of $\alpha_3$, and the loop connecting $\alpha_2$ with $\alpha_3$. Formation of the dimer buries $>900$ Å$^2$ of accessible surface area. Residues I265 to A269 of $\alpha_4$ pack against residues I265′ to A269′ of $\alpha_4'$ and the N-terminal part of $\alpha_3'$.

Beyond A269 the dimer interface resembles that of a coiled-coil between $\alpha_4$ and $\alpha_4'$ extending from C272 to I279. The first interaction in this region is an intersubunit disulfide bond connecting residues C272 and C272′. It is worth noting that C272 is conserved in all but one of the known KaiA proteins (Fig. 7). Because no noncrystallographic symmetry requirement was applied on C272, the calculated disulfide geometry was an asymmetric intermediate of the left-hand-spiral and right-hand-hook disulfide conformations (31), both in terms of average $\psi$ separation and sidechain dihedral angles adopted. The $^1$H resonances of C272 appear significantly broader in all of the NMR spectra collected, likely because of localized motions on a $\mu$s-to-ms timescale.

The $\alpha_4$–$\alpha_4'$ coiled-coil is further stabilized by an intersubunit salt-bridge between residues R276 and D226′. In addition, O$^{\delta2}$ of E273 was inferred to participate in a bifurcated hydrogen bond with O$^\gamma$ of S228′ and N of V229′. However, these hydrogen bonds do not occur in all of the structures of the ensemble; therefore, their presence is not certain. Although E273 is conserved among all KaiA molecules known, S228 is unique to T. elongatus KaiA.

**Possible Function of ThKaiA180C.** Dimer formation creates a groove between $\alpha_3$–$\alpha_3$ and $\alpha_4$–$\alpha_4'$ (dashed arrow in Fig. 3), thereby exposing several residues to solvent. Most are polar (D241, T242,
S244, K245, K248, S253, E254, D255, and R261) and a couple are hydrophobic (L249 and L258). A number of substitutions in KaiA that cause altered circadian phenotype in *S. elongatus* map at this exposed region, and others map either at or near the hydrophobic core or the dimer interface (Fig. 3) (32). The substitutions at the hydrophobic core and dimer interface likely alter normal circadian rhythmicity by destabilizing the structural integrity of the C-terminal domain of KaiA. We postulate that the solvent-exposed groove directly interacts with KaiC, and substitutions within that region weaken the interaction between the two proteins, thereby altering the clock mechanism.

**EM of KaiC.** We calculated a 21-Å-resolution hexameric KaiC structure from EM data collected on negatively stained *S. elongatus* KaiC by the single particles method (Fig. 6). KaiC appears to form a dumbbell-shaped symmetric hexamer with a ∼50-Å radius and a 100-Å height (Fig. 4A), in agreement with what has been previously reported (12, 13). KaiC has two domains (CI, residues 1–250, and CII, residues 252–519) (6), both of which interact with KaiA in vitro. Our electron microscope reconstruction of the KaiC particle clearly shows two globular domains connected by an ∼10-Å linker region. Openings on both ends of the hexameric structure (∼29-Å top opening diameter, ∼23-Å bottom opening diameter) can be seen in our reconstruction of the *S. elongatus* KaiC particle, in agreement with data from Mori *et al.* (13). However, an electron microscope reconstruction by Hayashi *et al.* (12) of the *T. elongatus* hexameric KaiC features a hole only at the top of the particle. The latter structure also does not display holes at the side of the particle, although significant depressions can be seen. The KaiC sequence in these two strains is highly conserved (81% identical residues, 92% similar residues); therefore, we do not expect that the differences observed in the electron microscope reconstructions reflect real differences in the protein structures but, rather, a difference in the analysis and interpretation of the data. The presence of two openings in our structure is consistent with experimental data that show binding of forked DNA by KaiC and those author’s proposal of the structure (13).

**Discussion**

The overall fold of the ThKaiA180C monomeric unit can be described as a right-turning, four-helix bundle (33) with similar topology to that of previously studied proteins, such as cytochrome *b₅₆₂* (34). However, the wide (∼54° on average) angle

![Fig. 2.](image)

**Fig. 2.** (A) Selected ¹³C–¹H strips from the ¹³C-edited/¹³C-filtered 3D NOESY spectrum (21) identifying NOEs at the dimer interface of ThKaiA180C. The spectrum was acquired in a heterodimeric sample of unenriched and ¹³C/¹⁵N-enriched protein. (B) The dimer interface of ThKaiA180C. Shown here is α₄ of one subunit in magenta, and in blue the C-terminal region of α₂, the loop between α₂' and α₃, and the N-terminal region of α₃' and α₄' are shown. Residues identified as being part of the dimer interface by the ¹³C-edited/¹³C-filtered 3D NOESY and the intersubunit disulfide bond and salt-bridge are also shown.

![Fig. 3.](image)

**Fig. 3.** Accessible surface area representation of the ThKaiA180C structure and residue substitutions producing an altered circadian oscillator phenotype identified in *S. elongatus* KaiA as mapped at this domain. The accessible surface area for each substituted residue is colored in magenta or blue for the two monomeric subunits. The groove formed between α₃ and α₄ and α₃' and α₄' is indicated by a dashed arrow. Several of the circadian phenotype substitutions map at this groove.
search algorithms (38–40) did not find any other stand-alone proteins of this class in the Protein Data Bank. This finding is in contrast to the structure of the N-terminal KaiA domain, which was found to adopt a very common α/β fold (9). In addition, the quaternary homodimeric structure of ThKaiA180C is unusual, because dimerization involves a combination of coiled-coil and helix-in-helix interactions.

Because the C-terminal domain of KaiA has a groove between the two monomeric subunits and many circadian phenotype substitutions map in that region, we propose that this groove of KaiA interacts with KaiC by bending the narrow linker connecting the two globular KaiC domains (putative CI and CII). This possible arrangement is shown in Fig. 4B, with one KaiA C-terminal domain bound to KaiC in the manner we propose and a second KaiA C-terminal domain indicating the proposed mode of approach. It is worth noting that our proposal is consistent with previous studies that identified putative KaiA-binding domains in KaiC close to the CI–CII interface (41) and provides a structural form-fitting basis for the mode of interaction.

In vivo studies have shown the stoichiometry of KaiA to KaiC to be ~1:20 (14) (corresponding to a stoichiometry of 1 KaiA dimer per 6.7 KaiC hexamers). Although this would seem to indicate that, on average, KaiA–KaiC complexes involve a single KaiA dimer, it is still possible that the active species of this complex has a higher KaiA:KaiC stoichiometry. Therefore, the KaiC phosphorylation state may be controlled simply through variation of the number of KaiA dimers bound to the hexameric KaiC particle or their complete absence. Similarly, amino acid substitutions in KaiA with altered circadian phenotype could attenuate KaiA stimulation of the KaiC autokinase activity, either through disruption of the KaiA structure or by weakening the KaiA–KaiC interaction. Notably, the phenotypes are longer free-running periods or arrhythmia (32), which suggests that a reduced ability of KaiA to enhance the KaiC phosphorylation rate directly lengthens the circadian period.

The size of the clock protein complexes varies in vivo during the circadian cycle (7). KaiA alternates between a complex of molecular mass approximately equal to that of the KaiA dimer during the subjective day and a complex with a mass >500 kDa during the subjective night. We expect that this latter complex contains one or more KaiA dimers bound to the hexameric KaiC particle in addition to other clock components, such as KaiB and SasA. Therefore, a possible mechanism of circadian oscillator phase resetting emerges, in which protein–protein interactions at the N-terminal pseudo-receiver domain are followed by signal propagation to the C-terminal KaiC-interacting domain of KaiA (9). The conformational change at the C-terminal domain would affect the strength of the KaiA–KaiC interaction, thereby altering the number of KaiA dimers bound to the KaiC particle and, thus, the phase of the oscillator.

We thank Dr. Satoshi Tabata (Kazusa DNA Research Institute, Chiba, Japan) for providing T. elongatus DNA. Dr. Karl Kosklaph for technical assistance with the NMR spectrometers, and Shannon Canales for helpful comments on the manuscript. Funding was provided to A.C.L. and S.S.G. by National Institutes of Health Grants GM064576 and GM062419, respectively. A.H. thanks the Office of the Vice President for Research for financial support. The NMR instrumentation in the Biomolecular NMR Laboratory at Texas A&M University was supported by National Science Foundation Grant DBI-9970232.

A

B

Fig. 4. (A) Twenty-one-angstrom-resolution EM structure of the hexameric Synechococcus elongatus KaiC particle. The structure was derived from analysis of ~1,000 single hexameric KaiC particle images. The ~29Å opening at the top of the structure and the separation between the two globular KaiC domains (putative CI and CII) can be seen. (B) Our proposed model of KaiA–KaiC interaction is shown. Depicted here is one C-terminal KaiA dimer in a molecular surface area representation modeled as “bound” to KaiC at the interface linker between the putative CI and CII domains. A second C-terminal KaiA dimer is also shown, and an arrow indicates the proposed approach.

References


Vakonakis et al.