

# Research Statement - Richard Yamada

UNIVERSITY OF MICHIGAN

<yryamada@umich.edu>

## Summary

My current research lies at the intersection of two disciplines: biology and mathematics. Recent advances in the biological sciences has necessitated the development of quantitative approaches to help interpret and analyze the wealth of experimental data. **My focus is on developing and using mathematical methods to understand the role of transcription regulation in various physiological processes.** At the same time, I strive to ensure that the models I develop are based on experimental reality. Currently, I am focusing on two main areas: (i) *elongation dynamics of transcription* (ii) *transcription regulation of circadian oscillators*. I employ quantitative methods in order to make predictions that can be experimentally verified. My career goal is to make lasting contributions on our understanding of biological processes through the use of mathematics.

## Current Research Projects

### 1. Elongation Dynamics of Transcription

RNA polymerase is the key enzyme of transcription, the step at which most regulation of gene expression occurs. Transcription consists of three distinct processes: initiation, elongation and termination. Of these processes, elongation is most amenable to a quantitative analysis. My doctoral dissertation [27] focused on developing a quantitative model for elongation dynamics of the RNA (Ribonucleic acid) polymerase enzyme during transcription. Understanding transcription is important because it is the process at which most regulation of gene expression occurs. During transcription genetic information in the form of double stranded DNA (Deoxyribonucleic acid) is converted to single stranded RNA; the molecular machine responsible for coordinating the events of transcription is the RNA polymerase enzyme. Transcription consists of three distinct stages: initiation, elongation, and termination. Much is known about initiation and termination, but little is known about elongation, the step where a rNTP (ribonucleotide triphosphate) is added to the nascent RNA chain. Conventional biochemical methods cannot study the movement of RNA polymerase along DNA during elongation; furthermore, the study of elongation requires a quantitative approach, since it is inherently a dynamical process. The introduction of novel biophysical tools [1, 12], however, has produced a wealth of experimental data, making it possible to test predictions from quantitative models of elongation.

My work develops the first chemical kinetic model of transcriptional elongation [26, 27, 28]. The model's novel concept is a *look-ahead* feature, in which nucleotides bind reversibly to the DNA prior to being incorporated covalently into the nascent RNA chain. The proposed model governs the discrete movement of the RNA polymerase along a DNA strand. A master-equation formulation together with a computational simulation using the Gillespie algorithm studying the behavior of the look-ahead model are introduced. Furthermore, several approaches to parameter estimation are tested on synthetic and also on actual experimental data. The most important feature of my model is its testable predictions, which can be verified with future experiments. By replicating the data analysis algorithms from the experimental procedure, the model produces velocity histograms, enabling direct comparison with published experimental results [1, 12]. Two additional experimental data are also considered: 1) the mean velocity of RNA polymerase as a function of the ambient concentrations of the ribonucleoside triphosphates; and 2) the distribution of time intervals between the forward steps of RNA polymerase. In particular, based on the parameter fitting of the latter type of data, we find that the data strongly supports a

*look-ahead* window that is approximately four bases in width.

From an evolutionary standpoint, the following question is of pertinent interest: what is the advantage of a *look-ahead* window? In principle, elongation is the step where most errors during transcription would occur. Is there a way to control or limit the number of errors during elongation? The formulation of the look-ahead model allows for a natural way to reduce errors during transcription. Recently, we have used mathematical analysis and computer simulation to study the rate at which transcriptional errors occur as a function of  $w$ , the window size of the *look-ahead* model. We find dramatic reduction in the error rate of transcription as  $w$  increases, especially for small values of  $w$ . The error reduction method provided by look-ahead occurs *before* hydrolysis and covalent linkage of rNTP to the nascent RNA chain, and is therefore distinct from error correction mechanisms that have previously been considered [29].

## 2. Transcription Regulation in Mammalian Time-Keeping

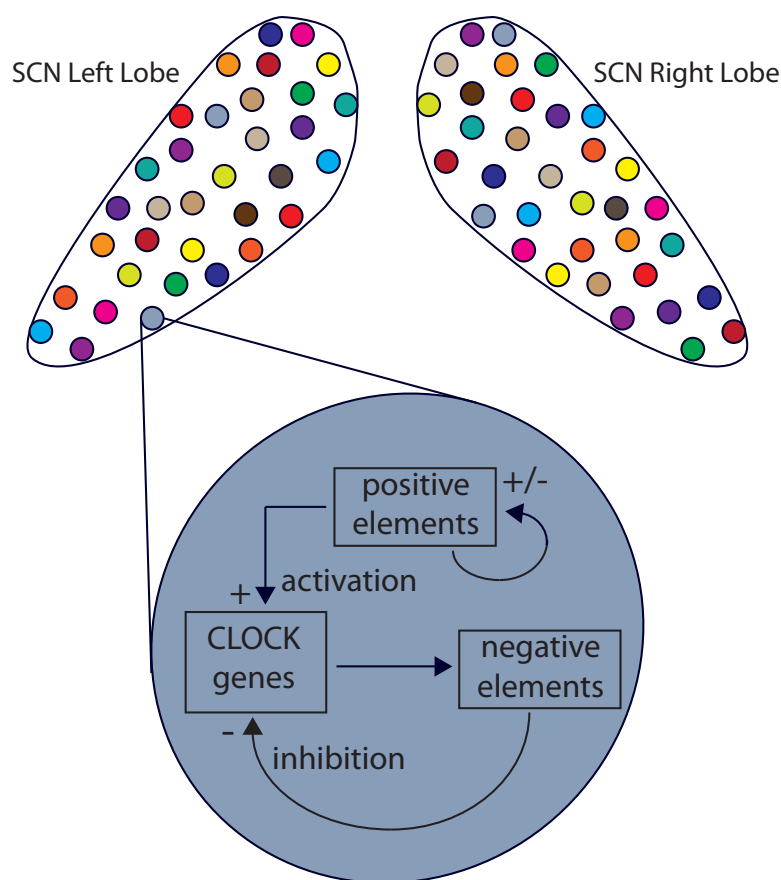


Figure 1: The Suprachiasmatic Nucleus (SCN) consists of 2 lobes of approximately 10,000 cells. Though each cell is an autonomous oscillator, governed by the presence of transcription feedback loops, there is a variation in the time-keeping of each cell, due to the presence of molecular noise. Thus no two cells will have the exact same timekeeping dynamics; inter-cellular coupling restores collective rhythms in the entire SCN.

Mammalian time-keeping is controlled by the Supra-Chiasmatic Nucleus (SCN), the master clock controlling all 24 hour (circadian) activity. Understanding the circadian clock mechanism of the SCN is important, because its abnormal function has been linked with many diseases and disorders, including

depression and insomnia. The SCN consists of 10,000 individual cells or autonomous oscillators. Each cell is governed by a complex biochemical network [13]. In mammalian cells, the basic physiology of circadian rhythms involves a negative feedback loop linking the events of transcription and translation. Furthermore, my work has shown that transcription regulation plays a central role in mammalian circadian clocks [14, 30]. From a mathematical standpoint, modeling the SCN is a challenging multi-scale problem, with many potential research opportunities [30].

My recent work has focused on studying an interesting mathematical question raised by Enright [8]: Can self-sustained rhythms in an "ensemble of sloppy circadian oscillators" arise even if the individual oscillators were damped? Enright answered this question in the affirmative without incorporating any knowledge of the detailed cellular mechanisms, that have since been discovered, into his model. First, I developed the most detailed single cell model of the SCN to date by extending the Forger-Peskin model for mammalian time-keeping. Important modifications of this model include 1) binding/unbinding of the activator complex to the specific repressor promoters 2) regulation of transcription activators by the repressor protein, CRY 3) use of published degradation rates of key molecular components. This model contains 263 coupled differential equations along with 30 free parameters; the large number of equations reflect the combinatorial complexity of all possible core clock component interactions in the model. A stochastic formulation of this same model was also developed using the Gillespie algorithm; this formulation allows the exploration of molecular noise in time-keeping. Next, using this single cell model as the individual oscillator, I coupled these cells together in an all-to-all manner, exploring both linear and non-linear coupling mechanisms.

Using this detailed model, experimental predictions of a key core clock component, BMAL, were tested and verified by the Takahashi Lab (Howard Hughes Medical Institute, UT-Southwestern Medical Center). We find that varying the level of the BMAL core clock protein acts as a tuning parameter, which determines whether stable oscillations in SCN network will occur. Furthermore, our model hypothesizes that when BMAL levels in individual cells is 20% of that in wildtype cells, these cells are not rhythmic; however when these cells are coupled together with a nonlinear mechanism, the rhythms are recovered, albeit with shorter periods and higher variance. From this detailed study involving BMAL, we have answered Enright's hypothesis in the affirmative. These findings, which involve a close interaction between theory and experiments, were published recently in [14]. More importantly, these simulations point to the hypothesis that molecular noise is amplified by a sensitive inter-cellular coupling mechanism that can kick start oscillations within the network; at the same time, sensitive transcriptional regulation mechanisms that promote rhythmicity are also likely to amplify the effects of noise.

From this detailed study of mammalian time-keeping in network, two key mathematical concepts have emerged: molecular noise and inter-cellular coupling. While detailed models can make testable hypotheses (and therefore of considerable use to biologists), they also are difficult to analyze and gain further intuition of the underlying dynamics. Put another way, biological phenomena can inspire models that can be simplified (as 'toy models') for further analysis. Thus, in order to fully understand the surprising results in [14], we are studying a simplified model that contains the key characteristics of the detailed model; in this way we can better understand the link between inter-cellular coupling and transcription. With undergraduates Kara Fulton and Lars Johnson, we have written the differential equations for a single cell model, and have also developed a stochastic version using the Gillespie algorithm. We are now looking at the effect of several inter-cellular coupling mechanisms to see how transcription of the repressor is being regulated. In another direction, I am investigating the role of molecular sequestration in promoting stable oscillations. While the concept of sequestration (or quenching) is not new in biochemistry [21], its novel role in maintaining stable time-keeping has been recently reported in the literature [7, 19]. Sequestration is a transcription regulation mechanism in which the repressor *sequesters* an activator, before it can bind to the DNA and activate transcription of the gene. We are currently writing a paper on the significance of this mechanism in promoting stable oscillations within circadian oscillators.

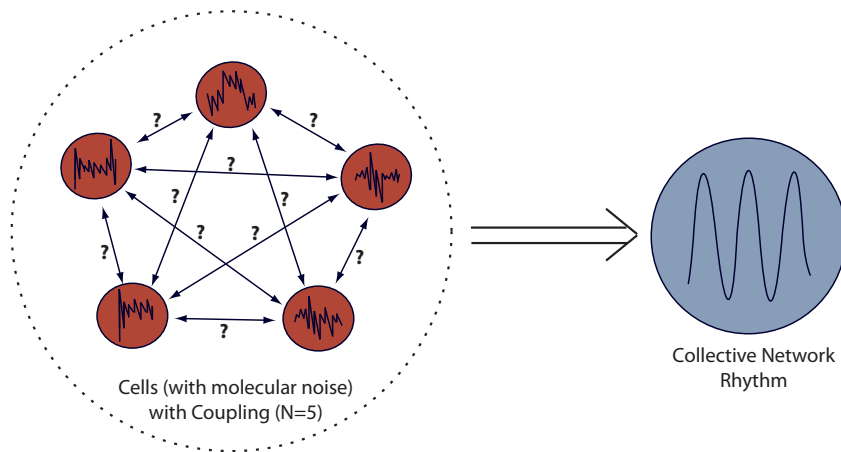


Figure 2: Here, individual cells in the SCN are coupled together in an all-to-all manner. The exact inter-cellular coupling mechanism is much debated. However, our work [14] has shown that both molecular noise, along with non-linear inter-cellular coupling, is crucial for generating rhythms in the SCN.

## Future Research Projects

### 1. Elongation Dynamics in Time-Keeping

My goal is to understand the importance of elongation in other physiological processes. In particular, I am interested in extending and applying the *look-ahead* model as a transcription regulation mechanism in the time-keeping of SCN cells [26, 27, 28]. New experimental data points to the possibility of a *look-ahead* feature in the RNA polymerase [12]. Most models of transcription do not model this process in detail [9, 10, 15]; instead, the complex process of transcription is simplified and reduced to a single mathematical term. Recent experiments, however, have indicated the importance of elongation dynamics in gene expression, leading to transcription pulsing [24]. This process can regulate the transcription-translation loop, ultimately affecting the sensitive time-keeping mechanism found in SCN cells.

### 2. Polymerase Backtracking During Transcription Elongation

One of the most interesting aspects of the RNA polymerase is its observed behavior during transcription elongation. Experiments indicate that polymerase motion is not continuous; instead, the motion is discontinuous, with intermittent bursts of rapid elongation, pausing and stopping. The current hypothesis is that long pauses allows the polymerase to correct for transcriptional errors, while short pauses are either random or due to another mechanism. The look-ahead model does not consider backwards motion of the polymerase that have now been observed in experiments. Thus, I am interested in incorporating this feature to study whether it will change the waiting time distribution.

### 3. Detailed Kinetic Models of Transcription Elongation

Currently I am working with the Burton Biochemistry Lab (Michigan State University, Biochemistry) and Prof. Peiru Wu (Michigan State University, Mathematics) to develop a detailed kinetic model of

transcription elongation based on recent experimental evidence. This data indicates that a look-ahead feature (pre-loading of rNTPs) is indeed a strong candidate for the mechanism of elongation dynamics. Using a detailed kinetic model, we are modeling the exact loading progression of a specific 4 base sequence of rNTPs. The experimental data in this project is not based on single molecule biophysics (unlike my work in [26, 27, 28]), but on detailed kinetic data using quench and pulse-chase techniques from the Burton Lab.

#### **4. Physical Limits to Biochemical Time-keeping**

Many important physiological processes operate with small numbers of molecules; thus molecular noise Bialek and colleagues have been studying the physical limits in biochemical signaling [5, 6]. In particular, they have re-investigated the work of Berg and Purcell [4] and have generalized their arguments to estimate physical limits to signaling processes within the cell. The SCN maintains accurate 24 hour rhythms with little variance. A natural question to ask is how this precision is possible, given the noisy nature of transcription in individual oscillators. I am interested in extending these ideas to understand whether and how circadian oscillators operate within the physical limits of nature. Furthermore, I want to understand the role of inter-cellular communication in maintaining precise time-keeping.

#### **5. Role of Electrophysiological Firing on Transcription in SCN Cells**

There has been considerable interest in understanding the connection between the electrophysiology of SCN cells and its affect on circadian gene expression. Electrophysiological events in the SCN occur at the millisecond time-scale whereas the biochemical clock, mediated by the transcription-translation negative feedback loop, occur at a 24 hour time-scale. Successful models have been developed separately for the electrophysiology of SCN cells [3, 23] and for the molecular clock [14]. Thus, the next goal would be to develop a model that combines the molecular core clock with the electrophysiology of SCN cells, with the goal of helping biologists guide their thinking about how best to experimentally test this connection. With Dr. Casey Diekman, we are developing a multi-cell model that combines the most detailed model of a mammalian SCN cell [14] with the most detailed model of the developed in [3]. We stress that the main goal of our model is to make testable experimental predictions.

#### **6. Role of SCN in Other Physiological Processes**

Experimental evidence points to a central role of time-keeping in many important physiological processes. These processes include (but are not limited to): metabolism, cell cycle, pituitary regulation, and development [2, 17, 20, 22, 31]. In particular, the link between circadian clocks and metabolism is particularly promising [2]. The role of the SCN, and furthermore the role of transcription within SCN cells in regulating these physiological processes remains to be elucidated. Currently, with Jae-Kyoung Kim, I am pursuing the development of a detailed circadian oscillator model involving core metabolic components. Insights obtained from this model will also help understand the role of circadian rhythms in diseases such as diabetes [18] and cancer [11].

### **Conclusion**

In conclusion, I aim for the tight integration of mathematics and biology in my research problems. I plan to search for new research opportunities by focusing on the subtleties inherent in biological phenomena and collaborating closely with experimental colleagues. From a mathematical point of view, problems arising from biological phenomena inspire the development of new (both numerical and analytical) tools. Only by working closely with experimentalists and developing mathematical models capable of positing experimental predictions will we make true progress on understanding the most interesting and complex problems in the biological sciences.

## References

- [1] E. Abbodanzieri, W. Greenleaf, J. Shaevitz, R. Landick, and S. Block. Direct observation of base-pair stepping by RNA polymerase. *Nature*. 2005. 438:460-465.
- [2] G. Asher, D. Gatfield, M. Stratmann, H. Reinke, et al. SIRT1 regulates circadian clock gene expression through PER2 deacetylation. *Cell*. 2008. 134:317-328.
- [3] M. Belle, C. Diekman, D. Forger, and H. Piggins. Daily electrical silencing in the mammalian circadian clock. *Science*. 2009. 326:281-284.
- [4] H. Berg and E. Purcell. Physics of chemoreception. *Biophys. J.* 1977. 20:193-219.
- [5] W. Bialek. Stability and noise in biochemical switches. *Advances in Neural Information Processing 13*. 2001. 103-109.
- [6] W. Bialek and S. Setayeshgar. Physical limits to biochemical signaling. *Proc. Natl. Acad. Sci.* 2005. 102:10040-10045.
- [7] R. Chen, A. Schirmer, Y. Lee, H. Lee, et al. Rhythmic PER abundance defines a critical nodal point for negative feedback within the circadian clock mechanism. *Mol. Cell* 2009. 36:417-430.
- [8] J.T. Enright. Mutual Excitation of Damped Oscillators and Self-Sustainment of Circadian Rhythms. In *Mathematical Models of the Circadian Sleep-Wake Cycle*, C. Czeisler and M. Moore-Ede editors, pages 1-15, New York, NY, USA, 1984. Raven Press.
- [9] D. Forger and C. Peskin. A detailed predictive model of the mammalian circadian clock. *Proc. Natl. Acad. Sci.* 2003. 100(25):14806-14811.
- [10] D. Forger and C. Peskin. Stochastic simulation of the mammalian circadian clock. *Prof. Natl. Acad. Sci.* 2005. 102(2):321-324.
- [11] L. Fu, H. Pelicano, J. Liu, P. Huang and C. Lee. The circadian gene Period2 plays an important role in tumor suppression and DNA damage response in vivo. *Cell*. 2002. 111:41-50.
- [12] E. Galburt, S. Grill, A. Wiedmann, L. Lubkowska, et al. Backtracking determines the force sensitivity of RNAP II in a factor-dependent manner. *Nature*. 2007. 446:820-823.
- [13] C. Ko and J. Takahashi. Molecular Components of the Mammalian Circadian Clock. *Hum. Mol. Genet.* 2006. 15:R271-R277.
- [14] C. Ko, Y.R. Yamada, D. Welsh, E. Buhr, et al. Emergence of noise-induced oscillations in the central circadian pacemaker. *PLoS Biology*. 2010. 8(10):e1000513.
- [15] J. Leloup and A. Goldbeter. Toward a detailed computational model for the mammalian circadian clock. *Proc. Natl. Acad. Sci.*. 2003. 100(12):7051-7056.
- [16] A. Liu, D. Welsh, C. Ko, H. Tran, et al. Intercellular coupling confers robustness against mutations in the SCN circadian clock network. *Cell*. 2007. 129:605-616.
- [17] S. Lim, L. Pnueli, J. Tan, Z. Naor, et al. Negative feedback governs gonadotrope frequency-decoding of gonadotropin releasing hormone pulse-frequency. *PLoS ONE*. 2009. 4(9):e7244.

- [18] B. Marcheva, K. Ramsey, E. Buhr, Y. Kobayashi, et al. Disruption of the clock components CLOCK and BMAL1 leads to hypoinsulinaemia and diabetes. *Nature*. 2010. 466:627-631.
- [19] J. Menet, K. Abruzzi, J. Desrochers, J. Rodriguez, and M. Rosbash. Dynamic PER repression mechanisms in the *Drosophila* circadian clock: from on-DNA to off-DNA. *Genes Develop*. 2010. 24:358-367.
- [20] Y. Nakahata, S. Sahar, G. Astarita, M. Kaluzova, and P. Sassone-Corsi. Circadian control of the NAD<sup>+</sup> salvage pathway by CLOCK-SIRT1. *Science*. 2009. 324:654-657.
- [21] M. Ptashne. *Genes and Signals*. Cold Spring Harbor. 2001.
- [22] K. Ramsey, J. Yoshino, C. Brace, D. Abrassart, et al. Circadian clock feedback cycle through NAMPT-mediated NAD<sup>+</sup> biosynthesis. *Science*. 2009. 324:651-654.
- [23] C. Sim and D. Forger. Modeling the electrophysiology of the suprachiasmatic nucleus neurons. *Jour. Biol. Rhy*. 2007. 22(5):445-453.
- [24] I. Swinburne, D. Miguez, D. Landgraf, and P. Silver. Intron length increases oscillatory periods of gene expression in animal cells. *Genes & Dev*. 2008. 22:2342-2346.
- [25] S. Yoo, C. Ko, P. Lowery, E. Buhr, et al. A Noncanonical E-box enhancer drives mouse Period2 circadian oscillations in vivo. *Proc. Natl. Acad. Sci*. 2005. 102(7):2608-2613.
- [26] Y.R. Yamada and C. Peskin. 2006. A chemical kinetic model of transcriptional elongation. 2006. q-bio.BM/0603012.
- [27] Y.R. Yamada. Quantitative models of transcriptional elongation. *Ph.D. Thesis*, Cornell University, Ithaca, NY. 2007. 173 pages.
- [28] Y. R. Yamada and C. Peskin. 2009. A look-ahead model for the elongation dynamics of transcription. *Biophys. J*. 2009. 96:3015-3031.
- [29] Y.R. Yamada and C. Peskin. 2010. The influence of look-ahead on the error rate of transcription. *Math. Mod. Nat. Phen*. 2010. 5(3):206-227.
- [30] Y.R. Yamada and D. Forger. 2010. Multi-scale complexity in the mammalian circadian clock. *Accepted, Curr. Opin. Genet. Develop*.
- [31] J. Zamborszky, C. Hong, and A. Nagy. Computational analysis of mammalian cell division gated by a circadian clock: quantized cell cycles and cell size control. *Jour. Biol. Rhy*. 2007. 22(6):542-553.