Life before the Clock: Modeling Circadian Evolution

Till Roenneberg and Martha Merrow
Institute for Medical Psychology, Chronobiology Division, Ludwig-Maximilians-Universität München, D-80336 Munich, Germany

Abstract A feedback loop that functions via transcription and translation is thought to be the mechanistic core of circadian rhythmicity. Numerous modeling efforts incorporate the identified components and their modifications to recreate the circadian clock in computer simulations. Several issues remain problematic, including the lack of precise quantitative kinetics and the likely existence of additional, as-yet-undiscovered components. Even without these complications, models and flow charts of the circadian system have reached high complexity. They attempt to reconcile all observations without violating current views and concepts. In this article, the authors consider the mechanisms that may have preceded the circadian system in evolution. Given that cellular metabolism and biochemistry were presumably already interconnected in cascading feedback reactions prior to the appendage of the transcription/translation feedback loop, a coordinated response to exogenous changes would be advantageous over unsystematic responses. The authors hypothesize that those mechanisms that allowed synchronization in spite of metabolic complexity form the basis for the evolution of circadian properties and are as fundamental to the circadian system as the transcriptional/translational feedback loop.

Key words circadian rhythm, model, evolution, negative feedback

In 1729, the French astronomer DeMairan discovered that leaf movements in plants continue with a 24-h rhythmicity in constant darkness. Since then, the description of the mechanism responsible for endogenous, circadian rhythms has come a long way. Mainly from the work of Büning, Aschoff, and Pittendrigh, we know the formalisms and rules of circadian timing. The anatomical location of the central circadian pacemaker is known for most animals (e.g., the suprachiasmatic nucleus [SCN]), whereas plants appear to lack central pacemakers. Circadian rhythmicity is a property of most cells, and a network of clock genes is being discovered.

The current conceptual view of the molecular circadian clock mechanism is based on the hypothesis that circadian rhythmicity is generated by a negative feedback loop that is closed via the expression of clock genes and their protein products (a transcriptional-translational loop [TTL]). This model was originally hypothesized for the Drosophila period gene (Hardin et al., 1990) and was strengthened by the fact that the Neurospora clock gene, frequency, and its protein also
form a negative feedback (Aronson, Johnson, Loros, et al., 1994). The circadian feedback loops described in cyanobacteria (Iwasaki and Kondo, 2000; Mori and Johnson, 2001), Neurospora (Loros and Dunlap, 2001; Merrow et al., 2001), Drosophila (Blau, 2001; Williams and Sehgal, 2001), mammals (Herzog and Tossini, 2001; Reppert and Weaver, 2001), and more recently in plants (Alabadi et al., 2001) use different sets of genes. However, even loops that use the same components (like in insects and mammals) are put together differently or the respective genes and their products serve different functions. There is no doubt about the important role of these clock genes in each of the circadian systems. Without them, circadian rhythmicity lacks many of its characteristics, including self-sustainability in constant conditions (Bargiello et al., 1984; Gekakis et al., 1998; Loros and Feldman, 1986), entrainability by light (Albrecht et al., 2001; Lakin-Thomas and Brody, 2000; Merrow et al., 1999; Stanewsky et al., 1998; Zheng et al., 2001), and temperature and nutritional compensation (Loros and Feldman, 1986). The known molecular feedback loops in different species are all tightly linked to light input pathways, which has led to the assumption that the circadian clock evolved from light transduction mechanisms (Crosthwaite et al., 1997). Yet, experiments in Neurospora and Arabidopsis show that other circadian characteristics (e.g. entrainment by temperature cycles) remain in the absence of clock genes (McWatters et al., 2000; Merrow et al., 1999), so additional components of the system are still to be identified—most probably in all model systems. A recent study strengthens the case for additional components. In cockroaches, aftereffects can only be elicited when the insects are kept for several days in light:dark cycles but not if they are submitted to temperature cycles (Page et al., 2001). These results indicate that a light-sensitive part may be distinct from a temperature-sensitive part of the circadian machinery. Our knowledge about the former is much more detailed than it is about the latter, so that each additional component will increase our understanding of the circadian machinery, especially if it opens up cellular mechanisms outside of the known TTLs.

One of the most conspicuous qualities of circadian systems is self-sustained oscillation in constant conditions. This feature was instrumental in the discovery of the circadian clock, in proving the rhythm’s endogenous nature, and in identifying its genetic components. But what do we learn from experiments in constant conditions in view of the fact that the evolution of the circadian system took place predominantly under the influences of a rotating earth? Theoretically, any damped oscillator mechanism would serve just as well with zeitgebers counteracting decreasing amplitudes. It is likely that self-sustained rhythmicity is a consequence of how the system is built to function optimally under entrained conditions rather than being the object of the evolutionary process (Roenneberg and Merrow, 2001). But, what does “optimal function” of the circadian system mean in the context of a 24-h day? Circadian systems are entrained rather than driven by zeitgebers. Unlike driven synchronization, entrainment results in a specific phase relationship to the zeitgeber cycle, which depends on the system’s endogenous period and its phase response curve as well as on the amplitude of both the zeitgeber and the endogenous oscillator. Entrainable systems are taken to be advantageous over purely driven ones by providing the possibility of anticipation, that is, preparing physiology before the external changes occur (sunrise, drop in temperature, food availability, etc.). Entrainable systems can vary their phase relationship with a zeitgeber cycle simply by changing the endogenous period or the strength by which the phase-resetting mechanism responds to a given zeitgeber (amplitude of the phase response curve). This provides flexibility in adjusting to different external changes. Although the sun rises at the same time each day (or predictably earlier/later within the course of the year), the timing of food sources might not always be directly linked to the earth’s rotation. Yet, food availability may still be predictable from its previous spatiotemporal history. A physiology that is simply driven by the external changes would, thus, be disadvantageous relative to one that is regulated by an endogenous rhythmicity with flexible and anticipatory qualities (Daan, 1981).

An inherent assumption in this line of arguments is that the circadian system was implemented as an “improvement” over purely driven systems. However, complex networks of metabolism, which must have existed in the most primitive organisms, may not be simply drivable by environmental cycles; they may have had to develop mechanisms to counteract chaotic responses of a complex intracellular network to exogenous zeitgebers. If one presumes that such a mechanism is necessary for reliable synchronization with the environment, organisms lacking this mechanism would be at a disadvantage in a periodic world. Therefore, the foundation on which the circadian clock is built might be those mechanisms that allow
physiology to be driven in spite of complexity. We have formalized these assumptions by modeling a circular network of feedback loops.

**METHODS**

Modeling biological functions can have very different goals, each with different merits for understanding an initial “back-box system.” Either they may attempt to mimic experimental results as closely as possible (Leloup and Goldbeter, 1998) or they can be a formal basis for hypotheses (e.g., the existence of singularity) (Winfree, 1970) that can then be verified experimentally (Taylor et al., 1982). Here, we use mathematical modeling to explore a possible scenario of circadian evolution. The model is essentially a coupled oscillator system. Several models of the circadian system have used multiple coupled oscillators, for example, to predict the interaction by global coupling between neurons within the SCN and their contribution to the in-toto phenotype (Achermann and Kunz, 1999). Other coupled oscillator models investigate the interaction between bilateral pacemakers (Petri and Stengl, 2001), between subpopulations of SCN neurons, or between the evening and morning oscillator (Daan et al., 2001; Oda et al., 2000). This article does not attempt to compare its algorithms with previous multioscillator models. Its major difference is that the underlying oscillators, in isolation, do not have periods in the circadian range and are not self-sustained. The model connects five oscillators in a circuit, which couples the individual components either by direct coupling (i.e., the product of one of the feedback reactions influences the kinetics of the next) or by a global coupling (i.e., the summed output of each feedback loop influences one of the components).

Models were implemented with the help of the Stella® Program for Macintosh, and model results were further analyzed and graphed with the help of the CHRONO program (Roenneberg and Taylor, 2000). All feedback loops ($FB_1$ to $FB_5$; see Fig. 1A) used in the network are based on the same set of equations describing the concentration changes of two state variables ($S_1$ and $S_2$) over time (see Table 1). The two “reactions” form a negative feedback loop by $S_2$ inhibiting the production of $S_1$. The algorithms used for the feedback loops are essentially the same as used in the previously published single feedback models (Roenneberg and Merrow, 1998, 1999):

$$S_1(t) = S_1(t - dt) + (p_1 - d_1) \cdot dt$$

$$S_2(t) = S_2(t - dt) + (p_2 - d_2) \cdot dt.$$

The model was run at time intervals ($dt$) of 1 h (using the fourth-order Runge-Kutta integration algorithm). The production of $S_1$ ($p_1$; column 2 in Table 1) is inhibited by $S_2$, whereas the production of $S_2$ ($p_2$; column 3 in Table 1) depends directly on the concentration of $S_1$. The degradation of each state variable depends on its current concentration, that is, the higher its concentration, the stronger its degradation ($d_1$ and $d_2$; columns 5 and 6 in Table 1). The feedback loops differ only in their turnover rates of $S_1$ ($r$; influencing both production and degradation; column 1 in Table 1).

A circular network between the five feedback loops is implemented by direct coupling. State variable ($S_i$) of one loop influences the production of state variable ($S_j$) in the next loop: $S_{2,1}$ influences $p_{1,2}$, $S_{2,2}$ influences $p_{1,3}$, ..., $S_{2,5}$ closes the circle by influencing $p_{1,1}$ ($c$; column 4 in Table 1).

$FB_1$ constitutes the input component of the network and is influenced by the zeitgeber (ZG), which affects the production of $S_{1,1}$ by adding a rate to $p_{1,1}$ for 12 h each 24 h (see figures for the different zeitgeber strength used). An integrated output of the network is modeled by simply adding the concentrations of all $S_2$ for a given time point ($\Sigma$; see the bottom trace in Fig. 1B). The input component, $FB_1$, is influenced not only by ZG but also via feedback by the zeitnehmer (ZN) ($ZN$ represents the integrated output, $\Sigma$, multiplied by the zeitnehmer strength, $I_{ZN}$; see equation [3]). The degradation of $S_{2,1}$ is increased by adding ZN from the previous time interval (i.e., with a delay of 1 h; see column 6 in Table 1 and figures for the different zeitnehmer strength, $I_{ZN}$, used):

$$ZN = I_{ZN} \cdot \Sigma (S_{2,1} \text{ to } S_{2,5}).$$

**RESULTS AND DISCUSSION**

**Metabolic Networks**

Some kind of negative feedback is the basis of most oscillations, but negative feedback alone does not ensure self-sustained rhythmicity. In addition, negative feedback is one of the most common control mechanisms within pathways at all levels (e.g., product inhibition in enzymatic reactions, gene regulation,
or neuronal circuits). Any metabolic network must, therefore, involve numerous negative feedback loops.

The five feedback loops in our model (Fig. 1A) are based on the same set of equations involving the pro-

Figure 1. Rhythmicity and synchronization of a model representing a theoretical organism with complex metabolism but without a circadian system. (A) Five negative feedback (FB) loops (for model algorithms) exemplify the numerous FBs in different parts of metabolism. The five FBs are identical except for the turnover of state variable $S_1$. (B) Individually and uncoupled, each of the FBs shows an oscillation that damps more or less rapidly. The top five time series in B represent concentration changes of $S_2$ in each of the five FBs. The bottom trace represents the sum of all $S_2$ concentrations at a given time ($\Sigma$). (C, D) When connected to a circular network, $\Sigma$ is arrhythmic (see linear and double plot representation in D).
duction and degradation of two state variables ($S_1$ and $S_2$) (Roenneberg and Merrow, 1998, 1999). The two state variables could, for example, represent RNA and protein in a TTL or any other two-step negative feed-back. The turnover rates of $S_1$ (see $r$ in Table 1) are given different values based on the rationale that individual TTLs or metabolic feedback loops are bound to have different qualities. When the five loops are not connected in a network, $FB_1$ to $FB_5$ show different kinetics of a dampened oscillation with distinct periods in the ultradian range (Fig. 1B). In our model, we have chosen a network of five feedback loops, but similar results can be achieved with three (data not shown). In presuming that each of the feedbacks is directly or indirectly involved in changing a metabolic variable (e.g., $Ca^{2+}$ or redox) (Johnson et al., 1995; Merrow and Roenneberg, 2001; Rutter et al., 2001), an integrated oscillation was calculated by simply adding the concentrations of each state variable $S_2$ at any given time (Fig. 1C and $\Sigma$ in Fig. 1B).

When the individual feedback loops are connected in a network, they become self-sustained (i.e., the integrated output does not converge to a steady state as does each of the individual loops when not connected). The system’s periodicity, however, is unstable—still in the ultradian range. A long-term recording of the integral of the $S_2$ values ($\Sigma$) shows a chaotic oscillation insofar as the pattern never repeats (tested for 200 “days”; data not shown). The recording resembles the activity patterns of arrhythmic organisms (Fig. 1D). Arrhythmic activity (e.g., Ewer et al., 1990; Kume et al., 1999; Ralph et al., 1990) is generally thought not to be generated by underlying circadian or ultradian oscillators. However, if some readout occurs in bouts, even at irregular intervals, some underlying mechanism must be responsible for these episodes. The “actogram” in Figure 1D supports the notion that oscillators that by themselves are regular but dampened can form a network where they become irregular but self-sustained. In fact, very early on in analyzing the activity patterns of per$^{0}$ flies, it became clear that the arrhythmicity was much more systematic than could be expected from inspection (Dowse et al., 1987).

### Synchronizing Networks to a Cyclic Environment

Given an appropriate input pathway, each of the simple, unconnected feedback loops is driven by a zeitgeber (as shown for $FB_1$ in Fig. 2A). When the feedbacks form a network, however, neither the individual oscillations nor the integrated rhythm is drivable with any stable output (classical protocols (i.e., single perturbations given at different circadian times of a free run in constant conditions).

The results shown in Figure 2 indicate that the complex network is in a different state every time the zeitgeber goes on or off, due to its inherent dynamics, leading to a different response and, thus, resulting in unreliable synchronization. One possibility to counteract unreliable synchronization of the network is to

### Table 1

Variables used to implement the feedback loops, $FB_1$ to $FB_5$.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>$FB_1$</td>
<td>100</td>
<td>$ZG + r*e^{-0.025(S_{2,1} + c)}$</td>
<td>$S_{1,1}$</td>
<td>$S_{2,5}$</td>
<td>$2/3,r + 0.018,S_{1,1}$</td>
<td>$ZN + 4 + 0.018,S_{2,1}$</td>
</tr>
<tr>
<td>$FB_2$</td>
<td>200</td>
<td>$r*e^{(-0.025(S_{2,2} + c))}$</td>
<td>$S_{1,2}$</td>
<td>$S_{2,1}$</td>
<td>$2/3,r + 0.018,S_{1,2}$</td>
<td>$4 + 0.018,S_{2,2}$</td>
</tr>
<tr>
<td>$FB_3$</td>
<td>300</td>
<td>$r*e^{(-0.025(S_{2,3} + c))}$</td>
<td>$S_{1,3}$</td>
<td>$S_{2,2}$</td>
<td>$2/3,r + 0.018,S_{1,3}$</td>
<td>$4 + 0.018,S_{2,3}$</td>
</tr>
<tr>
<td>$FB_4$</td>
<td>400</td>
<td>$r*e^{(-0.025(S_{2,4} + c))}$</td>
<td>$S_{1,4}$</td>
<td>$2S_{2,3}$</td>
<td>$2/3,r + 0.018,S_{1,4}$</td>
<td>$4 + 0.018,S_{2,4}$</td>
</tr>
<tr>
<td>$FB_5$</td>
<td>500</td>
<td>$r*e^{(-0.025(S_{2,5} + c))}$</td>
<td>$S_{1,5}$</td>
<td>$2S_{2,4}$</td>
<td>$2/3,r + 0.018,S_{1,5}$</td>
<td>$4 + 0.018,S_{2,5}$</td>
</tr>
</tbody>
</table>

NOTE: $r$ = turnover rate, $p_1$ = production rate of state variable $S_1$, $d_1$ = degradation rate of state variable $S_1$, $c$ = connectivity between the individual feedbacks, $p_2$ = production rate of state variable $S_2$, $d_2$ = degradation rate of state variable $S_2$, $ZG$ = zeitgeber, $ZN$ = zeitnehmer.
make the input pathway itself (represented in our model by $FB_1$) sensitive to the integrated oscillation ($\Sigma$; Fig. 1C). This additional, global feedback modulates the component of the network that receives the zeitgeber (time giver) and, thus, serves as zeitnehmer (time taker) (Roenneberg et al., 1998). By closing the zeitnehmer loop, the system becomes stably entrained (compare panels E and F with B and C in Fig. 2). Note, that even with a closed zeitnehmer loop, the system remains arrhythmic in constant conditions (data not shown, but similar to the zeitnehmer-less network shown in Fig. 1D)—clearly not an accurate representation of a circadian system. The stabilizing effect of a zeitnehmer is also seen when all feedback loops are given the same ($r = 100$; data not shown), rather than different, turnover rates (see Table 1).

Stable synchronization depends on zeitgeber strength. Systems obviously do not synchronize when a zeitgeber is too weak, but they may also lose synchronization when it is too strong (Gonze and Goldbeter, 2001). Does the zeitnehmer loop stabilize the network so that synchronization occurs over a wider range of zeitgeber strengths, or does it simply reduce the impact of the zeitgeber? We studied the effects of different zeitgeber and zeitnehmer strengths systematically. At high zeitgeber amplitudes, even the zeitnehmer-less network can be synchronized; however, concentrations increase steadily over time until they oscillate around a high mean level (Fig. 2D). This trend is absent when the zeitnehmer loop is closed. Due to the inherent kinetics of the network, synchronization breaks down at intermediary zeitgeber
strength (around $ZG = 65$; Fig. 3A). This breakdown can be prevented by increasing the impact of the zeitnehmer loop (Fig. 3B).

From Synchronization to Entrainment

Evolution is thought to proceed in small steps rather than in big jumps (Dawkins, 1997). We propose that the ancestral components from which a circadian clock evolved consisted of mechanisms that enabled complex cellular networks to be driven by environmental cycles. From these, only small changes may have been necessary to turn a drivable network that is arrhythmic in constant conditions (Fig. 1D) into an entrainable, self-sustained circadian system (Fig. 4). In our model, we exemplified this small “evolutionary” step by simply strengthening the connectivity between $FB_2$ and $FB_1$ (changing it from 1 to 6). Under these conditions, the network becomes self-sustained in the circadian range (with an average $\tau$ of 17.3 h; Figs. 4A, 4B). With stronger connectivity in parts of the network, the system is even less apt to be driven without the compensating effect of a zeitnehmer loop (Fig. 4C). When the zeitnehmer loop is closed, the period of the free-running integrated rhythm depends on zeitnehmer strength—the stronger its impact, the shorter the period (data not shown). This leads to the fact that the entrained network adopts different phase angles with different zeitnehmer strengths and even shows transients typical for circadian systems (Figs. 4D, 4E).

In the model used here, changes in entrainment phase can be achieved simply by varying the connectivity between the individual feedback loops of a complex network or by altering the strength of the zeitnehmer loop (Figs. 4D, 4E). The circadian systems of both day- and night-active organisms (e.g., mouse and man) are apparently built by the same components, and even within the same organism, day and night activity can change depending on conditions. Day-active migratory birds, for example, which travel long distances at night, change their activity phase at the appropriate times of the year (Gwinner, 1996), and humans, which are “larks” both as a child and as an adult, tend to be “owls” during adolescence (Carskadon et al., 1998). It is likely that such developmental or history-dependent changes in entrainment phase are due to physiological/metabolic changes rather than to changes in the genetic clock components. In our model, these changes are implemented by changing the impact of networked components on
each other, resulting in different production and degradation rates. The model suggests that the circadian qualities currently attributed to a single TTL could also be implemented by multiple feedback loops with noncircadian periods connected in a network. The idea that a circa-24-h rhythmicity may be a product of coupled ultradian oscillators was put forward approximately 25 years ago (Dowse and Ringo, 1987) but has been overwhelmed by the apparent power in explaining circadian qualities by the interactions of a handful of clock genes in a TTL (Darlington et al., 1998; Reppert and Weaver, 2001). Now that different clock mutants are being submitted to more rigorous physiological analysis, we are gathering more detailed information about exactly which circadian qualities are defective in a given clock mutant or in combinations of clock mutants, and multiple oscillator models may be revisited.

**Experimental Predictions**

Models such as this one are only useful if they produce experimental ideas. With weaker direct coupling in parts of the circuit, the model simulates an arrhythmic organism (Fig. 2), whereas increasing coupling strength leads to an output resembling that of a circadian organism (Fig. 4). The main difference between the two model versions lies in the fact that the latter consolidates a circadian rhythmicity in the absence of a zeitgeber whereas the former does not. Under entrained conditions, the two models behave similarly, except for the typical circadian transients
present in both the self-sustained models. Circadian research has traditionally based many of its experiments on the self-sustainment in constant conditions, whereas the presented model makes a strong case for experiments using systematic entrainment protocols to probe for the mechanisms underlying circadian behavior. The necessity of employing entrainment conditions in investigating circadian systems was often stressed by Jürgen Aschoff.

If circadian systems were built on mechanisms that ensure complex metabolism to be drivable by zeitgebers, then the components of this mechanism can be identified in organisms without self-sustained circadian rhythmicity. Such organisms include those that (1) did not evolve a circadian system, (2) express self-sustained rhythmicity only in as-yet undiscovered outputs or only under specific constant conditions that have not been found, or (3) have been “made” arrhythmic by mutation. The first two classes are difficult to separate because the lack of self-sustained circadian rhythmicity may only be due to the fact that the right conditions or the right output variables have not yet been discovered. Cyanobacteria and Caenorhabditis elegans are good examples of organisms that were historically regarded as “clockless” but were subsequently found to be circadian (Johnson et al., 1996; Kippert et al., 2002; Saigusa et al., 2002). Our model results predict that circadian components should exist in all three classes of arrhythmic organisms.

The model predicts that organisms that are arrhythmic in constant conditions show circadian properties under entrained conditions. Systematic phase relationships should be measurable in any of the clockless organisms for changing both zeitgeber lengths and strengths, and arrhythmicity should be conditional in many of them. These predictions can easily be tested in organisms of all three classes mentioned above, and some confirming results already exist. Systematic phase angles in zeitgeber cycles of different amplitudes and lengths have already been shown for clock mutants in Neurospora (Merrow et al., 1999), and release experiments show remaining circadian qualities in Arabidopsis clock mutants (McWatters et al., 2000). Neurospora clock mutants that are initially arrhythmic develop a self-sustained rhythm when grown on long race tubes (Aronson, Johnson, and Dunlap, 1994; Loros and Feldman, 1986), and mouse clock mutants are arrhythmic in constant darkness but not in constant light (Steinlechner et al., 2002). Interestingly, the model shown in Figure 2 can also be made self-sustained when constant light is simulated (data not shown).

In addition, mutagenesis of putative clockless organisms or further mutagenesis of clock mutants should reveal additional components beyond the known TTLS. A reporter gene could be used to monitor key metabolic outputs, for example, in temperature cycles. Prior to mutagenesis, expression patterns under zeitgeber conditions should be similar to those shown in Figures 2E and 2F (reliable synchronization). Following mutagenesis, one would screen for mutants with a phenotype similar to those shown in Figures 2B and 2C (unreliable synchronization). Once genes have been identified that enable complex metabolic networks to be synchronized by zeitgebers, their homologues can be deleted in organisms with an intact circadian system to investigate their role in the circadian machinery. Because several of these components may be important for both the (extended) clock system and for viability, strategies have to be applied to prevent lethal mutations (e.g., overexpression). With growing possibilities of coping experimentally with complex systems, we will be able to incorporate components that contribute to the circadian phenotype in a less legotype fashion than before.

ACKNOWLEDGMENTS

Our work is supported by the Eppendorf Company, Hamburg, the Deutsche Forschungsgemeinschaft, the Dr.-Meyer-Struckmann-Stiftung, and the Friedrich-Baur-Stiftung.

REFERENCES


