UNIVERSITY OF CALIFORNIA SAN DIEGO

Signaling Pathways in Living Systems

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

in

Physics with a Specialization in Quantitative Biology

by

Aravind Rao Karanam

Committee in charge:

Wouter-Jan Rappel, Chair
Daniel P. Arovas
Elena F. Koslover
Julian I. Schroeder

2024

The Dissertation of Aravind Rao Karanam is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2024

DEDICATION

Dedicated to all my teachers who put faith in me, nudged me to do the harder things, and expanded the realm of what is achievable.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to thank my advisor and the chair of the committee Dr. Wouter-Jan Rappel for nurturing and challenging me as a scientist, for his critical and timely feedback, and most of all for his patience and investment in my growth. I also thank Dr. Richa Karmakar, Dr. Man-Ho Tang, Dr. Timothy Tyree, Dr. Dorsa Elmi, Dr. Mahesh Mulimani, and other members of the Rappel Lab for their many valuable inputs, instructive discussions, and an unwavering support.

I would also like to acknowledge the support of Prof. Julian Schroeder, members of the Schroeder lab and David He for a fruitful, albeit remote, collaboration through the pandemic.

I would also like to thank the physics department at UC San Diego for providing a welcoming and stimulating work environment and an invaluable opportunity to hone my teaching skills. Special thanks to the Quantitative Biology program and Prof. Terence Hwa for introducing me to biophysics and supporting me in the initial years of the program.

Chapter 1, in part, is a reprint of the material as it appears in Plant Physiology 2021 and Quantitative Plant Biology 2022. Karanam, A., He, D., Hsu, P.K., Schulze, S., Dubeaux, G., Karmakar, R., Schroeder, J.I. and Rappel, W.J., 2021, Plant Physiology 2021, and Karanam, A. and Rappel, W.J., 2022, Quantitative Plant Biology 2022. The dissertation author was the primary investigator and author of both papers.

Chapter 2, in full, is a reprint of the material as it appears in the Physical Review E 2021. Karmakar, R., Tang, M.H., Yue, H., Lombardo, D., Karanam, A., Camley, B.A., Groisman, A. and Rappel, W.J., American Physical Society, 2021. The dissertation author was a co-author of this paper.

Chapter 3, in full, is a reprint of the material as it is submitted for publication and as it may appear in the Physical Review Letters, 2024. Karmakar, R., Karanam, A.R., Tang, M.H. and Rappel, W.J., American Physical Society, 2024. The dissertation author was the primary investigator and author of this paper.

VITA

2016	Bachelor of Science (Research), Indian Institute of Science, Bangalore
2016–2017	Master of Science, Indian Institute of Science, Bangalore
2018–2023	Teaching Assistant, Department of Physics University of California San Diego
2018–2024	Research Assistant, University of California San Diego
2024	Doctor of Philosophy, University of California San Diego

PUBLICATIONS

"Cellular memory in eukaryotic chemotaxis depends on the background chemoattractant concentration." Physical Review E 103.1 (2021): 012402.

"Boolink: a graphical interface for open access Boolean network simulations and use in guard cell CO_2 signaling." Plant Physiology 187.4 (2021): 2311-2322.

"Boolean modelling in plant biology." Quantitative Plant Biology 3 (2022): e29.

"Eukaryotic chemotaxis under periodic stimulation shows temporal gradient dependence." Submitted to Physical Review Letters, archived at bioRxiv(2023): 2023-10.

FIELDS OF STUDY

Major Field: Physics with Specialization in Quantitative Biology

Studies in Mathematical Methods
Professor Massimo Vergassola

Studies in Classical Mechanics
Professors Daniel Dubin and Massimiliano Di Ventra

Studies in Electromagnetism
Professor Michael Fogler

Studies in Statistical Physics
Professor Daniel Arovas

Studies in Quantum Mechanics

Professors Elizabeth Jenkins and Kenneth Intrilligator

Studies in Quantitative Physics
Professor Thomas Murphy

Studies in Fluid Dynamics
Professors Patrick Diamond and Massimo Vergassola

Studies in Information Theory
Professor Massimo Vergassola

Studies in Population Dynamics
Professor Massimo Vergassola

Studies in Physics of the Cell
Professor Elena Koslover

ABSTRACT OF THE DISSERTATION

Signaling Pathways in Living Systems

by

Aravind Rao Karanam

Doctor of Philosophy in Physics with a Specialization in Quantitative Biology

University of California San Diego, 2024

Wouter-Jan Rappel, Chair

Signaling networks are at the heart of most biological processes. In this dissertation, two classes of models with applications are discussed in the following chapters.

In most signaling networks with several components, either the connections between the components or the parameters governing the reaction kinetics are not known. Given this uncertainty, Boolean networks, in which each component is either on or off, have emerged as viable alternatives. Open-source platforms of Boolean models for community use are desirable. Here, we present Boolink, a freely available graphical user interface that allows users to easily construct and analyze existing Boolean networks. We demonstrate its application using a previously published network for abscisic acid (ABA)-driven stomatal closure in *Arabidopsis*

thaliana, and by extending the network to include CO_2 regulation of stomatal movements. Predictions of the model were experimentally tested, and the model was iteratively modified based on experiments showing that ABA effectively closes Arabidopsis stomata at near-zero CO_2 concentrations.

When cells of the social amoeba *Dictyostelium discoideum* are starved of nutrients they synthesize, secrete, and relay the chemical messenger and chemoattractant cyclic Adenosine Mono Phosphate (cAMP), resulting in the establishment of periodic waves. Cells aggregate through chemotaxis towards the center of these waves. In the process, they experience an elevated background concentration of cAMP as well as multiple waves of a fixed period. We investigated in two separate studies the effect of these two changes, using waves of cAMP generated by a microfluidic device.

We found that the chemotactic ability of the cells increases for small to moderate levels but collapses to zero for sufficiently high concentrations. Secondly, we found that the chemotactic ability of cells rises with the number of waves encountered by the cells provided the wave period is not large. We developed mathematical models to explain the observed trends, building on earlier work on the Local Excitation Global Inhibition (LEGI) class of models. We showed that a temporal gradient sensing mechanism underlies the wave-period-dependent rise in the chemotactic ability. The observed trends in the chemotactic ability are relevant to *Dictyostelium* in aiding its aggregation.

Introduction

Modeling in Biology

"All models are wrong, but some are useful." – George Box

The purpose of scientific theories is to create a coherent view of the world around us that can, ultimately, be refined by testing the predictions from the theories against experiments. Since theories are by design generalizable, the object or phenomenon a theory seeks to describe is an abstraction of its whole. By abstracting out the irrelevant details, a model captures the essential facts in the phenomenon it describes. But the abstraction ultimately limits the applicability of the model. Treating the Sun, the Earth and the Moon as point masses is sufficient to calculate the periods of revolution but this abstraction cannot explain tides though both are governed by the laws of gravity.

An useful model is explanatory, quantitative, and predictive. The phenomenon the modeler seeks to explain should arise out of the dynamics of the components at a lower level of complexity. The model should be able to make quantitative and falsifiable predictions, which can be checked against new observations. When one sets out to model biological phenomena at the cellular level, the spatiotemporal and energy scales situate its constituent processes in the realms of chemical reactions, classical mechanics, and non-equilibrium thermodynamics. Elementary reactions that synthesize and modify the chemical composition of cells, transport processes, elastic and inelastic deformations, and modes of exchange of matter and energy with the surroundings are the building blocks of higher order form and function.

The evolutionary history shared by organisms across kingdoms means that many bio-

chemical components, motifs, and their functions are conserved. This leads to two important consequences for the modeler: one, there is no need to reinvent the wheel for every system; the underlying biochemistry remains the same. Two, life processes can be treated as modular, and the complexities in form and function arise out of interactions among the modules.

Decades of systematic studies in biology identified the building blocks of life – such as DNA sequences, RNA and protein, metabolic and signaling molecules – and established pathways in which matter is manipulated and information flows. In the next stage, one would like to quantify this knowledge. This is desirable for a few reasons: quantification puts into perspective the extent and the relative importance of individual processes, which is a critical input to modeling; the nature of interactions can, in some cases, be inferred from their kinetics. When we know which variables are critical and which are not, for an effective description of the phenomenon at hand, we are able to not only formulate simpler models that can easily be tested and refined but also obtain ways to manipulate the system for medical or industrial needs.

In the following, I discuss two model paradigms and the underlying biological phenomena they are applied to. First, I describe a Boolean reaction network modeling the intracellular processes in guard cells in leaves. Through changes in their shape, guard cells regulate the size of a tiny pore between them known as stomata through which gas exchange takes place between the atmosphere and the interior of the leaf. Then, I describe two studies on chemotaxis in the social amoeba *Dictyostelium* and show how the motion of the cells can be described by models of gradient sensing using differential equations.

Boolean Modeling of Guard Cells in Plants

To model a biological pathway, one needs to know the identity of the components and the topology of their connections. In pathways where strengths of interactions are known in the form of rate constants, one can construct a detailed model consisting of rate equations for each component [1, 2]. If a few constants are unknown and the number of components in the model is not high, the values of the unknown constants can be estimated using data fitting algorithms[3, 4].

In both cases, the forms of the interaction functions (the right hand side of the rate equation) and the rate constants yield valuable information on the nature of the pathway.

The level of detail as noted above may not always be available as some parameters may be hard to measure. If parameters are collectively fit against experimental data, care should be exercised in interpreting the parameters, as individual parameters often poorly constrained [5]. Moreover, the desired outcome of a model could just be qualitative (high/low levels of concentration, or a process being active/inactive) instead of quantitative. Boolean models are a viable alternative to modeling in these situations [6].

A Boolean model is a network, or a graph, of interactions among the constituent components, or nodes, in which each component can be in one of two states at any given time: active (1 or ON) or inactive (0 or OFF). The state of a component depends on the states of the components it is connected to. The exact dependence is given by the so-called update equations, that are composed of elementary Boolean functions NOT, OR, and AND. The update equations do not contain rate constants. Indeed, this is one of the strengths of the Boolean modeling approach in that one need not input rate constants, which are often difficult to measure from experiments; nor one does need to perform a cumbersome fitting process to obtain a plausible set of parameters for a reaction network model. The output of a Boolean network is qualitative in that we only know the state of the output (active or inactive) when each of the input variables is active or inactive. Despite these significant simplifications, Boolean models have been shown to provide insights into genetic [7, 8, 9, 10], protein [11], and cellular regulatory networks [12, 13].

At the base of plant leaves exist tiny pores known as stomatae, through which gas exchange takes place between the interior of the leaf and the atmosphere. The opening or closing of each stomata is regulated by a pair of guard cells which change their shape to constrict or dilate the stomata. Stomatal closure can be triggered by external factors like light and carbon dioxide (CO_2) [14, 15]. Drought results in accumulation of the plant hormone Abscisic Acid (ABA) in guard cells, which also leads to stomatal closure [16, 17]. The reaction network of stomatal closure triggered by ABA, established through hundreds of biochemical studies, is quite

large: it has more than eighty components and more than 150 connections [18]. Modeling the entire set of reactions as differential equations is not only expensive but also impractical, given that the kinetic parameters of many reactions are unknown. For these reasons, the signaling network in guard cells lends itself to Boolean modeling.

We developed a general framework for simulating and analyzing any user-defined Boolean network through a software package named Boolink [19], which includes a graphical user interface (GUI) to execute the commands and to visualize the network and plot the simulation results. We used Boolink to simulate a published model of the ABA network [18], reproduced its findings, and extended it to include the effect of CO_2 either by itself or in conjunction with ABA. Though several components are shared between the two closure pathways, some components are unique, and the effects of the inputs on intermediate nodes could be different [20, 21, 22]. Through an iterative process of modeling and simulation, we proposed a pathway that correctly predicts the effect of ABA and CO_2 acting individually (one in the absence of the other) or together [19].

Gradient Sensing Models in Eukaryotic Chemotaxis

Chemotaxis is the directed motion of cells in response to chemical signals in their environment. It is shown to be involved in several critical processes such as wound healing, embryonic development, cancer metastasis, etc [23, 24, 25, 26, 27]. Cells chemotax by either moving up the gradient of a *chemoattractant* or down the gradient of a *chemorepellent*. The physics of chemoreception, beginning from the dynamics of the chemoattractant binding to the cell-surface receptors[28, 29] to the cascade of reactions that leads to the cell motion, either by the rotation of flagellar motors or the deformation of the cell membrane, has been the subject of several theoretical and experimental investigations [30, 31, 32, 33].

We investigated the chemotactic ability of the eukaryotic model organism, the social amoeba *Dictyostelium discoideum*. A population of these cells undergo a transition from a unicellular to multicellular state when deprived of nutrients. They achieve this by aggregating

through chemotaxis using a small molecule cyclic Adenosine Mono Phosphate (cAMP) as a chemoattractant[34]. Each cell in the population acts as a relay of cAMP; it synthesizes and secretes more cAMP into its surroundings after receiving it through its receptors. At a population level, this relay process manifests itself as the propagation of a chemoattractant wave, with successive peaks in cAMP concentration separated by a time period T of $6-10$ minutes[35, 36, 34]. At the same time, cells chemotax towards the source by moving up the cAMP gradient.

When a *Dictyostelium* cell encounters a wave, the concentration of cAMP in its surroundings rises from a small value to the peak amplitude in the front half and decreases to the earlier value in the back half. If a cell were to only follow the local gradient of the chemoattractant, it would oscillate with zero net displacement over one wave period. But cells do have a way of ignoring or being less sensitive to the gradient in the back half of the wave. Careful observations of isolated chemotaxing cells have shown that cells display a positive (towards the source of the wave), zero, or a negative (away from the source of the wave) displacement while experiencing the falling concentration of cAMP [3, 37, 33]. The development of an internal asymmetric response to an externally symmetric rise and fall in cAMP concentration, and the development of a directional memory, albeit with a dependence on the wave period, was modeled by a Local Excitation Global Inhibition (LEGI) reaction network motif coupled to a bistable memory module (M)[3, 37]. Similar memory phenomena have also been reported in other biological systems, including chemotactic neutrophils [38, 39]. Biomolecules that are plausible candidates for the components of the LEGI model are idetified[40]: the response element corresponds to the activated form of Ras, Ras-GTP, which ultimately leads to cellular displacement through membrane deformation; the activator and the inhibitor are the enzymes that respectively catalyze the addition and removal of an activated phosphate group to Ras. A candidate for the bistable memory is yet to be identified.

To further our understanding of the cellular response to chemoattractant waves, we performed two studies. In the first study, we examined the effect on the chemotactic ability

of cells upon adding a non-zero concentration of cAMP in background besides presenting the cells with the chemoattractant wave[4]. In the second study, we examined the effect on the chemotactic ability of cells upon exposure to multiple waves of a given wave period[41]. The two profiles of the chemoattractant presented to the cells have significance in the context of *Dictyostelium* aggregation through chemoattractant waves. As cells begin to aggregate, the local density of cells increases; consequently the local concentration of cAMP also increases[42], as not all cAMP that is produced is degraded [43]. Secondly, every cell experiences several waves before it reaches the aggregation center[33].

In these studies, we took special care to preclude the effects of neighboring cells and self-secreted cAMP. Cells were placed on a substrate with one-dimensional tracks. Crossflow washed away the cAMP secreted by cells during the course of the experiment[44]. cAMP waves were generated using a bespoke microfluidic device by which the concentration of the wave and the background level, and the period of the wave can be controlled. By tracking the cells and computing a measure of chemotactic ability termed chemotactic index (CI)[3], we found that a small background concentration of cAMP has a positive effect on CI. When the background concentration is very high, the chemotactic response is entirely shut down. These trends are also backed up by modeling using a LEGI model coupled to a bistable memory module[4].

In a second study, we studied using a similar setup the effect of exposing the cells to multiple waves of the same period T. We observed that the chemotactic index of cells improves significantly as they encounter more waves when the wave period is small. For longer periods, the improvement is zero or very weak. The improvement is explained by the slow growth in concentration of a directional memory component over several wave periods. This frequency dependent growth in chemotaxis shows the presence of a temporal gradient sensing mechanism operating in parallel with the spatial gradient sensing pathways like LEGI[41]. Our study shows for the first time the presence of temporal gradient sensing in *Dictyostelium* . An earlier study on migrating myeloid cells has shown that those cells too can sense temporal dynamics of chemoattractant concentrations [45]

Both experiments suggest that the chemotactic ability of cells increases once cells are committed to moving to a specific center of cAMP. This has relevance for *Dictyostelium* cells in their natural setting. The initial increase increase in cAMP concentration as well as the occurrence of multiple identical waves in the same direction reinforce the directionality of the chemotaxing cell. Once the aggregating center is established, it is advantageous for cells to reach the center as soon and efficiently as possible.

Chapter 1

Signaling pathways in guard cells - A study using Boolean modeling

Introduction

Many cellular processes in biology are controlled by a large number of components that are part of complex signaling networks [46]. Examples include the pathways controlling cell polarity, cell motility, cell division, and differentiation as well as the gene networks that underlie a myriad of biological processes. The biological function in question frequently arises out of the connections and dependencies among physical and chemical processes which may be relatively simple and well understood. Technological advances in the last few decades have contributed to a proliferation of data at the level of individual genes and metabolites [47, 48], paving the way for synthesizing the knowledge to produce a systems-level understanding.

Models of biological networks attempt to recast the systems in a mathematical form and their level of detail depends on the amount of available data as well as its requirements [49]. In its optimal form, quantitative modeling can replace often laborious experiments by carrying out *in silico* experiments during which one or more components of the pathway or the interactions between components are altered. Even if this is not possible, modeling can often reveal the role of a particular component in the pathway and can, thus, predict the effect of removing or making it constitutively active.

Constructing models for biological pathways requires knowledge about their topology.

In other words, one needs to know whether component A affects component B. This is equivalent to answering the question whether B is downstream or upstream of A. Furthermore, the "sign" of the interaction between these components is required: does A activate (corresponding to a positive interaction) or inhibit (negative interaction) B? Ideally, one would also like to know the strength of the interaction: how much will B increase or decrease when A is present?

For pathways in which all connections and strengths are known, it is possible to construct a mathematical model that represents concentrations of the pathway components as continuous quantities that can take on all positive values. This type of model can provide significant insights, particularly for small systems made up of a handful of simple reactions where all the interactions are known [1, 2]. For these systems, parameters like rate constants, dissociation and association constants can be inferred by monitoring the formation of the product or the decay of the substrate. Often, however, and especially for pathways that contain many components, it is not possible to quantify the type of interaction and the strength between the different components. After all, quantifying this for, say, A and B typically requires a systematic variation of the level A and measuring the response in B. This type of experiment is not always possible for all components and models with a large number of unknown parameters, and interactions can quickly lose their predictive and mechanistic value.

An alternative to creating continuous models is to construct Boolean models [6]. In a Boolean model, each element (alternatively called a node) can only take on one of two values: 0 and 1. The dynamics of these nodes are no longer determined by solving equations that involve rate constants but are updated using logical operations. These operations encode the connections between the different components using the elementary logical functions: identity, AND, OR, and NOT. A Boolean network is then obtained by connecting a number of such nodes in a meaningful manner. Despite these significant simplifications, Boolean networks have been shown to be able to provide insights into genetic networks [7, 8, 9, 10], protein networks [11], and cellular regulatory networks [12, 13]. More importantly, from a practical point of view, Boolean models have several advantages: they can be simulated relatively quickly, even on daily-use desktop

computers, and several software packages are freely available [6, 19]; modifying the network and simulating variants of the original network are easy tasks, and thus Boolean modeling can be used to i) generate hypotheses that can be tested by experiments, and ii) systematically explore variants of a network that 'predict' or lead to an observed phenotype. These ideas have been explored in several studies as will be described later [50, 19].

In this chapter, we focus on Boolean modeling in plant biology. We start with a brief overview of Boolean logic and how one can deduce a Boolean network from rate equations as well as from experimental data. We then discuss software packages that can be used to simulate Boolean networks, after which we discuss applications of Boolean modeling to gene regulatory networks in plants. We then review how Boolean modeling can be used to probe the pathways in guard cells that lead to stomatal closure in response to the plant hormone Abscisic Acid (ABA) and carbon dioxide (CO_2), and end with a brief conclusion and outlook.

Boolean logic and networks

In this section, we will first describe in more detail how Boolean equations are evaluated, provide a simple example, and show how truth tables are a convenient way to analyze and comprehend small Boolean networks. We will then describe how a Boolean network can be constructed from experimental data and describe the various updating schemes developed for this type of network. We will also show how one can translate rate equations into Boolean equations and finish by discussing available software for the simulation of Boolean networks.

Truth tables

The nodes in a Boolean network can only take on values 0 (OFF) and 1 (ON). The ON state of a variable corresponds to high activity or concentration and the OFF state corresponds to low activity or concentration. The interactions between the nodes are given by a combination of the logical functions AND, OR and NOT acting on the input nodes that feed into the output node. The future state of the output node (say at time $t+1$) is obtained by evaluating its

corresponding Boolean function that takes the current states (say at time t) of its input nodes as inputs. To simplify notation, we write the output node and the update rule together as an equation, commonly known as the update equation of the output node. We do not explicitly specify time because i) the update rules do not change with time, and ii) the states of the input nodes specify, through the update equation, the state of the output node in the succeeding time step only.

As a simple example, consider the activation of gene B by a transcription factor A. In this case, when the concentration of A is high, the gene is on while when it is low, B is off. This process can be mathematically expressed using an ordinary differential equation, which describes the rate of change of B, dB/dt, as a function of the concentration of A. In its simplest form, this differential equation is written as

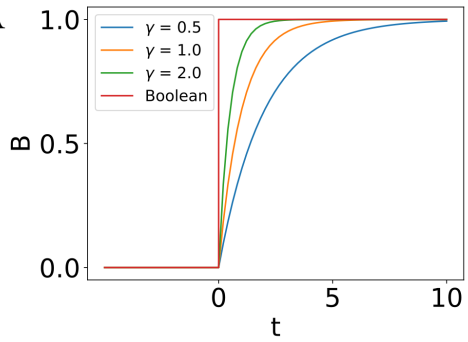
$$\frac{dB}{dt} = f(A) - \gamma B$$

Here γ is a degradation constant, determining how B is removed, and the function $f(A)$ describes how the production rate of the gene depends on the transcription factor concentration A. This function is often taken to be a Hill function $f(A) = \beta A^n/(A^n + K^n)$, with n the (integer) Hill coefficient, β the maximum production rate, and K the activation coefficient. If we take n to be very large, we can approximate $f(A)$ to be a so-called step function: $f(A) = 0$ if $A < K$ and $f(A) = \beta$ if $A \geq K$. Thus, when $A < K$, B will be 0, while for $A > K$, the time dependence of B is found by solving the differential equation

$$\frac{dB}{dt} = \beta - \gamma B \tag{1.1}$$

The steady state value, achieved after a long time, can be found by setting the left-hand side of this equation to zero, resulting in $B = \gamma/\beta$. Furthermore, assuming that A is set above the threshold value K at $t = 0$, the solution of this equation can be found to be $B(t) = \frac{\gamma}{\beta}(1 - e^{-\gamma t})$. This solution is shown in Fig. 1.1A where we plot B as function of time for different values

A



B

A	C*= A	C*= NOT A
1	1	0
0	0	1

C

A	B	C*= A OR B	C*= A AND B
0	0	0	0
0	1	1	0
1	0	1	0
1	1	1	1

D

A	B	C	NOT A	B OR C	X*= (NOT A) AND (B OR C)
0	0	0	1	0	0
0	0	1	1	1	1
0	1	0	1	1	1
0	1	1	1	1	1
1	0	0	0	0	0
1	0	1	0	1	0
1	1	0	0	1	0
1	1	1	0	1	0

Figure 1.1. A. Comparison of the output of a continuous model (Eq. 1) and a Boolean model (Eq. 2) for the activation of a gene. In the former, the output can take on any value between 0 and 1 and depends on the model parameters while in the latter the output is either 0 or 1 and is independent of parameters. B-D: Truth tables of elementary Boolean functions. B. Identity gate, which copies the value of the input to the output ; NOT gate, which copies the inverted value of the input to the output. C. OR and AND gates, which take two inputs. D. An example of a Boolean function that is a combination of the elementary functions. The output X can be determined by evaluating the parts recursively.

of the degradation constant and using $\gamma/\beta = 1$ for simplicity. When the transcription factor is turned on, B approaches its steady state value at a timescale that depends on γ. In this description of gene activation, B can take on all possible values between 0 and 1.

Consider, on the other hand, a simplification of the model in which A and B can only take on values of 0 or 1 and in which the presence of A causes an instantaneous rise in B from 0 to 1. This model can be simply formulated without any parameters by a Boolean equation, which defines how the value of B is updated given the value of A. This equation can be compactly written as

$$B^* = A \tag{1.2}$$

where we have adopted the convention that the variable with an asterisk is being updated. In

other words, if $A = 0$, then B is updated to 0, independent of its current state. If $A = 1$, on the other hand, B is updated to 1, again independent of its current state. The time course of this Boolean equation is shown in red in Figure 1.1A, where A is changed from 0 to 1 at $t = 0$. In contrast to the differential equation, B is immediately turned on when A is set to 1.

The above example is very simple and does not involve any of the elementary Boolean functions. To illustrate these functions, let us now consider the nodes A and B as the input nodes and C as the output node. A useful way to characterize the logical operations is to construct the so-called truth tables, which list the output values for all possible combinations of input values. The truth tables for the elementary logical functions are listed in Fig. 1.1B and C. For example, the AND function ($C^* = A$ AND B) only returns $C = 1$ if both A and B are ON and will return 0 for all other input combinations. Similarly, an OR gate returns an output of 1 if at least one or both of the inputs is 1. Obviously, the identity gate copies the current state of the only input node to the future state of the output node. Similarly, the NOT gate only has a single node as input and it inverts the current state of the node. It can be shown that by compounding these three elementary gates it is possible to encode all Boolean functions [51], including some commonly encountered ones in electronics, such as XOR (exclusive OR). In Fig. 1.1C, we show the truth table of the compound Boolean function X^* = (NOT A) AND (B OR C). This table also illustrates how the output node X is updated by evaluating parts of the function recursively.

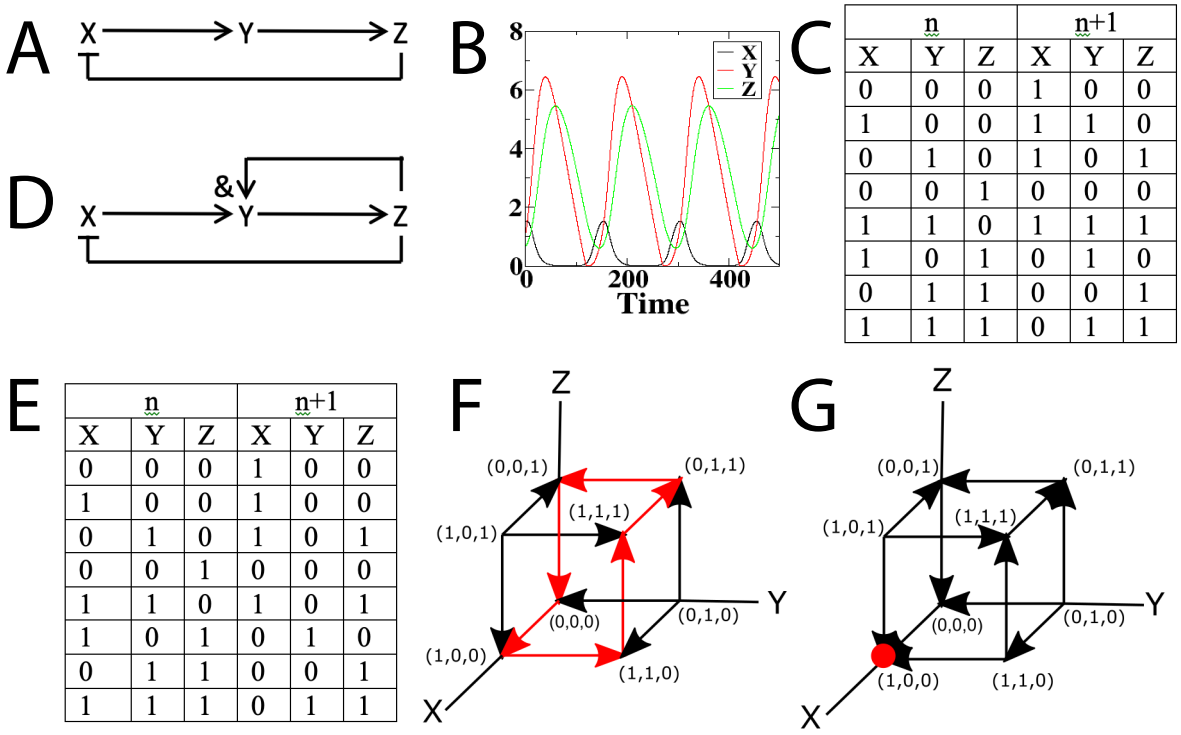
Update rules

Once a Boolean network is constructed, the nodes are updated following a particular update scheme. This is a choice the investigator needs to make because a Boolean model contains neither a natural time scale nor a specified order in which the reactions of the model take place. In the existing literature on Boolean models, three types of update schemes have been used: synchronous, asynchronous, and probabilistic update schemes [6]. In synchronous Boolean models, all the components are updated at the same time, i.e., the states of all the nodes at time step $t + 1$ are determined by their states at time step t [52, 53, 54, 55]. This also means that the

evolution of a synchronous Boolean model is deterministic: a particular input will always result in the same output.

An asynchronous Boolean model orders the updates of the nodes one after another in either a pre-determined or a stochastic manner. There are a number of ways to implement this scheme [56, 57, 58, 53]. For instance, one can follow a **random order asynchronous** update rule wherein all the nodes are updated exactly once but in a random order in each iteration (also called time step). This can be done by generating a random permutation of $\{1, 2, \ldots n\}$, where n is the number of nodes, at the beginning of each iteration. Alternatively, one can follow a **general asynchronous** update rule in which the element that is updated is randomly drawn from the sequence $\{1, 2, \ldots n\}$. Thus, some nodes can get updated, by pure chance, twice or more before another node gets its turn. These two update methods will result in outcomes that are stochastic. This is in contrast to the **deterministic asynchronous** method in which nodes are updated using a fixed sequence [59, 60] or at pre-determined time steps set by the rates the corresponding reaction.

For biological applications, the synchronous update scheme is most likely not appropriate; it is rare that all components in a network change their value at the same time and that all processes take the same duration of time to be completed. Asynchronous updating can in principle implement data on timing and kinetics. However, this type of data is not always available, in which case it is unclear which type of asynchronous updating rule should be used. For a comparison between synchronous and asynchronous update schemes and its consequences, we refer to a study by Faure *et al* [53]. This study applied both schemes to a model for the mammalian cell cycle and also proposed a hybrid scheme, combining both synchronous and asynchronous updating. A third method of updating a Boolean model, also resulting in stochasticity, is through the use of so-called probabilistic Boolean networks [61]. In this updating method, each node in the network has a set of update equations to choose from. At the beginning of a time step, an equation for each node is randomly chosen, after which the nodes are updated synchronously. It thus combines a rule-based determinism for Boolean networks



n			n+1		
X	Y	Z	X	Y	Z
0	0	0	1	0	0
1	0	0	1	1	0
0	1	0	1	0	1
0	0	1	0	0	0
1	1	0	1	1	1
1	0	1	0	1	0
0	1	1	0	0	1
1	1	1	0	1	1

n			n+1		
X	Y	Z	X	Y	Z
0	0	0	1	0	0
1	0	0	1	0	0
0	1	0	1	0	1
0	0	1	0	0	0
1	1	0	1	0	1
1	0	1	0	1	0
0	1	1	0	0	1
1	1	1	0	1	1

Figure 1.2. Examples of Boolean networks. A: Example of an oscillatory network. Arrows indicate activation and flat-edge symbols indicated inhibition. B: The components of the network in A as a function of time, modeled using rate equations (Parameter take from Ref. [64]: k_1=0.1, k_2=0.2, k_3=0.1, k_4=0.05, k_{-1}=0.1, S=2, K_m=0.01, p=4.). C: Truth table for synchronous updating of the network shown in A D: Modified network in which Y depends on X and Z. E: Truth tables for synchronous updating of the network shown in D. F&G: State space and dynamics, represented by arrows, for asynchronous updating of the networks shown in A and C. Fixed point attractors are indicated by red dots while the oscillatory cycle is shown by the red arrows.

with stochasticity arising from the uncertainty from the choice of the update equation. For a review of this type of Boolean model, including its applications, we refer to Trairatphisan *et al* [62, 63].

Translating rate equation models into Boolean models

To see how a signaling network may be encoded using Boolean logic, let us examine one of the simplest three-component systems that can give rise to oscillations [64]. This network is shown in Fig. 1.2A and has only three components X, Y, and Z. The network is wired such that X activates Y, Y activates Z, and Z inhibits X. This is shown in the figures, where activation is

indicated by arrows ($\rightarrow$) and inhibition by a line and a perpendicular bar ($\dashv$). This system can be translated into mathematical equations in which the concentration of the components can take on arbitrary positive values. The resulting set of ordinary differential equations is written as:

$$\begin{aligned} \frac{dX}{dt} &= k_1 S/(1+Z^p) - k_{-1}X \\ \frac{dY}{dt} &= k_2 X - k_3 Y/(K_m+Y) \\ \frac{dZ}{dt} &= k_4(Y-Z) \end{aligned} \tag{1.3}$$

In these equations, k_1,..., k_4 and k_{-1} are the activation and degradation rates, respectively, K_m is a dissociation constant, p is an integer representing the non-linear inhibition of X, and S is an input signal [64]. Simulating these equations for particular sets of parameters results in an oscillatory state as shown in Fig. 1.2B.

To write this network in terms of Boolean operators, it is simplest to examine the diagram of Fig. 1.2A. Note, however, that there are also more systematic ways to derive Boolean networks from ordinary differential equations [65, 66]. This diagram can be translated into the following set of Boolean operators: $Y^* = X$, $Z^* = Y$ and $X^* = \text{NOT } Z$. We can then perform simulations of this Boolean network using synchronous update rules. As mentioned in Sec. 1, for synchronous updating of small networks, it is most convenient to construct the truth table. The table for this diagram is displayed in Fig. 1.2C, which shows that it also exhibits oscillatory cycles. Specifically, starting at $(X,Y,Z) = (0,0,0)$, the sequence is $(0,0,0) \rightarrow (1,0,0) \rightarrow (1,1,0) \rightarrow (1,1,1) \rightarrow (0,1,1) \rightarrow (0,0,1) \rightarrow (0,0,0)$ while $(1,0,1) \rightarrow (0,1,0) \rightarrow (1,0,1)$ is also a cycle.

As a second example, let us consider the previous signaling network but now changed such that the activation of Y depends on both X and Z. This can be easily incorporated by changing the rate equation for Y into:

$$\frac{dY}{dt} = k_2 XZ - k_3 Y/(K_m+Y) \tag{1.4}$$

while keeping the equations for X and Z unchanged. Now, there are two possible solutions: the oscillatory state, similar to the one shown in Fig. 1.2B, and a stationary state given by $Y = Z = 0$ and $X = k_1 S/k_{-1}$. The latter is stable and the resulting state of the system depends on the initial conditions.

The Boolean network corresponding to this slightly altered network is shown in Fig. 1.2D. The only difference between this and the previous network is that the Boolean update equation for Y is now written as $Y^* = X$ AND Z. The truth table for this network, corresponding to the synchronous update scheme, is given in Fig. 1.2E. This table reveals that (1,0,0) is a fixed point of the system: once in this state, the network will remain in it indefinitely. Note however, that this fixed point is only reached for certain initial conditions ((0,0,0), (1,0,0), (0,0,1), (0,1,1), and (1,1,1) to be precise). Thus, as in the continuous version of the network, the binary Boolean network displays a steady state solution in which both Y and Z are zero and in which X has a non-zero value. Obviously, for the continuous system, this value depends on the model parameters while for the parameterless Boolean network it is simply one. As in the continuous system, the Boolean network also exhibits an oscillatory state: $(0,1,0) \rightarrow (1,0,1) \rightarrow (0,1,0)$, which is reached from initial conditions (0,1,0), (1,0,1) and (1,1,0).

Let us now examine these two Boolean networks using asynchronous update rules. In this case, each element can be changed independently. Since our networks contain only three elements, this process can be visualized using the cubes shown in Fig. 1.2F&G, where each node represents a particular state of the system and the edges represent transitions between the states. Here, the arrows indicate the transition between the different nodes according to the rules of the Boolean network. The dynamics of the Boolean network can then be determined by following these arrows.

For the network of 1.2A, it is easy to see that the asynchronous update scheme also results in the same oscillatory cycle as the synchronous update scheme. This cycle is shown in Fig.1.2F by the red arrows. Contrary to the synchronous update scheme, however, the asynchronous update scheme for the second network does not exhibit an oscillatory state. For this update

scheme, regardless of the initial conditions, the network always transitions to the same node, (1,0,0). Thus, the steady state of the system corresponds to a fixed point, indicated by the red dot in Fig. 1.2G. Finally, we should also point out that is possible to go "backwards" and transform a Boolean model into a continuous model [67]. The resulting ordinary differential model could then be used to provide quantitative information regarding, for example, the concentrations of network components.

Encoding a Boolean network from experiments

The task of encoding a Boolean network based on experimental data is not trivial. It requires the identification of the relevant components (nodes in the network) as well as the correct update rules and thus requires biochemical, genetic, and pharmacological data. While identifying components is typically not that difficult, determining the interactions between these components is challenging since the number of possible update equations grows exponentially in the number of nodes in the network [68]. Furthermore, to define these interactions requires careful consideration of experimental data. This task is especially difficult since available experimental information is generally incomplete. To elaborate, consider a node in a Boolean network with n nodes upstream. To formulate the update equation unambiguously, we need the response of the node for the whole set of 2^n inputs. When such information is available, formulating the equation is straightforward [19]. Generally, however, such extensive data are unavailable and simplifying assumptions about the nature of interactions are required.

A classical algorithm to infer a Boolean network, called REVerse Engineering ALgorithm (REVEAL)[69], computes quantities encountered in information theory [70], such as joint entropy and mutual information. The advantage of REVEAL over earlier methods is that one only needs a small fraction of all possible input-output relations to obtain a Boolean network with a very small error rate. The method is, of course, exact when one uses all the 2^n input-output relations for a network of n nodes. To include a more realistic scenario in which one allows for noise in gene regulation, either inherent or caused by measurement techniques, the so-called *Best-Fit*

Extension method [71, 72] can be employed.

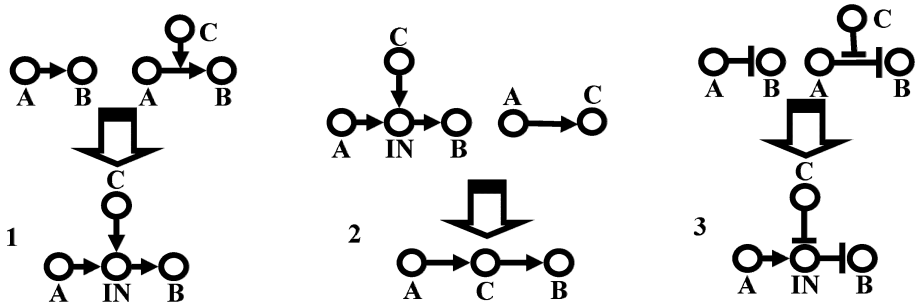


Figure 1.3. Inference rules for the construction of Boolean networks. Experimental data is synthesized to be represented in graphs with the least number of nodes and edges, i.e., as a sparse representation. This sometimes requires an introduction of an intermediary node, as in graphs 1 and 3, but when additional information becomes available, the graph can in fact simplify, as in going from graph 1 to graph 2. For further details, see text (From [73]).

We highlight here another approach used in constructing a large network, following an extensive literature search, to model guard cell dynamics in *Arabidopsis* in response to ABA [73, 18]. Mathematically, this approach relies on developing a graph with the smallest number of nodes and edges consistent with all established qualitative relationships [74]. It formulates a number of inference-based rules, shown schematically in Fig.1.3. In the first graph, experimental data have identified that component A promotes B (and is not a directbiochemical reaction) but also that C promotes the interaction between A and B. In that case, it is assumed that there is an intermediary node (IN) of the $A-B$ pathway and that C acts on this intermediary node. If it is also known that A promotes C, then this intermediary node can be identified as C (graph 2 in Fig 1.3). Finally, if A inhibits B and C inhibits the interaction between A and B, then the logical rule can be interpreted as A promotes an intermediary node IN, which inhibits B, while C inhibits IN (graph 3 in Fig.1.3). Using these rules, it was shown that the developed network was able to capture existing experimental data [73, 18].

Dynamics of Boolean Networks

Often, the goal of modeling is to determine the steady state of the system. That is to say, what is the outcome of the system for long times?Any deterministic Boolean model, when

simulated for long enough time, converges to a limit cycle or an attractor. A limit cycle is a subset of the states of the network over which the state of the system repeats over and over in a cyclical fashion. The length of the limit cycle is the number of states in the limit cycle. An attractor is a state of the system whose 'future' state is identical to the current state; the system gets locked-in once it reaches an attractor state.

We have already seen examples of these two possible outcomes when discussing the networks presented in Fig. 1.2. The set of states that converges to a particular attractor constitutes its so-called basin of attraction. Since the evolution of the network is deterministic, no two basins of attraction share a common element; they are said to be disjoint. Likewise, a limit cycle – along with transient states that feed into it – is disjoint with the next one. Thus, the entire state space can be carved up into disjoint basins of attraction and basins of limit cycles and each trajectory of the system is a subset of the basin its initial state belongs to. Determining the number of attractors, together with their basins of attraction, is an active area of research in the mathematical field of graph theory and we direct the interested reader to several recent studies [75, 76, 77, 78].

Defining and analyzing the basins of attraction for non-deterministic Boolean networks (e.g., using random order asynchronous or general asynchronous update methods) is not as straightforward since the trajectory of the system is no longer deterministic. Furthermore, the set of attractors and limit cycles found using asynchronous updating can be different from the ones found using synchronous updating. This was highlighted by Saadatpour *et al* [56], which carried out a comparative study of the dynamics and steady states of the ABA-induced stomatal closure network under synchronous and the three aforementioned asynchronous update schemes. To enumerate the fixed points and limit cycles, they reduced the system using Markov chains [79] and by simplifying the Boolean update equations. They found that both types of update schemes exhibited a fixed point. However, for synchronous updating, they found large basins of attractions for two limit cycles. These limit cycles, and their basins of attractions, were not found using asynchronous updating, unless strict limitations regarding the timing of several processes

were implemented.

Software tools

Once a network and the update schemes are defined, a Boolean network can be simulated to obtain the trajectory and steady states of the system, to visualize the network, and to determine the activity levels of its components. A large number of computational tools have been developed to simulate Boolean networks on personal computers, as reviewed recently by [6]. In addition, software packages are available to determine and computationally identify the attractors of a Boolean network [80, 81]. Most of these packages, but not all [82], are open source and can thus be freely used. Some of these tools, however, do not use a graphical interface, which makes it more challenging to construct and visualize the network [83, 84, 85, 86, 87, 88, 89]. Other packages only allow synchronous updating and can thus not implement an asynchronous update scheme [90, 91]. Finally, some packages are only able to run a single initialization at a time, which means that probing a large set of initial conditions, especially valuable for large scale networks with asynchronous updating, is challenging [92, 93].

We have developed Boolink, a simulation platform for Boolean networks that is based on a graphical user interface (GUI) and is completely open-source [19]. Specifically, the software allows users to define the nodes and connections in the Boolean network, visualize the network as a tree, set various simulation parameters including the number of time steps and initial conditions, plot the activity of a few chosen nodes, and to analyze the trajectory of the system as a whole. Boolink is written in Python and C++, and the source code is freely available from the GitHub repository github.com/Rappel-lab/Boolink-GUI, along with its documentation, to use, modify, and distribute. We have also packaged the software as a Docker container [94], which is a self-contained system that comes with all the software dependencies and runs straight out of the box. In its original presentation, Boolink was only able to simulate a Boolean network using the physiologically relevant asynchronous update scheme. Later, however, we extended Boolink to include the ability to simulate networks using a synchronous update scheme.

Boolean networks and gene regulation in plants

Creating a regulatory framework based on the available data on gene expression is essential to understanding gene expression. A network that is inferred from gene expression data is termed Gene Regulatory Network (GRN) [95]. Several methods have been developed to construct GRNs from data sets that are extensively available, including Boolean models, information-theory based models, machine-learning based models, etc. These methods and their suitability to different contexts have been reviewed earlier [96, 97, 98, 99]. Here we limit our discussion to Boolean models. The first application of Boolean modeling was carried out by Kauffman when he described a genetic network [100]. In a Boolean gene network, a gene is either turned on (i.e., has value 1) or turned off (with value 0), while the topology of the network specifies how and if a gene interacts with other genes. In plants, Boolean networks have been applied to a number of genetic networks. We will discuss here three different examples: Boolean models for flower development, for induced systemic resistance induced by microbes, and for the root stem cell niche. These models have introduced modifications to the simple implementations of Boolean networks described so far. These modifications will be discussed as the systems are introduced.

Flower development

One of the first examples of Boolean modeling studied early flower development in the model plant *Arabidopsis thaliana* [101]. In this model, 12 genes were considered and the topology of the network was determined based on experimental data. The model was slightly more involved than the simple Boolean implementation we described in Section 1 in that the modified model is known as a threshold Boolean model. Each node of the model still takes binary values (0 or 1) but interactions between any pair of nodes are encoded by *weights* between them; excitatory interactions carry a weight of $+1$ whereas inhibitory interactions carry a weight of -1. The update equation of a node in this model is not Boolean but algebraic, consisting of the sum of weighted interaction terms. When a node is updated, the sum of all of its interactions

with other nodes is calculated. If the sum exceeds the *threshold* of the node, then the node is updated to 1; if not, to 0.

The update scheme for this model is in-between the synchronous and asynchronous update schemes as described before, and is termed semi-synchronic, and block-sequential and block-parallel in later iterations [102, 59]. Instead of updating all the nodes at once or one after another in some order, the nodes in the model are grouped into *blocks*. All the nodes in a block are updated at once, and the blocks themselves are updated sequentially. This method makes use of qualitative experimental data, such as the order of activation of different parts of the genetic network.

The model was found to have 6 attractors, 4 of which were consistent with the gene expression patterns observed in *A. thaliana* [101]. One of the remaining two was not able to flower and the sixth one, while not observed, could be induced experimentally [101]. Since this Boolean threshold model was published, several studies have further analyzed its dynamics. These studies revealed that it is possible to reduce its complexity while maintaining its steady state behavior [103, 104] and highlighted the crucial role of the plant hormone gibberellin in normal flower development [103].

Induced systemic resistance

Recently, the induced systemic resistance (ISR) in *A. thaliana* plants triggered by beneficial microbes was investigated using Boolean modeling [105]. ISR is an important defense mechanism of plants against harmful pathogens [106] and the study investigated how the bacterium *Paraburkholderia phytofirmans* PsJN can trigger ISR and protection from the bacterial pathogen *Pseudomonas syringae* DC3000 [107, 108]. It used the temporal experimental expression patterns of 8 key genes following inoculation of PsJN and asked which threshold Boolean network was able to reproduce the time series data. Parameters of the model, including the weights among the nodes and their threshold values, were fitted to experimental data using an algorithm called Differential Evolution [109], which belongs to a class of fitting algorithms

called Genetic Algorithms [110]. The study inferred 1000 networks from the data. One of these networks was chosen and pruned using biological reasoning. The robustness of the pruned network was then tested by determining how mutations of fundamental genes affected the ISR response. These virtual mutation experiments produced responses that were consistent with available experimental data [105]. Additionally, the study found that the pruned consensus network is robust because it requires an unlikely event of a triple mutation to the network before the ISR is lost. Furthermore, the authors argue that, in the presence of errors in gene expression data, the Differential Evolution algorithm used to derive the gene regulatory network fared better than classical algorithms to infer Boolean networks, including REVEAL [69] and Best Fit Extension [71].

Root stem cell niche

The examples above applied Boolean modeling to determine the most probable network that is consistent with experimental data. In doing so, these studies found missing links or were able to determine the most critical network components. As a result, these Boolean models were often able to predict novel components or connections between components and could suggest new experiments. To further illustrate the ability of Boolean models to guide experiments, we focus here on another example of a gene network studied using Boolean modeling, the root stem cell niche (SCN) in *A. thaliana* [111, 112]. The root SCN in *A. thaliana* is well studied and is located at the root apical meristem [113]. It consists of a so-called quiescent center (QC), comprised of four infrequently dividing cells, and, in immediate proximity, active stem cells that are called initials. The divisions of these initials result in different types of differentiated cells and in tissue growth of the plant [113]. The question thus arises, how can the undifferentiated cells of the QC give rise to several differentiated cell types?

Modeling, and in particular Boolean modeling, is ideal to address this question. Experimental work has identified a number of molecular and genetic components that play a role in the maintenance of the SCN [114]. Furthermore, the interactions between some, but not all

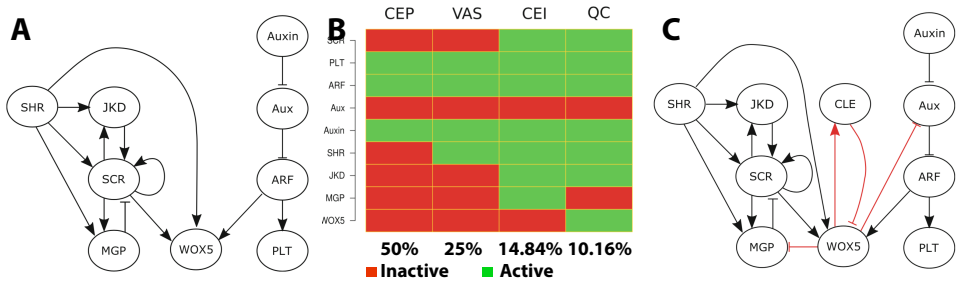


Figure 1.4. Boolean modeling of gene networks. A: Example of a putative network that maintains the SCN in *Arabidopsis*. Arrows indicated activation and flat-edge symbols correspond to repression. For the definition of the different components, see [112]. B: Attractors of the Boolean network shown in A. Green represents an active and red represents an inactive gene. The labels at the top of the diagram represent the attractors and correspond to the phenotypes observed in experiments (CEP: columella epidermis initials, VAS: vascular initials, CEI: cortex-endodermis initials, QC: quiescent center). (From [112]). C: Modified network based on novel experimental and computational results. (From [112]).

of the components can also be deduced from experimental work. It is therefore possible to construct a putative wiring diagram as in Fig. 1.4A, which shows the components along with their interactions as either arrows, indicating activation, or flat-end symbols, indicating repression. The dynamics of this network should then allow steady state solutions with gene expression that is consistent with the different cell types of the SCN. In terms of Boolean modeling, this means that the network should display attractors corresponding to these different cell types.

Simulating the Boolean network in Fig. 1.4A using synchronous updating revealed four different attractors as shown in Fig. 1.4B. In this diagram, active genes are displayed in green while inactive genes are displayed in red. Each of the four attractors correspond to a different set of genes that are on or off and, thus, to a different cell phenotype. For example, the gene SCR (SCARECROW) is on (and thus has a value of 1) in the phenotype corresponding to CEP (columella epidermis initial) and VAS (vascular initial) but is off and has value 0 in CEI (cortex-endodermis initial) and QC (quiescent center cell). The diagram also shows the size of the basin of attraction, expressed as the percentage of initial conditions that resulted in the attractor.

Further testing of this model and comparing the outcomes to experimental results showed

that certain interactions were missing. For example, this analysis revealed the need for a repressor of WOX5 (WUSCHEL RELATED HOMEBOX 5) and an additional component with an inhibitory link was predicted [111]. This prediction was verified in experiments, which showed that WOX5 is negatively regulated by CLE40 (CLAVATA-like-40) [115]. Additional predictions resulted in the modified network displayed in Fig. 1.4C, where the postulated interactions are shown in red [111]. After this study and once new experimental findings became available, this network has been modified and extended further [116]. These studies showed the power of Boolean modeling: once a Boolean network has been constructed, it is fairly straightforward to modify and extend it and to generate experimental predictions. These modifications and extensions are much easier to implement than in continuous models based on rate equations. In those type of models, a modification typically requires refitting and adjusting the model parameters, which can be an arduous task [4].

Boolean networks and signaling in plants

Biological signaling pathways can be very complex, containing numerous components and multiple feedback loops. Such complex pathways can also be addressed by Boolean modeling and examples include T-cell signaling [117], molecular pathways of neurotransmitters [118], and cancer pathways [119, 120]. A prime example of a complex signaling network is found in plants, where the network regulating phytohormone ABA-induced stomatal closure contains a large number of interconnected components. Below, we will review studies that attempt to cast this closure pathway into a Boolean network. Furthermore, we will also discuss recent efforts to extend this signaling network to include CO_2 signaling.

ABA signaling network

Stomata are pores in the epidermis of leaves that regulate gas exchange, including CO_2 for photosynthesis and loss of water vapor. Each stomata is formed by a pair of guard cells and its aperture is modulated in response to environmental changes such as light and CO_2[14, 15].

Furthermore, drought results in the accumulation of ABA in guard cells, which leads to stomatal closure [16, 17]. The network that underlies ABA-induced stomatal closure in *A. thaliana* is complex and contains a large number of components (> 80). Consequently, the number of rate constants is also very large and, not surprisingly, many are not quantified.

We applied Boolink to the ABA-induced stomatal closure network formulated by Albert *et al* [18]. The input file for this network, containing all components by name and their interactions, can be found in the subfolder "sample_data_files/ ABA_data_files/" of the repository for the Boolean equations and for the names of the nodes https://github.com/dyhe-2000/Boolink-GUI or https://github.com/Rappel-lab/Boolink-GUI. The reconstruction of the published ABA signaling network [18] within the Boolink interface here, will enable any user to use and manipulate components of this network and develop experimental predictions, as well as modify the Boolean network depending on experimental outcomes or to predict outcomes for modified network models. A screenshot of our implemented network encoded within the Boolink GUI is presented in Fig. 1.5, with the input ABA node shown in red and the "Closure" output node shown in green. This network contains 81 nodes, including input and output nodes, and was constructed by and adapted from Albert et al. [18] following an extensive survey of more than one hundred peer-reviewed articles. As in the simple example, the interactions between nodes in the GUI are color-coded, with green arrows representing positive interactions and red arrows representing negative interactions. Using the Boolink GUI interface, the user can move nodes around by simply dragging them to a new location. Furthermore, to facilitate examining inter-node connections, double-clicking on a node reveals all downstream and upstream interactions of that node (e.g. Fig. 1.5A). (See https://github.com/dyhe-2000/Boolink-GUI or https://github.com/Rappel-lab/Boolink-GUI; Detailed instructions can be found in the Supplementary Text.)

Results of the Boolink simulations for 25 time steps and averaged over 2,500 initial conditions are presented in Fig. 1.5B. In this, and subsequent curves, and following Albert et al. [18], we chose to illustrate the predicted stomatal conductance level as a function of simulation (time) step. However, this time step does not equate to "real time" and only the steady state

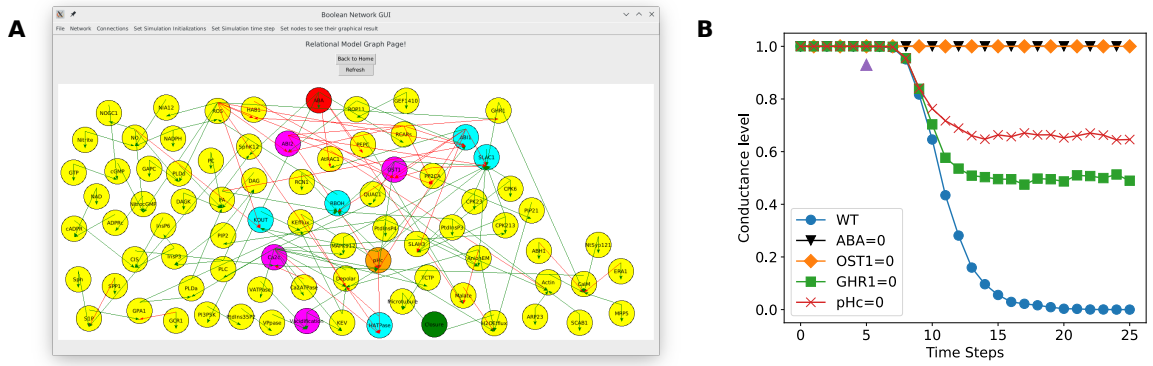


Figure 1.5. Implementation of ABA into Boolink. A: Visualization of the Boolean network for ABA-induced stomatal closure, rendered here by the Boolink GUI. The input ABA and output Stomatal Closure are colored in red and green, respectively. The node denoting cytoplasmic pH (pHc), colored in orange, has its upstream nodes colored in magenta and downstream nodes in cyan. Connections for any node can be viewed by double-clicking on the node of interest. B: Stomatal conductance as a function of time steps in the simulation for wild type and the mutants ost1 and ghr1, and alteration of cytosolic pH (pHc). A conductance level of 1 corresponds to maximal (open) stomatal conductance while 0 represents complete stomatal closure. The triangle shows the point in the simulation where ABA is introduced (except for the case ABA=0).

following a change in network architecture or input can be compared to experimental results. This conductance level, computed as 1-Closure, varies between 0 (corresponding to closed stomata and Closure=1) and 1 (corresponding to open stomata and Closure=0) and facilitates comparison with experiments in which the stomatal conductance is presented (e.g., Fig. ()). In the absence of ABA, simulated by setting the input node ABA to 0 throughout the simulation, the output node Closure is 0 for all time steps, corresponding to no stomatal closure and a conductance level of 1 (orange diamond curve). In the presence of ABA, modeled by changing ABA from 0 to 1 at time step 5, the network reaches a conductance level of 0 (stomatal closure) after approximately 15 time steps (blue curve, labeled as wild type (WT)). By implementing the model of Albert et al. [18], we have also computed the relative stomatal conductance for several mutants, labeled in Fig. 1.5B, following the introduction of ABA after 5 time steps. Knocking out the Open Stomata 1 protein kinase (OST1), which corresponds to forcing the node OST1 to 0 at all times, results in a conductance level that remains 1 (no stomatal closure)

after setting ABA=1 (orange curve) Furthermore, knocking out GHR1 (Guard cell hydrogen peroxide resistant 1) results in a conductance level of 0.5 (50% stomatal closure) following the introduction of ABA (green curve) and manipulating the cytosolic pH through the node pHc leads to a conductance level of approximately 0.65 (red curve). These control predictions are identical to the ones obtained in the original publication [18], further validating reconstruction of this network in our interactive graphical user interface.

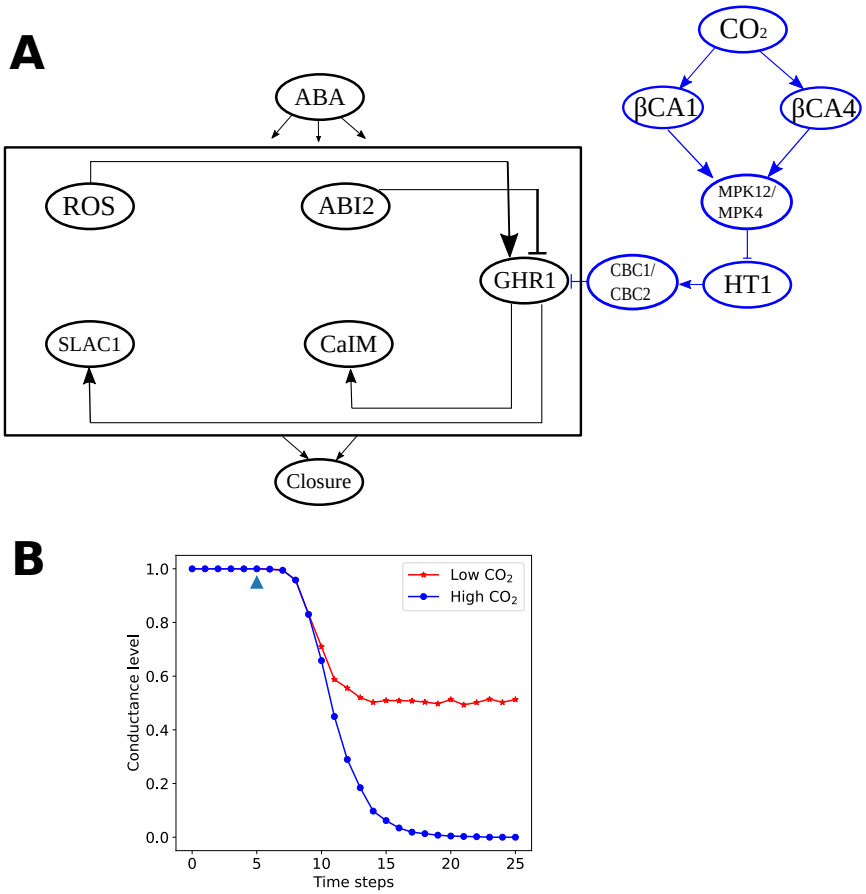


Figure 1.6. Stomata 2.0 and Boolink A: ABA-driven stomatal closure model extended with a CO_2 branch, indicated in blue, which positively regulates GHR1. The box denotes all the intermediate nodes of the original ABA network shown in Fig. 1.5 with only GHR1 and its immediate upstream and downstream nodes shown. B: Predicted relative stomatal conductance levels obtained by implementing Stomata 2.0 into Boolink for two concentration levels of CO_2: low (CO_2 = 0; red line and symbols) and high (CO_2 = 1; blue line and symbols). The triangle shows the point in the simulation where ABA is introduced (ABA = 1).

CO_2 network

We next applied our GUI interface to examine and explore the guard cell signaling network. We used the original ABA network of Albert *et al.* and first extended it with a putative branch that models the input of carbon dioxide (CO_2). Elevated CO_2 triggers stomatal closure and some elements of CO_2 signaling overlap with those of ABA signaling, whereas others affect stomatal closure through separate pathways [20, 21, 22]. Based on previous experimental data, we modeled the CO_2 branch to be upstream of GHR1 in the ABA network [121, 122]. The added branch contains CO_2 as input, which is then catalyzed by the beta carbonic anhydrases βCA4 and βCA1 in parallel [123, 124]. These then activate the node MPK12/MPK4 via yet unknown mechanisms. This node inhibits the negative-regulator of CO_2-induced stomatal closing HT1, which, in turn, regulates CBC1/CBC2 either directly or indirectly [125, 121, 126]. Finally, CBC1/CBC2 enters the ABA network through an assumed inhibitory link to GHR1 (Fig. 1.6 A).

The above summarized CO_2 branch can be translated into the following Boolean equations:

$$\begin{aligned}
\mathrm{CO}_2 &= \mathrm{CO}_2 \\
\beta CA1 &= \mathrm{CO}_2 \\
\beta CA4 &= \mathrm{CO}_2 \\
MPK12/MPK14 &= \beta CA4 \,|\beta CA1 \\
HT1 &= \sim MPK12/MPK14 \\
CBC1/CBC2 &= HT1 \\
GHR1 &= \sim ABI2 \;\&\; ROS \;\&\; \sim CBC1/CBC2
\end{aligned} \tag{1.5}$$

Note that the equation for GHR1 is adapted from Albert *et al* and takes into account the existing connections from the ABA network (from ABI2 and ROS) and the new input from the CO_2 branch. This simplified CO_2 signaling model, which can be accessed in the online Boolink

repository, termed Stomata 2.0, includes presently identified and confirmed early CO_2 signaling mechanisms that have been found to function in the CO_2 signaling pathway upstream of the merging with the ABA-induced stomatal closing pathway [21, 127, 22].

Based on this model, we then determined how the extended network responds in simulations to ABA under high and low CO_2 conditions (Fig. 1.6B), simulated by setting the input node for CO_2 to either 0 (very low concentration) or 1 (high concentration). For CO_2=1, the introduction of ABA at time step 5 results in a decrease of conductance level from 1 to 0 (Fig. 1.6B, blue curve), identical to the WT response shown in Fig. 1.7 (blue curve). This can be understood as the added CO_2 branch having little or no effect on the ABA network since CBC1/CBC2 is 0 when CO_2 = 1.Note that in this model the starting steady-state stomatal conductance of 1 is similar at each background CO_2 concentration, which will be addressed in updated simulations further below. When simulating very low CO_2 conditions, our simulations predicted that introduction of ABA also induces stomatal closure and, thus, a decrease in the conductance level. However, the conductance level in the presence of low (nominally zero) CO_2, was found to be reduced from 1 to 0.5 (Fig. 1.6B, red curve).

To test our predictions experimentally, we analyzed ABA-mediated stomatal closure under either 400 ppm or near 0 ppm (1.5 ppm) [CO_2] by conducting gas-exchange experiments with ABA application to the transpiration stream of excised intact leaves [128]. Our results show that application of 2 μM ABA induced robust stomatal closure in leaves exposed to 400 ppm [CO_2] as expected (blue curves Fig. 1.7A-D). As stomatal responses are known to show biological noise, and as ABA-induced stomatal closing in Arabidopsis thaliana has not been previously analyzed at near zero CO_2, we conducted four independent sets of experiments (Fig 1.7). In all four experiments, leaves exposed to 1.5 ppm [CO_2] showed robust stomatal closing in response to 2 μM ABA, with a degree of expected biological variation (red curves Fig. 1.7A-D). By analyzing the steady state stomatal conductance in leaves it appears that the response to ABA at low CO_2 was reduced in 3 of these experiments (Fig. 1.7A, C&D). We also compared the difference (change) in steady-state stomatal conductance before and after applying ABA (Fig.

1.7 E-H). This analysis shows that independent of whether leaves are exposed to 400 ppm CO_2 or 1.5-2 ppm CO_2, the ABA responses had a similar magnitude in three of the experiments and a stronger ABA response in one of the experiments (Fig. 1.7F). Furthermore, our data show that in the absence of ABA, leaves exposed to low CO_2 show a higher stomatal conductance than leaves exposed to 400 ppm CO_2. This is consistent with a reduction in stomatal conductance upon CO_2 elevation. In addition, analyses of an early time point of the ABA response, 10 minutes after ABA addition, show a slightly, but significantly slowed initial ABA response in 3 of 4 experiments (Fig. 1.7 E,G,H). Taken together, our experimental data show that the steady-state stomatal conductance responses to ABA in Arabidopsis remain to a large degree intact even at very low near zero CO_2 concentration, and as an approximation indicate a steady-state additive effect of low CO_2 on the stomatal conductance prior to ABA exposure.

A comparison of our experimental results and model predictions in the absence of ABA reveals that Stomata 2.0 (Fig. 1.6B) is not able to fully capture the dependence of the steady-state conductance levels on CO_2. This suggests that additional modifications of the ABA network are required to reproduce the observed reduction of stomatal conductance in the presence of CO_2. To explore possible modifications that result in model predictions that are more consistent with our experimental steady-state response results, we utilized the ability of Boolink to easily modify, simulate, and visualize modified Boolean network outcomes. We found that we were able to better reproduce experimental data if we modified the Boolean equations for four network components (Fig. 1.8A). Specifically, we modified the nodes Ca2c (cytosolic calcium), which is linked in the ABA model [18] to CaIM (Ca2+ influx across the plasma membrane), Microtubule (Microtubule depolymerization) and H_2OEfflux (water efflux through the plasma membrane) to:

$$
\begin{aligned}
Ca2c &= \sim \text{Ca2ATPase \& (CIS} \mid \text{CaIM)} \mid \mathbf{(ABA \ \& \ CO_2)} \\
CaIM &= \sim \text{ABH1 \& (NtSyp121} \mid \text{MRP5} \mid \text{GHR1)} \mid \sim \text{ERA1} \mid \text{Actin} \mid \mathbf{CO_2} \\
\text{Microtubule} &= \text{TCTP} \mid \text{Microtubule \& } \mathbf{ABA} \\
\text{H}_2\text{OEfflux} &= \text{AnionEM \& PIP21 \& KEfflux \& } \sim \text{Malate} \mid \mathbf{CO_2}
\end{aligned} \tag{1.6}
$$

where the modifications are shown in boldface. Introducing a CO_2 dependence on calcium signaling was motivated by experimental evidence that cytosolic calcium involved in CO_2-induced stomatal closure [129, 130, 131]. Furthermore, findings that anion efflux and water efflux functions in the CO_2 response and that GHR1 functions in CO_2-induced stomatal closure were included [121, 122]. Addition of an ABA component to microtubule function was motivated based on recent findings [132, 133]. These modifications, their associated Boolean logic operations, and their effect on stomatal closure are further detailed in the Supplementary Text. When simulating this modified and updated network, termed Stomata 2.1, we find that the steady-state conductance level now depends on CO_2 and is reduced for high CO2 conditions (Fig. 1.8B). Furthermore, and also approximately consistent with the experiments, the introduction of ABA reduces stomatal conductance by similar absolute conductance changes for both low and high CO_2 conditions (Fig. 1.8B). The Stomata 2.1 network is able to better reproduce experimental results. We have included the network files for both Stomata 2.0 and 2.1 in the folder sample_data_files/ in the Boolink repositories. We anticipate that community members will be able to use Boolink as a starting point to easily introduce modifications and iteratively test predictions.

Discussion

Large signaling networks are common in biology in general, and plant physiology. The published ABA signaling network we implemented into Boolink, for example, contains more

than 80 components [18]. The vast majority of the interaction strengths and kinetic parameters between these components is not known, making it difficult to formulate mathematical models of these networks. Motivated by the simplicity and utility of Boolean networks and the challenges associated with formulating detailed rate equation-based models for these large networks, we have presented here a software package with a graphical user interface (GUI) that can simulate, visualize, and plot the results of a user-defined Boolean network. Our package, named Boolink, is free to use and distribute, and is built from free and open-source software. The interface is intuitive and users do not require extensive coding knowledge to use it. The Supplementary Text contains detailed instructions on downloading, installing, and running Boolink on Windows, Mac, and Linux-based machines. In addition to this open-source version, we also packaged the software in a *Docker container*, which allows execution of Boolink in an even more facile and direct computer operating system-independent fashion. Instructions detailing how to obtain the Docker container is also provided in the Supplementary Text.

Boolink platform and advantages

As reviewed by Schwab et al. [134], a number of computational platforms exist that are able to simulate Boolean networks. These platforms utilize a variety of programming languages, both high-level and low-level, including Python [91, 84, 89, 87], Java [135, 93, 92, 136, 137, **?**], MATLAB [82], R [88], and C++ [85, 86]. Our platform is distinct from these platforms for several reasons. First, it is the only C++-based platform that offers a GUI interface. Implementation in a low-level programming language like C++ renders simulations faster than implementation in higher level languages, including Java and Python [138, 139]. This can become particularly attractive for large-scale networks that need to be run multiple times, as is the case for our CO_2 signaling network. Coupled to a user-friendly graphics interface, our C++ implementation should thus provide a fast and user-friendly platform for the exploration of Boolean networks. Furthermore, our platform is operating system independent and can be installed on Linux, Windows, and Mac operating systems. The explicit instructions and the Docker container

we provide should facilitate this installation. Finally, our platform implements asynchronous updating, chooses the order randomly, and does not require the explicit specification of the order of node updates.

Boolink uses asynchronous and random order of updates, which is best suited to simulate a network of chemical reactions in which the outcome of one reaction then affects the outcome of another in the near future (hence asynchronous), and when the relative rates of different reactions in the network are unknown (hence random order of update). Besides nodes and connections, users may also specify the number of time steps or iterations to run the simulations and the number of initial conditions to get a statistically robust sample output. Once a system is defined, it may be visualized as a network in Boolink. Visualization includes identifying the upstream and downstream nodes of a given node and the type of connections (activating or inhibiting) between them, by simply double-clicking on the node of interest. The steady states of nodes of the system after simulation can be quickly plotted within Boolink. The trajectory of the entire simulation is stored in a NumPy array; a Jupyter notebook is provided with the package that can be used to further analyze the system starting from the NumPy array, including producing publication-ready plots of the simulation.

Guard cell CO_2 signaling iterative modeling

We first tested and verified Boolink using a published advanced model for stomatal closure in guard cells as mediated by abscisic acid (ABA) and verified that our simulations were consistent with those of Albert et al. [18]. We then extended the network to include the effects of CO_2 on stomatal movements. Previous research indicated that CO_2 might mediate signal transduction via the OST1 protein kinase, as ost1 mutant leaves were impaired in their stomatal response to CO_2 elevation [140, 141]. However, more recent studies unexpectedly showed that, in contrast to abscisic acid, CO_2 elevation does not activate the OST1 protein kinase [21, 22]. This research further provided experimental evidence that basal OST1 protein kinase activity and basal ABA signaling are required for WT-like CO2-induced stomatal closure [21, 22] (Figs.

1.6 and 1.7). Thus, merging CO_2 signaling into the extant ABA signaling model is reasonable. The biochemical link by which CO_2 signaling merges with ABA signaling is proposed to lie downstream of the OST1 protein kinase, but remains unknown. In the present study, to test a simplified model merging the ABA signaling and CO_2 signaling networks, we modeled this link to occur at the level of the transmembrane receptor-like (pseudo)kinase GHR1 [142, 143].

The simulations of this simplified network predicted that the response to ABA should depend on the CO_2 concentration (Fig. 1.6). This prediction was then subsequently analyzed and ABA-mediated stomatal closure of intact Arabidopsis leaves was measured while leaves were either exposed to ambient 400 ppm [CO_2] or near zero (1.5 ppm) [CO_2] (Fig. 1.7). Interestingly, our data show that leaves exposed to 1.5 ppm [CO_2] showed a robust response to ABA. Under low CO_2 conditions the stomatal conductance remained higher prior to ABA application at steady state than at 400 ppm CO_2 (Fig. 1.7A). When comparing true steady state stomatal conductance responses, it appears that CO_2 and ABA may in part have additive responses in Arabidopsis as a first order approximation (note that basal ABA signaling amplifies or accelerates the CO_2 response [21], such that the starting stomatal conductance was much higher at low CO_2 due to the lack of CO_2-induced stomatal conductance reduction (Fig. 1.7A).

In contrast to our experimental data (Fig. 1.7), Stomata 2.0 predicted identical conductance levels for low and high CO_2 concentration in the absence of ABA (Fig. 1.6B). In an illustration of the use of Boolink, we modified the ABA network further, with as goal to incorporate CO_2 dependence on steady-state conductance levels when ABA is absent. Creating this updated network, Stomata 2.1, was greatly facilitated by the ability of Boolink to easily implement changes and generate predictions. We introduced CO_2 dependence on calcium signaling based on experimental evidence that cytosolic calcium is involved in CO_2-induced stomatal closure [129, 130, 131]. Furthermore, it is well-established that anion efflux and water efflux from guard cells are essential for the CO_2–induced reduction in stomatal conductance. Furthermore, findings that GHR1 functions in CO2-induced stomatal closure [121, 122] were expanded to include GHR1 predictions of the original ABA signaling model [18]. Addition of an ABA

component to microtubule function was motivated based on recent findings on roles of guard cells microtubules [132, 133]. The output of Stomata 2.1 was able to better incorporate effects of CO2 and the present experimental data (Fig. 1.6B). As described earlier, important gaps exist in the understanding of the CO_2 signaling pathway, including that the primary CO_2/bicarbonate sensors remain unknown in guard cells and the mechanisms by which HT1, CBC1 and CBC2 link to one-another and to stomatal closing mechanisms are unknown. Expansion of the present model will be required.

In the present study, ABA responses were analyzed at near-zero CO_2 in the C3 model plant *Arabidopsis thaliana.* The robust decrease of steady-state stomatal conductance at near-zero CO_2 by ABA addition was also found in the C3 species *Avena sativa, Gossypium hirsutum*, and *Xanthium strumarium*, but was not found in the C4 species *Amaranthus powelli* and *Zea mays* [144], suggesting a species variability in the response. The converse response was also analyzed in which CO_2 responses were analyzed in the absence of exogenously added ABA in these C3 and C4 species [144]. Interestingly, a variation among species was detected, in which either CO_2 responses proceeded in the C4 species or were impaired in the C3 species [144]. These findings are consistent with recent findings that CO_2 signaling requires basal ABA signaling and would be predicted to depend on variation in basal levels of ABA in guard cells depending on plant species and growth conditions [21, 22]. These classical findings [144, 145] correlate with a model in which ABA and basal ABA signaling plays an important role for other stomatal closing stimuli.

Our proposed additions to the existing ABA network, which illustrate the potential use of Boolink, are meant as a starting point for further explorations and further research is needed to determine the precise mechanism by which CO_2 signaling merges with abscisic acid signal transduction. Nevertheless, several improvements can be suggested. First, it is conceivable that CO_2 affects yet unknown mechanisms. Second, the CO_2 pathway may contain feedback loops, which can be easily implemented within Boolink. Finally, we should point out that Boolean networks do not incorporate explicit rate constants and contain nodes that can only take one of

two values (0 or 1). Therefore, these networks are not able to address the time-dependence of responses nor how they respond to graded inputs.

Importantly, the GUI platform and stomatal signaling model developed here can be used and altered by users to test diverse predictions and to add expanded components or to modify the ABA and CO_2 signaling models. Furthermore, the methods and software tools presented here can be of interest to the wider plant biology community interested in physiological pathways.

Materials and methods

Software

Boolink is implemented using Python and C++ and can be freely downloaded from the GitHub repository https://github.com/Rappel-lab/Boolink-GUI. It requires a current version of Python and C++, and a detailed manual, including installation instructions, is provided in the GitHub repository. These instructions are provided for Windows, Mac, or Linux-based computers. Boolink can also be run as a Docker container, a self-contained environment that includes all the required packages and utilities, on MacOS and Linux-based systems. The advantage of this method is that users only need to install the desktop client for Docker and not the dependencies like C++ and Python. A detailed explanation of instructions to install the Docker client and the required script to run the container is given in the Supplementary Text and in the GitHub repository.

Experiments

Plants of the *Arabidopsis thaliana* accession Columbia (Col-0) were grown as described in [21]. Stomatal conductance (Gs) measurements in response to ABA were performed in detached intact leaves of 5.5- to 7-week-old plants in Arabidopsis leaves following the procedure described previously [128] using a LI-6800 Portable Photosynthesis System with an integrated Multiphase Flash Fluorometer (6800-01A; LI-COR Biosciences, Lincoln, NE, USA). Detached leaves were clamped in the leaf chamber and kept at $\sim$ 1.5 ppm or 400 ppm [CO_2], 135 $\mu Mm^2/s$

red light combine with 15 $\mu Mm^2/s$ blue light, 70 ± 0.5% or 65 ± 0.5% relative air humidity, 21 °C heat exchanger temperature, and 500 $\mu M\,s^{-1}$ incoming air flow rate for at least 2 hour until stomatal conductance is equilibrated and stabilized. Stomatal conductances were recorded every 30 sec under ∼1.5 ppm or 400 ppm [CO_2] for 10 min. ABA (2 μM) was then added to the transpiration stream via the petiole, and stomatal conductances were recorded as shown in the figure panels. In each independent set of experiments, intact leaves from independent plants were analyzed per experimental condition.

1.1 Supplementary Information

Installing and running Boolink

Boolink, the graphical user interface (GUI) developed to simulate user-defined Boolean networks, is free to download, use, and distribute. Detailed instructions to install the software and its dependencies for all operating systems – Windows, MacOS, and linux distributions – are available on the GitHub page of the project at github.com/Rappel-lab/Boolink-GUI. We recommend using the *Docker container* option. Docker provides a self-contained environment for the software wherein all the dependencies and their right versions are already installed. The instructions to install and run the Docker container can also be found on the GitHub project page.

Stomata 2.1

As detailed in the main text, our initial attempt Stomata 2.0 to include CO_2 signaling into the ABA network involved a branch upstream of GHR1. However, the experimental results are inconsistent with the results of Stomata 2.0, which motivated us to seek ways to improve the network. Our experimental results showed that the steady-state conductance level in the absence of ABA is higher for low CO_2 than for high CO_2. In the Boolean simulations, we vary the input nodes ABA and CO_2 between 0 and 1, corresponding to low and high values, respectively. The output is the node Closure, which can be related to stomatal conductance

through Conductance=1-Closure. Thus, we want the model to reflect the following experimental observations:

1. Leaves equilibrated at low CO_2 concentrations have a higher conductance than those equilibrated at high CO_2 concentrations prior to the application of ABA. This should correspond, in the model, to higher conductance for (CO_2=0, ABA=0) than for (CO_2=1, ABA=0). Our experimental data suggest that the stomatal conductance for low CO_2 concentrations is approximately twice as large as for high CO_2 concentrations.

2. The steady state conductance of the leaves after the application of ABA is greater for the lower CO_2 concentrations than for higher CO_2 concentrations. This should correspond, in the model, to higher conductance for (CO_2=0, ABA=1) than for (CO_2=1, ABA=1). Our experimental data suggest that the stomatal conductance for low CO_2 is roughly identical to the stomatal conductance for high CO_2 before the application of ABA.

To translate these experimental findings into a quantitative Boolean state of the output node, we require the model to reproduce the following observations:

ABA	**CO_2**	**Closure / Conductance**
0	0	0
0	1	0.5
1	0	0.5
1	1	1

The Closure node in the ABA network model is affected by two nodes, Microtubule and H2OEfflux, through the equation Closure = H_2OEfflux AND Microtubule. To achieve an intermediate level of closure, required for the conditions ABA=0, CO_2=1 and ABA=1, CO_2=0, we need one of the two nodes at 100% activity and the other at 50% activity (fluctuating between 0 and 1). For full closure, we need both nodes at 100% while for full conductance we need both nodes at 0. In Stomata 2.1, this is achieved through the following modifications:

$$
\begin{aligned}
Ca2c &= \sim \text{Ca2ATPase} \;\&\; (\text{CIS} \mid \text{CaIM}) \mid \mathbf{(ABA \;\&\; CO_2)} \\
CaIM &= \sim \text{ABH1} \;\&\; (\text{NtSyp121} \mid \text{MRP5} \mid \text{GHR1}) \mid \sim \text{ERA1} \mid \text{Actin} \mid \mathbf{CO_2} \\
\text{Microtubule} &= \text{TCTP} \mid \text{Microtubule} \;\&\; \mathbf{ABA} \\
\text{H}_2\text{OEfflux} &= \text{AnionEM} \;\&\; \text{PIP21} \;\&\; \text{KEfflux} \;\&\; \sim \text{Malate} \mid \mathbf{CO_2}
\end{aligned} \tag{1.7}
$$

where the modifications are shown in boldface. As a reminder, the symbol & represents AND logic and the symbol | represents OR logic.

The motivation for these modifications can be described as follows

1. In the original Albert version of the model, Ca_2c (cytosolic calcium) has 50% activity due to oscillations between Ca_2c and $Ca_2ATPase$, if and only if ABA=1. Since the original network has no CO_2 input, this is independent of the state of CO_2. We added input from CO_2 so that Ca_2c has 100% activity if both ABA=1 and CO_2=1. Adding CO_2 to calcium was motivated by evidence that cytosolic calcium is involved in CO_2-induced closure.

2. CO_2 is added to CaIM through an OR gate to ensure that Microtubule=0.5 when CO_2=1, even in the absence of ABA.

3. In the original ABA network model, Microtubule was always either 0 (if ABA=0) or 1 (if ABA=1), even though Ca_2c=0.5. This is because of the feedback loop from Microtubule onto itself. To achieve Microtubule=0.5, we make the feedback loop dependent on ABA. As a result, Microtubule=0.5 if ABA=0 and CO_2=1 but Microtubule=1 when both ABA and CO_2 are 1.

4. In the original ABA network model, H2OEfflux is 0 if ABA is absent, independent of the state of CO_2. With this modification, H_2OEfflux=1 when ABA=0 and CO_2=1.

As a result of these modifications, the network is able to reproduce the experimental

results (See Fig. 1.8). It is straightforward to simulate and visualize the network using Boolink, which can be used to analyze the response:

When ABA=1 and CO_2=1, H_2OEfflux = 1 as all the terms are 1. Furthermore, Ca_2c oscillations are superseded by a sustained activation because of the ABA&CO_2 term in modification #1. Hence Microtubule is maintained at 1 as well and Closure=1.

When ABA=1 and CO_2=0, H2OEfflux is at 50% since AnionEM is at 50%, and PIP21, KEfflux, and Malate are at 100% activation. Ca2c is at 50% (due to oscillations) in the absence of CO_2 but the positive feedback of Microtubule is activated in the presence of ABA, making it 100% active. Thus, Microtubule=1, H_2OEfflux=0.5, and Closure=0.5.

When ABA=0 and CO_2=1, H_2OEfflux=1 due to the CO_2 branch implemented in Stomata 2.0 and also present in Stomata 2.1. Similarly, the addition of the CO_2 branch to CaIM causes Ca_2c oscillations even in the absence of ABA. Thus, Ca_2c is at 50% activity and Microtubule=0.5, as the positive feedback is shut down in the absence of ABA. Thus, Microtubule=0.5, H_2OEfflux=1, and Closure=0.5.

When ABA=0 and CO_2=0, both H_2OEfflux and Microtubule are shut down, and there will be no activity of closure. In other words, Microtubule=0, H_2OEfflux=0, and Closure=0.

Chapter 1, in part, is a reprint of the material as it appears in Plant Physiology 2021 and Quantitative Plant Biology 2022. Karanam, A., He, D., Hsu, P.K., Schulze, S., Dubeaux, G., Karmakar, R., Schroeder, J.I. and Rappel, W.J., 2021, Plant Physiology 2021, and Karanam, A. and Rappel, W.J., 2022, Quantitative Plant Biology 2022. The dissertation author was the primary investigator and author of both papers.

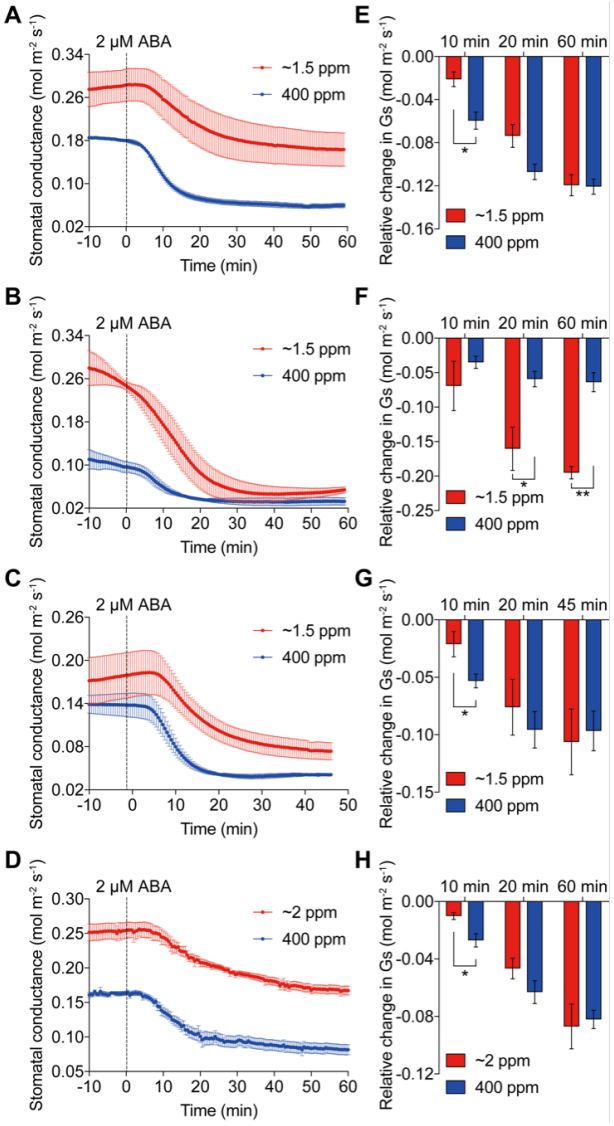


Figure 1.7. ABA-mediated stomatal closing responses of WT leaves during CO_2 "starvation". A–C, Intact excised leaves of wild-type plants (n = 3 to 4 independent leaves per treatment and experimental set) were equilibrated at 400 ppm CO_2 or $\sim$ 1.5 ppm CO_2 for 60 min prior to stomatal conductance measurements. Stomatal conductances were measured with the LI-6400XT Portable Photosynthesis System. D, Time-resolved stomatal conductance responses to ABA in the intact excised leaves (n = 5 independent leaves per treatment) equilibrated at $\sim$ 2 ppm CO_2 or 400 ppm CO_2 for 2 hr prior to ABA application. Experiments were carried out using the LI-6800 Portable Photosynthesis System. In each experimental set, 2 µM ABA was applied through the transpiration stream via the petiole at time = 0 min. CO_2 concentrations in the intercellular spaces of leaves (Ci) equilibrated under CO_2 starvation were computed using the gas exchange analyzer (see Methods), with Ci values < 20 ppm (A to C) and < 3 ppm (D) before application of ABA. E to H, Changes (differences) in stomatal conductance at the indicated time points after ABA application compared to 0 min. Data present mean $\pm$ SEM. $*P < 0.05$ and $**P < 0.01$ student's t-test in E to H.

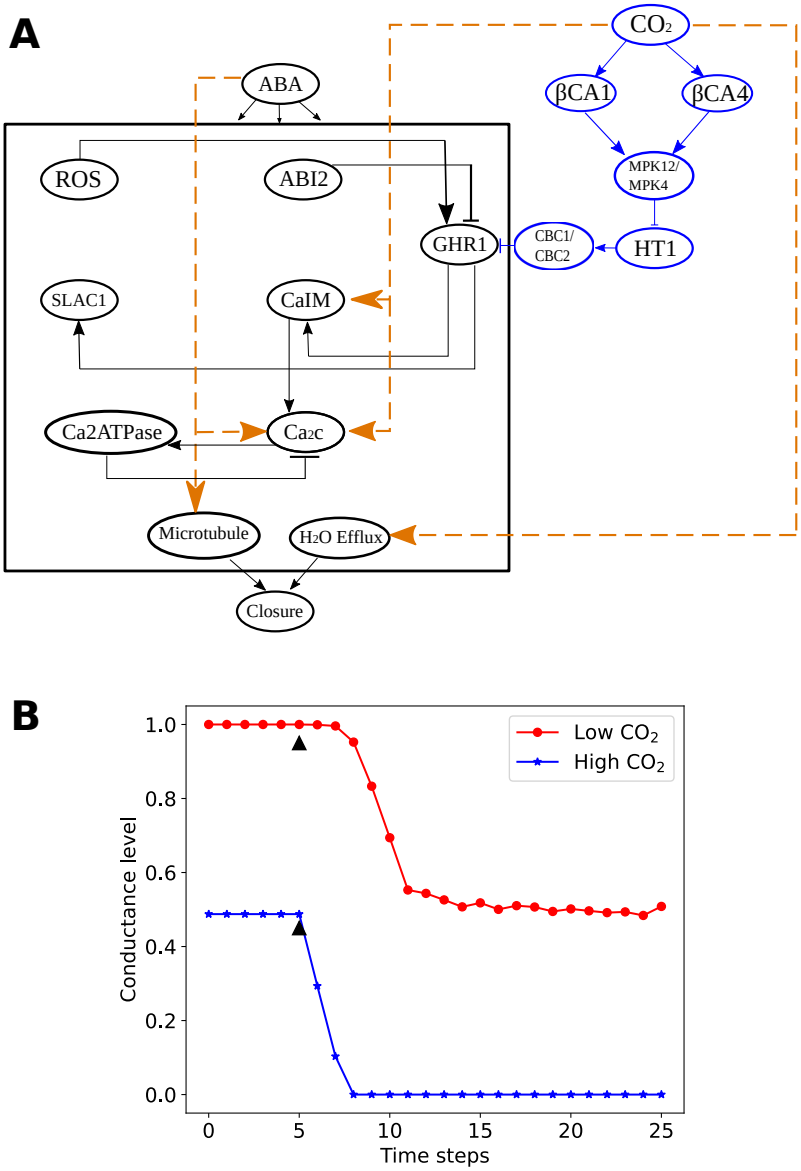


Figure 1.8. Stomata 2.1 and Boolink A: ABA-driven stomatal closure model extended with a CO_2 branch, indicated in blue, which positively regulates GHR1, and additional modifications represented by the orange links (see Text for details). B-C: Predicted relative stomatal conductance levels obtained by implementing Stomata 2.1 into Boolink for two concentration levels of CO_2: low ($CO_2 = 0$; red line and symbols) and high ($CO_2 = 1$; blue line and symbols). The triangle shows the point in the simulation where ABA is introduced.

Chapter 2

Cellular memory in eukaryotic chemotaxis depends on the background chemoattractant concentration

Introduction

Chemotaxis, the movement of cells guided by chemical gradients, plays an important role in many biological processes including tumor dissemination, wound healing, and embryogenesis [26, 146, 147, 148]. One of the most studied chemotaxis model organisms is the social amoeba *Dictyostelium discoideum*. Following starvation, *Dictyostelium* cells secrete a chemoattractant, cAMP, in a periodic fashion [34]. This chemoattractant signal is relayed by neighboring cells resulting in waves that sweep over the cell population with periods that range from 6-10 minutes [35, 36, 34]. These waves spontaneously organize themselves in spiral or target waves, leading to large-scale patterns of cell migration and eventually generating aggregation centers that attract tens of thousands of cells. Within the resulting aggregates, cells differentiate, with the majority turning into spore cells.

Multiple aspects of this biological system have been investigated using computational and mathematical modeling [149]. Models have addressed instabilities responsible for large scale migration patterns [150, 151], the coupling between intracellular signaling and morphological changes [152, 153, 154], and the topology of signaling pathways responsible for guided motion [155, 156, 157]. Furthermore, models have addressed potential mechanisms of gradient sensing

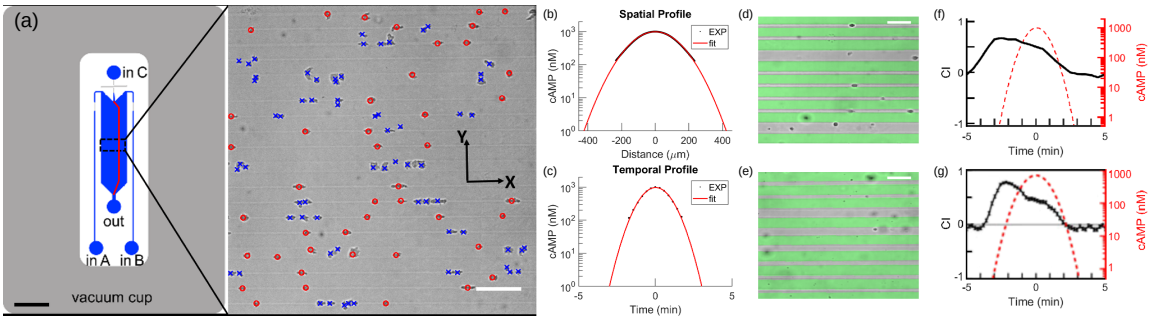


Figure 2.1. (a): Left: Schematic of the microfluidic wave device, with observation region indicated by black box (Scale bar: 3 mm). Right: Snapshot of cells moving on the micropatterned substrate, with symbols f corresponding to cells identified by the machine learning algorithm (red circles: cells used in our analysis; blue X's: excluded cells that are too close to one another; the blurry spots are out-of-focus dirt particles and other irregularities that are not identified as cells by the machine learning algorithm; scale bar: 100 μm). (b-c): Spatial (b) and temporal profile (c) of the cAMP wave, determined from the fluorescent intensity of the dye, and the result of the Gaussian fit. (d-e): Images of the two substrate patterns used in this study, with green highlighting the location of the PEG-gel stripes. The pattern consists of either 4 narrow ($\sim 10\mu$m) and 1 wide ($\sim 25\mu$m) untreated stripes (d) or of 6 variable width stripes, ranging from $\sim 6\mu$m to $\sim 25\mu$m (e). In both patterns, the untreated stripes are separated by 30 μm wide non-adhesive PEG-gel stripes. Scale bar: 50 μm. (f-g): The CI as a function of time for the current experiment using a micropatterned substrate (f) and in a previous study [3], using a non-patterned substrate (g). The results are qualitatively similar, indicating that restricting the cells to 1D stripes does not affect their chemotactic behavior.

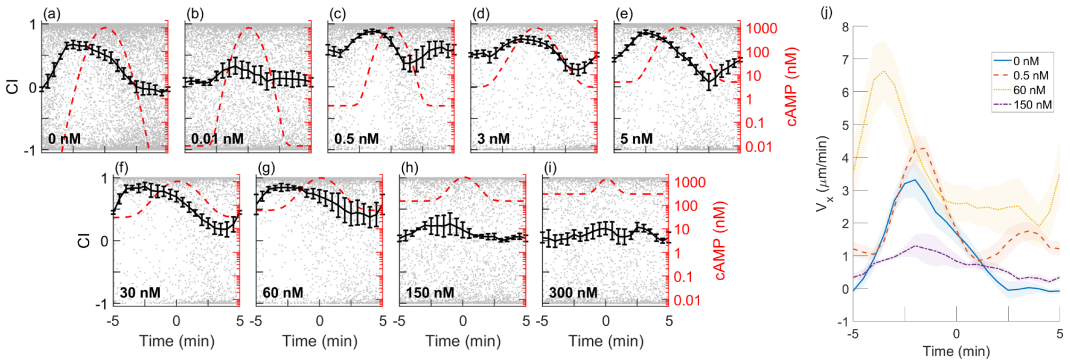


Figure 2.2. (a-i) Experimentally determined average CI as a function of time (measured relative to peak of wave) for different concentrations of background cAMP (0-150nM). In each panel, gray dots represent the CI of individual cells, the black curve is the binned average over N=3-4 different experiments, and the dashed red line is the cAMP concentration of the wave. (j) Average x-component of the velocity of cells for different concentrations of background cAMP. Time is binned in intervals of 0.5 min. Error bars in this figure represent the standard error of the mean obtained using bootstrapping.

[158, 149] while a number of studies have examined the role of noise in the chemotactic response[159, 160, 161].

Several experimental and modeling studies have also addressed the so-called back-of-the-wave problem in the chemotactic response to traveling waves [33, 162, 3]. If cells respond only to spatial gradients, they would move forward in the front and backward in the back of the traveling wave, preventing aggregation. These studies have shown that cells exhibit memory, responding directionally to the front but not the back of the wave, enabling them to move efficiently toward the wave source [162, 3]. For wave periods shorter than 10 min, this memory completely prevented reversals of cell migration, whereas for longer periods, cells started reversing their migration direction in the back of the wave [3]. A mathematical model, consisting of an upstream adaptive module and a downstream bistable module, was able to explain the response of cells to periodic waves of chemoattractant [3]. Similar memory phenomena have also been reported in other biological systems, including chemotactic neutrophils [38, 39].

In the experimental studies of memory in *Dictyostelium* chemotaxis, the cAMP waves were applied exogenously, with the cAMP concentration reaching nearly zero in the troughs of the

waves. In cAMP waves that are endogenously produced by starving populations of *Dictyostelium* , however, the background cAMP concentration, $[\text{cAMP}]_{\text{bg}}$, increases from cycle to cycle [42]. This increase occurs because secreted cAMP is not completely removed by phosphodiesterases (PDEs), enzymes that are responsible for the degradation of cAMP and that are also secreted by the cells [43]. Hence, the question how a non-zero $[\text{cAMP}]_{\text{bg}}$ affects chemotaxis is relevant to the aggregation, sporulation, and survival of *Dictyostelium* . Naively, one would expect a decrease in the cell's ability to migrate towards the source of the wave since the fractional gradient across the cell body, and thus the signal-to-noise ratio, decreases for increasing $[\text{cAMP}]_{\text{bg}}$[163]. Another possibility is that the cells fully adapt to $[\text{cAMP}]_{\text{bg}}$, rendering the ability of cells to respond to gradients independent of the background concentration [164].

To experimentally study this question, we used a modified version of the microfluidics device from Ref. [3] in which a traveling, bell-shaped wave of cAMP with a peak of 1000 nM periodically sweeps across a gradient channel at a constant speed (Fig. 2.1a; see Appendix A for further details). The fluorescent intensity profile with $[\text{cAMP}]_{\text{bg}}$=0 was well fitted using a Gaussian (Fig. 2.1b-c and Section 2.1) and the resulting wave profile is similar to the one measured for natural waves of cAMP [165, 166]. Importantly, and in contrast to previous studies, the background concentration of cAMP was a variable parameter. Another major modification was that the glass substrate in the gradient channel was micropatterned with $\sim$1.5 μm thick stripes of cell adhesion-blocking polyethylene glycol (PEG) gel. Examples of the two types of micropatterned substrates used in this study are shown in Fig. 2.1d-e. These substrates limit the adhesion and migration of *Dictyostelium* cells to $\sim$6-$\sim$25 μm wide stripes of non-PEG treated glass oriented in the x-direction, along the gradient and perpendicular to the flow. A detailed description of these micropatterned substrates can be found in Ref. [44], where it was shown that cells are nearly exclusively constrained to glass stripes. As a result, cell migration was effectively one-dimensional (1D), either up or down the gradient (positive or negative x-direction), greatly facilitating the collection and analysis of data as compared to 2D chemotaxis on a standard glass substrate.

In our experiments, we exposed cells to repeated waves of cAMP and recorded their movement, excluding the first wave. We used cells of the axenic *Dictyostelium discoideum* strain AX4 that were transformed to express a fusion of GFP to LimE (Δ coil LimE-GFP) and a gene encoding a fusion of RFP to Coronin (LimE GFP/corA RFP) [167]. The cells were grown in submerged shaking culture in HL5 medium (35.5g HL5 powder (Formedium, Norfolk, UK) and 10mL Penicillin-Streptomycin (10,000 U/mL; Gibco, Thermo Fisher Scientific, USA) per liter of DI water) [168]. For starvation, when cells reached their exponential growth phase (3-4 $\times$ 10^6 cells/mL), they were harvested by centrifugation at 3000 rpm for 5 min, resuspended in KN_2/Ca buffer (14.6 mM KH_2PO_4, 5.4 mM Na_2HPO_4, 100 μM $CaCl_2$, pH 6.4), collected by centrifugation, and re-suspended in KN_2/Ca at 10^7 cells/mL. Cells were developed for 5 h with pulses of 50 nM cAMP added every 6 min. Differential interference contrast (DIC) images were taken every 15 s in four fields of view spanning the width of the chemotaxis channel, 2800 μm away from the cAMP inlet, on a spinning-disk confocal Zeiss Axio Observer inverted microscope using a 10X objective and a Roper Cascade QuantEM 512SC camera. Images were captured and analyzed by using Slidebook 6 (Intelligent Imaging Innovations).

Cells were tracked with a custom-made machine-learning algorithm detailed in Section 2.1. Only cells that were at least 33 μm from neighboring cells were used in the analysis (marked by red circles in Fig. 2.1a) while cells that were part of clusters were not taken into account (marked by blue crosses in Fig. 2.1a). This algorithm was able to capture more than 90% of single cells, as determined by manual counting. Using cell tracks, we quantified the directional response by computing the chemotactic index CI, defined as ratio between the velocity in the x direction and the speed, computed as the difference in the x-position 3 frames prior and 3 frames forward (a 90 s interval): CI=V_x/V. Thus, this quantity ranges from +1 (cells with velocity perfectly aligned to the $+x$ direction), to -1 (cells with velocity perfectly aligned to the $-x$ direction).

Experimental Results

We first examined how cells responded to chemoattractant waves with a period of 10 minutes and very low background concentration; $[\text{cAMP}]_{\text{bg}}$=0 and 0.01 nM). This wave period was chosen since it corresponds to the largest period for which the CI in 2D assays remained positive in the back of the wave [3]. The CI for these background concentrations, shown as a black line in Figs. 2.2a and b, computed by averaging over different experiments and over 30s time-intervals, was nearly zero for cells ahead of the wave front, steeply increased to a maximum of $\sim$0.7 as cells were exposed to the wave front, stayed high after the peak of the wave has passed, and gradually decayed to near zero but never became negative. The response for zero background is qualitatively similar to that in 2D *Dictyostelium* chemotaxis assays on plain substrates [3], indicating that constraining the cells to narrow stripes does not change their behavior and that the cellular memory reported in 2D assays is fully manifested in 1D assays as well (Fig. 2.1f-g).

Next, we exposed cells to the same periodic waves but with larger background concentrations ($[\text{cAMP}]_{\text{bg}}$=0.5, 3, 5, 30, and 60 nM) (Fig. 2.2c-g). We found that the average CI improved for these values of $[\text{cAMP}]_{\text{bg}}$: the CI remained much greater than 0 during the entire wave cycle. Furthermore, the CI showed a clear minimum after the peak of the wave has passed and *increased* towards the end of the wave cycle even though $[\text{cAMP}]_{\text{bg}}$ was at its lowest level. For the two largest values of $[\text{cAMP}]_{\text{bg}}$ tested, $[\text{cAMP}]_{\text{bg}}$=150 nM and 300 nM, the CI was reduced and remained close to zero throughout the entire wave cycle (Fig. 2.2h-i). Both the increase in average CI for small values of the background concentration (0.01 to 0.5nM transition) and the decrease in average CI for large values of the background concentration (60 to 150nM transition) were significant ($p < 10^{-4}$, using a z-test). Thus, the background cAMP concentration has a profound effect on the chemotactic response, with intermediate/large values of $[\text{cAMP}]_{\text{bg}}$ enhancing/suppressing the response.

The effect of the background concentration was also evident from the quantification of

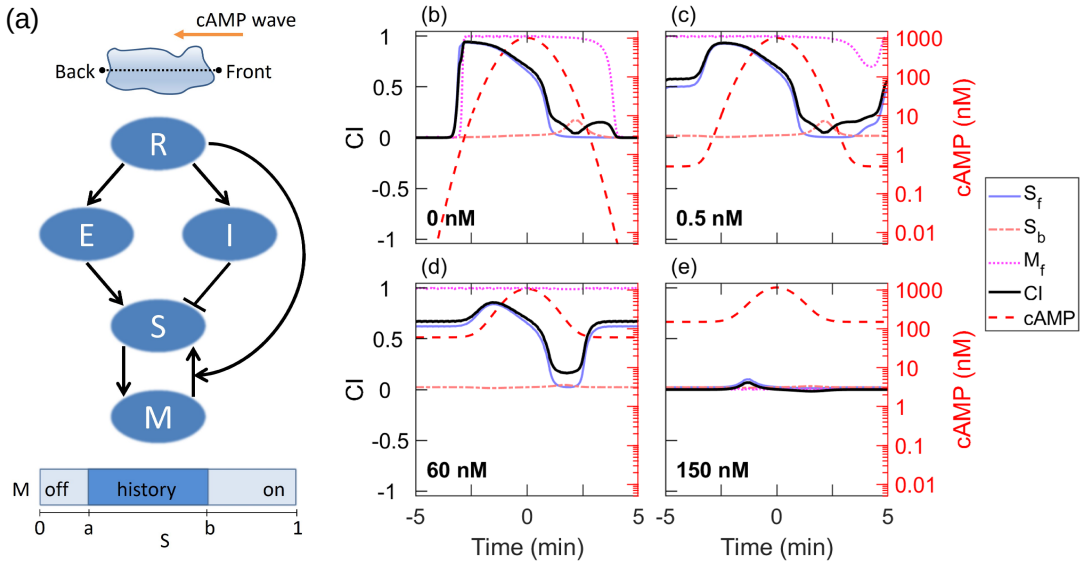


Figure 2.3. (a) Schematic diagram of the chemotactic model, consisting of a receptor R, an activator E, an inhibitor I, a response element S, and a memory component M. Simulations are carried out in a 1D geometry (top drawing). As indicated by the bottom bar, M is bistable, with a low and a high state, determined by parameters a and b. (b-e) Model results for different background cAMP concentrations added to a periodic wave, shown as a dashed red line. The black line represents the CI, the blue (light gray)/red (dash-dotted) line is the response S at the front/back of the cell, and the dotted magenta line corresponds to the memory M at the front.

the x-component of the velocity, V_x. This quantification is shown in Fig. **??**e where we plot V_x, also averaged over different experiments and over 30s time-intervals, as a function of time for different $[\text{cAMP}]_{\text{bg}}$. While this velocity component remained positive or close to zero during the entire cycle for all values of $[\text{cAMP}]_{\text{bg}}$, its maximum value is clearly larger for intermediate values of $[\text{cAMP}]_{\text{bg}}$ than for $[\text{cAMP}]_{\text{bg}}$=0nM. Furthermore, V_x is significantly reduced for the large background concentrations ($[\text{cAMP}]_{\text{bg}}$=150 nM).

Model Results

To investigate plausible mechanisms for this enhanced cellular memory, we turned to modeling. Specifically, we asked whether the cellular memory model developed by Skoge *et al.* [3] can reproduce the experimental results. This model describes the chemotaxis pathway in terms of abstract variables, although, for some, identification with biochemical components may be possible. Key features of this model are perfect adaptation upon uniform stimulation and cellular memory in gradients [40, 3]. It is schematically shown in Fig. 2.3a and contains an adaptive module, which incorporates an incoherent feedforward Local Excitation Global Inhibition (LEGI) mechanism [156, 155] and consists of a receptor R, an activator E, an inhibitor I and a response element S (Fig. 2.3a). In addition, the model contains a memory module, which is assumed to be bistable such that its component M can be either in a low or high state. The transition between these state is determined by two thresholds, a and b and M feeds back to S. Importantly, this feedback depends on R and this non-adaptive link may be thought of as representing parallel pathways for chemotaxis described in experimental studies [169]. For simplicity, we neglect the detailed morphology of the cell and model it as a 10 μm line with the two endpoints representing the front and back, respectively (Fig. 2.3a). At the front, the model is

written as

$$\frac{dR_f}{dt} = k_R(\text{cAMP}+\text{cAMP}_{\text{bg}})(R_f^{tot}-R_f)-k_{-R}R_f \tag{2.1}$$

$$\frac{dE_f}{dt} = k_E R_f - k_{-E}E_f \tag{2.2}$$

$$\frac{dM_f}{dt} = -k_{Mem}M_f(M_f-M_f^{tot})\left(M_f-M_f^{tot}\frac{b-s_f}{b-a}\right) \tag{2.3}$$

$$\frac{dS_f}{dt} = k_S E_f \frac{S_f^{tot}-S_f}{K_{m1}+S_f^{tot}-S_f} - k_{-S}I\frac{S_f}{K_{m2}+S_f} + k_{S2}M_f R_f \frac{S_f^{tot}-S_f}{K_{m3}+S_f^{tot}-S_f} \tag{2.4}$$

and a similar set of equations applies for the components at the back, labeled with subscript b. The first equation describes the binding/unbinding dynamics of cAMP to the receptor with on and off rates k_R and k_{-R}, respectively. Here, cAMP is the time-varying concentration due to the wave and its dynamics is taken from a Gaussian fit to the wave profile (Fig. 2.1b-c). The second equation models the activator dynamics, parametrized by the activation rate k_E and the degradation rate k_{-E}. The third equation describes the bistable memory module, which has two stable steady states 0 and M_f^{tot}. The output of the model, S, obeys Michaelis-Menten kinetics and is activated with rate k_S by the activator E. It is de-activated with rate k_{-S} by an inhibitor I, which diffuses within the interior of the cell. We will assume that I is uniform throughout the cell, corresponding to a high diffusion rate, and is activated by the average of R between the front and the back $R_{av} = (R_f + R_b)/2$, along with a small basal activity k_{basal}, and can spontaneously degrade with rate k_{-I}:

$$\frac{dI}{dt} = k_I(k_{basal}+R_{av})-k_{-I}I \tag{2.5}$$

We set the Michaelis constants, K_{m1} and K_{m2}, to be small ($K_{m1}, K_{m2} \ll 1$) to achieve near zero-order ultra-sensitivity, which, consistent with experiments, can result in the amplification of shallow external cAMP gradient[170, 171]. The last term in the equation for S_f describes the feedback from the memory module to the activation of S_f, parametrized by the activation rate k_{S2} and Michaelis constant K_{m3}. For simplicity, we will take $M_f^{tot} = M_b^{tot} = S_f^{tot} = S_b^{tot} = R_f^{tot} =$

$R_b^{tot} = 1$. The equations were simulated using an adaptive Runge-Kutta method with variable step size. To prevent the memory from becoming trapped in a fixed point, we reset M_f and M_b every t=15s to δ or $1-\delta$ if their values are smaller than δ or larger $1-\delta$ (δ=0.01).

Model parameters were determined using a fitting procedure (detailed in Section 2.1), which minimized a loss function L that compared simulation results (x_{sim}) to experimental results (x_{exp}) with uncertainty σ_{exp}: $L = \sum_{n=1}^{N} |x_{exp} - x_{sim}|/(N\sigma_{exp})$. Here, $N = 56$ with 46 data points chosen from previous experiments using different microfluidic devices [40, 3] and the remaining 10 data points chosen from the current experiments. Details of the previous data points used in the fitting are presented in Section 2.1 while the new data points consisted of data for $[\text{cAMP}]_{\text{bg}}$ of 0, 0.01, 0.5, 60, and 150nM that were chosen since they represent the three qualitatively distinct responses observed in the experiments. Specifically, in our fitting we required that M_f 2.5 minutes before and 5 minutes after the peak to be either 0, corresponding to a small experimental value of CI, or 1, corresponding to a high CI in the experiments. Simulated annealing was used to find possible global minima, followed by a pattern search to obtain the local minima using the Matlab routine patternsearch. Importantly, the parameter values for the bistable and adaptive module were taken from previous studies [40, 3] and only parameters associated with the memory module (a, b, k_{Mem}, k_{S2} and K_{m3}) were adjusted. The parameter values obtained by our fitting procedure are listed in Table 2.1 while a fit to previous data is shown in Fig. 2.4a-b.

Simulation results for $[\text{cAMP}]_{\text{bg}}$=0 (Fig. 2.3b) show that the chemotactic response, quantified by CI (computed, following our earlier study [3], as a linear combination of S and M: CI$= 0.1831(M_f - M_b) + 0.8169(S_f - S_b)$ is in qualitative agreement with the experimentally measured CI (cf. Fig. 2.2a). In this case, as the wave approaches the cell, the small difference in cAMP between the front and back is greatly amplified because of the ultra-sensitivity of the response. The resulting large increase in S_f (blue (light gray) line) causes a transition of M_f (dotted magenta line) to the high state. Since S_b (dash-dotted red line) remains low, M_b stays in the low state, and the CI is high (black line). After the wave sweeps over the cell, S_f decreases while M_f remains high for several minutes, resulting in cellular memory and an elevated CI.

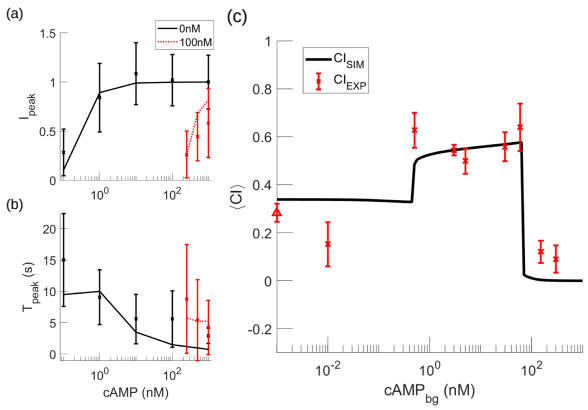


Figure 2.4. (a-b): Comparison between experimental results from Ref. [40] (symbols) and model results. Shown are the maximum change in intensity I_{peak} (a) and its corresponding time T_{peak} (b) as a function of the uniform change in cAMP concentration for two different cAMP pretreatment concentrations. (c): Average CI (red symbols) versus background cAMP concentration.

Eventually, however, the low values of S_f cause a transition of M_f to its low state and the CI decreases to zero (see below and Fig.2.5).

Intermediate values of $[\text{cAMP}]_{\text{bg}}$ ($[\text{cAMP}]_{\text{bg}}$=0.5-69 nM) result in higher values of S_f, maintaining M_f in a high state for the entire wave period (Fig. 2.3c-d). This increase in S_f is due to the feedback from M_f to S_f, which is, through the non-adaptive link, proportional to the receptor occupancy R_f. Since this occupancy is an increasing function of $[\text{cAMP}]_{\text{bg}}$, the feedback loop between S_f and M_f contributes to a sustained positive CI throughout the entire wave cycle (Fig. 2.3c-d), in agreement with the experimental results. This feedback from high M_f is also responsible for the increase of S_f in the back of the wave (Fig. 2.3c-d). As a result, the CI shows a distinct increase near the end of a cycle, which is, again, consistent with the experimental results.

For $[\text{cAMP}]_{\text{bg}}$>69 nM, both S_f and CI remain close to 0 during the entire cycle, which is consistent with the experimental results for $[\text{cAMP}]_{\text{bg}}$=150 and 300 nM. The reason for this is that the amplification of S_f due to the ultra-sensitivity of the response is reduced for increased values of $[\text{cAMP}]_{\text{bg}}$ (see below and Fig. 2.5). As a result, the value of S_f is not sufficiently large to bring M_f to its high state. Consequently, M_f remains in the low state, leading to a short and weak response of S_f and an overall low CI (Fig. 2.3e). Note that for our parameter values, the

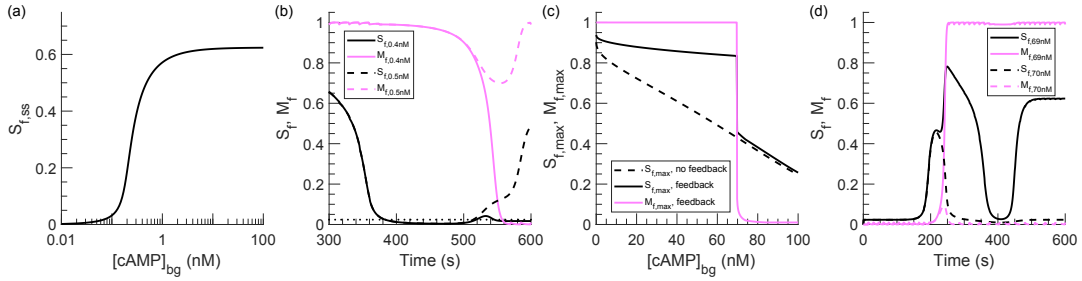


Figure 2.5. A: Steady-state value of S_f as a function of uniform $[\text{cAMP}]_{\text{bg}}$ for the case M_f=1. B: S_f and M_f as a function of time in the full model for a value of $[\text{cAMP}]_{\text{bg}}$ that does ($[\text{cAMP}]_{\text{bg}}$=0.5nM; dashed lines) and does not ($[\text{cAMP}]_{\text{bg}}$=0.4nM; solid lines) result in persistent memory. C: Maximum value of S_f and M_f in a wave as a function of $[\text{cAMP}]_{\text{bg}}$. There is a sharp, switch-like transition at $[\text{cAMP}]_{\text{bg}}\approx$69nM. D: Response of the full model for $[\text{cAMP}]_{\text{bg}}$ just below ($[\text{cAMP}]_{\text{bg}}$=69nM; solid lines) and just above ($[\text{cAMP}]_{\text{bg}}$=70nM; dashed lines) the switch-like transition.

dissociation constant for the receptors is $K_d \approx 408$nM. Thus, the absence of a strong chemotactic response for high values of $[\text{cAMP}]_{\text{bg}}$ is not due to receptor saturation but is directly linked to the bistability and the ultra-sensitivity of the pathway.

We also used the modeling results to compute the chemotactic index averaged in time over the entire wave cycle, $\langle \text{CI} \rangle$. The dependence of $\langle \text{CI} \rangle$ on $[\text{cAMP}]_{\text{bg}}$ has three distinct regimes (Fig. 2.4c). For very small values of $[\text{cAMP}]_{\text{bg}}$ ($[\text{cAMP}]_{\text{bg}} < 0.5$ nM) $\langle \text{CI} \rangle$ remains close to its value at $[\text{cAMP}]_{\text{bg}}$=0. For these values of $[\text{cAMP}]_{\text{bg}}$, the memory is only turned on during part of the wave cycle. In contrast, for intermediate values of $[\text{cAMP}]_{\text{bg}}$ (0.5 nM $\leq [\text{cAMP}]_{\text{bg}} \leq$ 69 nM), M_f is in the high state during the entire cycle, resulting in a nearly constant and high $\langle \text{CI} \rangle$. Finally, for $[\text{cAMP}]_{\text{bg}} >$ 69 nM, M_f is always in the low state and $\langle \text{CI} \rangle$ is close to 0. Notably, due to the bistable dynamics of our memory module, the transitions between these regimes are very abrupt and switch-like. To determine whether our experimental data also exhibited this switch-like behavior, we computed the average CI in the experiments. The results are plotted as symbols in Fig. 2.4c and are close to the results of the model (line). Most importantly, $\langle \text{CI} \rangle$ decreases in a switch-like fashion from as much as $\sim$0.6 at $[\text{cAMP}]_{\text{bg}}$=60nM to as little as $\sim$0.1 at $[\text{cAMP}]_{\text{bg}}$=150nM, in agreement with the bistable dynamics of the memory module.

To understand the transition in the model behavior for small values of $[\text{cAMP}]_{\text{bg}}$, shown

in Fig. 2.4c, it is useful to examine the response of S_f with $M_f = 1$. For this case, we plot in Fig. 2.5a the steady state solution for S_f as a function of uniform $[\text{cAMP}]_{\text{bg}}$(i.e., in the absence of a wave). S_f increases from 0 to approximately 0.6 as the background cAMP concentration increases. This increase is due to the feedback from M_f to S_f, which is proportional to the receptor occupancy R_f, an increasing function of $[\text{cAMP}]_{\text{bg}}$. The dependence of S_f on $[\text{cAMP}]_{\text{bg}}$ explains why for small values of $[\text{cAMP}]_{\text{bg}}$ the chemotactic index CI reduces to zero following the wave (see Fig. 2.3a and Fig. 2.4c). This is further illustrated in Fig. 2.5b where we plot S_f and M_f for $[\text{cAMP}]_{\text{bg}}$=0.4nM and $[\text{cAMP}]_{\text{bg}}$=0.5nM. For $[\text{cAMP}]_{\text{bg}}$=0.4nM, S_f remains close to the lower threshold of the bi-stable module (indicated by the dotted line), which causes M_f to transition from the high to the low state, resulting in a vanishing CI. For $[\text{cAMP}]_{\text{bg}}$=0.5nM, however, the feedback from M_f results in higher values of S_f such that the memory stays in its high state.

To analyze the switch-like behavior of Fig. 2.5, we can compute the maximum value of S_f in a wave as a function of $[\text{cAMP}]_{\text{bg}}$. The result is plotted in Fig. 2.5c, both for the full model (solid line) and for the case without feedback from M_f, i.e. $k_{S2} = 0$ (dashed line). As a result of the ultra-sensitivity of the equation, the asymmetry in the response of S is much larger than the asymmetry in the external gradient and the receptor occupancy. As is evident from the plot, this amplification of S_f is a decreasing function of $[\text{cAMP}]_{\text{bg}}$. Therefore, above a critical value of $[\text{cAMP}]_{\text{bg}}$, S_f is no longer able to cause a transition of M_f from the low to the high state. For the values of $[\text{cAMP}]_{\text{bg}}$ above the critical value, the CI will thus remain low (see Fig. 2.4). Furthermore, due to the bi-stability of the memory module, M_f exhibits a switch-like transition at a critical value of $[\text{cAMP}]_{\text{bg}}$ (magenta (light gray) line). This is also shown in Fig. 2.5d where we plot the response of the full model for a value of $[\text{cAMP}]_{\text{bg}}$ just below and just above the critical value. For $[\text{cAMP}]_{\text{bg}}$=70nM, above the critical value, M_f does not transition to the high state and S_f remains close to 0 for the second half of the wave. For $[\text{cAMP}]_{\text{bg}}$=69nM, on the other hand, M_f switches from its low to its high state, which leads to an increased response. After the wave passes, S_f decreases but then increases to $\approx$0.6 due to the persistent memory at

the front.

The switch-like transitions are also clearly visible when examining the phase diagram of the response of the model in the *a*-*b* parameter plane (corresponding to the two thresholds in the memory module). Here, we plot the time averaged difference between the memory at the front and back of the cell, $\langle M_f - M_b \rangle$, for different values of a and b using a color scale (Fig. 2.6). Since we take $a \leq b$ we only present one-half of the space. The parameter values corresponding to Table 2.1 and used in our study are indicated by the symbol. For $[\text{cAMP}]_{\text{bg}}$=0nM (left panel), the sharp transitions between three different regions are clearly visible. In the dark blue regions, corresponding to large values of b, neither the front nor the back memory is activated, resulting in $\langle M_f - M_b \rangle = 0$. For intermediate values of b and small values of a, the memory at the front is permanently activated while $M_b = 0$, which results in $\langle M_f - M_b \rangle = 1$ (yellow region). In the other regions, the front memory is activated only during part of the wave, giving rise to values of $\langle M_f - M_b \rangle$ that are between 0 and 1. Note that our parameters fall within this region of the phase space. For $[\text{cAMP}]_{\text{bg}}$=60nM, the sharp boundary of the yellow region has moved to larger values of a, such that our parameter values now fall within the region in which $\langle M_f - M_b \rangle = 1$ (middle panel). Finally, for $[\text{cAMP}]_{\text{bg}}$=150nM, the yellow region has almost completely disappeared and nearly the entire phase diagram, including our parameter set, correspond to $\langle M_f - M_b \rangle = 0$.

Summary

In summary, we find that the background concentration of the chemoattractant, $[\text{cAMP}]_{\text{bg}}$, has a profound effect on the cellular memory of chemotaxing *Dictyostelium* cells. For intermediate values of $[\text{cAMP}]_{\text{bg}}$ this memory is greatly enhanced, leading to substantially more efficient chemotaxis under periodic waves of cAMP. It is worth noting that the experimentally estimated value of $[\text{cAMP}]_{\text{bg}}$ during the natural aggregation process of *Dictyostelium* cells is $\sim$10nM [42]. This value is within the intermediate range, suggesting that aggregation may be facilitated by increased cellular memory due to accumulating cAMP. For larger values of $[\text{cAMP}]_{\text{bg}}$, the chemotactic response and cellular memory are suppressed. Our experimental results, and in

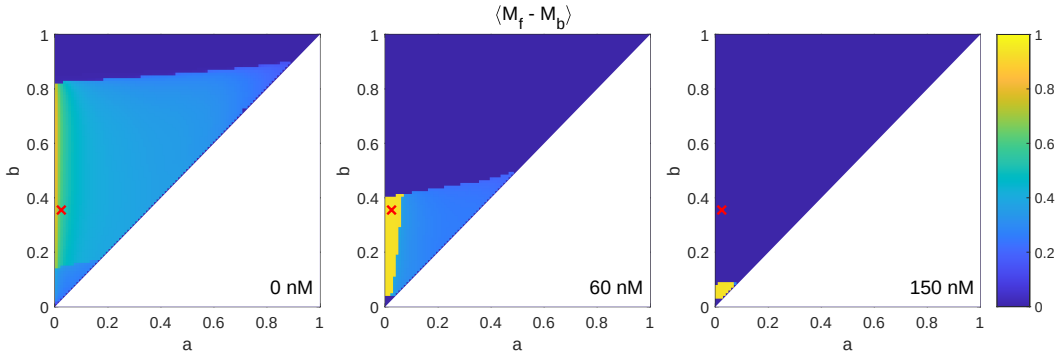


Figure 2.6. Phase diagram in the $a-b$ space quantifying $\langle M_f - M_b \rangle$ for three different values of $[\text{cAMP}]_{\text{bg}}$. Three regions, with sharp transitions, can be identified: $\langle M_f - M_b \rangle = 1$ (yellow regions) for which the front memory is always high while the back memory is always low, $\langle M_f - M_b \rangle = 0$ (dark blue regions) corresponding to low memory at the front and the back, and intermediate values of $\langle M_f - M_b \rangle$ for which the front memory is high during part of the wave. The values of a and b corresponding to our study are marked by a red X.

particular the switch-like behavior of the average CI, are fully consistent with our mathematical model. Crucial elements of this model are a bistable memory module, which allows cells to ignore the back of the wave, an ultra-sensitive response, responsible for the amplification of the chemoattractant gradient, and a direct, non-adaptive link between input signal and response, which explains the long-lasting memory for intermediate values of $[\text{cAMP}]_{\text{bg}}$. Future work will be required to identify the precise biochemical components that are responsible for the observed behavior.

2.1 Supplementary Information

Microfluidic device, device preparation, and experimental setup

The microfluidic system produces bell-shaped pulses of concentration that periodically sweep over cells on the substrate, thus emulating the spatial and temporal pattern of concentration in a periodic traveling wave. The microfluidic system is comprised of a microfluidic device and a flow-driving setup, which are modified versions of the device and setup described in our previous publication [3]. The microfluidic device consists of a micromachined polydimethylsiloxane (PDMS) chip and a #1.5 microscope cover glass substrate with a periodic micropattern of stripes

of polyethylene glycol (PEG) gel. The PEG-gel stripes are all $\sim$1.5 μm high and $\sim$30 μm wide and one of two substrate patterns are used: 1) one 25 μm wide glass (non-treated) stripe for every four 10 μm wide glass stripes or 2) variable width stripes, repeated every 6 stripes, ranging from $\sim 6\mu$m to $\sim 25\mu$m (Fig. 2.1d-e). We have verified that our results are independent of the stripe width. Because developed *Dictyostelium* cells cannot adhere to the surface of PEG-gel, adherent *Dictyostelium* cells are confined to glass stripes, making their migration along the 10 μm wide glass stripes nearly one-dimensional (1D) [44].

The microfluidic device has three inlets (A-C) and one outlet (see Fig. 2.1). Its main functional area is an $\sim$100 μm deep chemotaxis channel with a width of 2200 μm and an $\sim$7 mm long rectilinear part. At the entrance of the chemotaxis channel, a $\sim$200 μm wide stream of a concentrated solution of cAMP (coming from inlet C) is squeezed (hydrodynamically focused) between two streams of plain buffer or of cAMP solutions with a lower concentration (coming from the inlets A and B), which together fill the rest of the 2200 μm wide chemotaxis channel. As the focused stream of cAMP advances down the chemotaxis channel (along the y-axis), the diffusion of cAMP molecules across the boundaries of the stream gradually changes the lateral (across the flow, x-axis) profile of cAMP from its initial rectangular shape to a bell-shaped curve. At the standard experimental conditions, [cAMP] in the concentrated solution is 1400 nM, the mean flow velocity in the gradient channel is $\sim$300 μm/s, and the chemotaxis is observed $\sim$3 mm downstream from the gradient channel entrance, where the maximal [cAMP] is reduced to $\sim$1000 nM, and the effective width of the stream of concentrated cAMP (full width at half height) is increased to $\sim$230 μm, both due to diffusion of cAMP across the stream [3]. The microfluidic system operates in a cyclic fashion. In the beginning of a cycle, the volumetric flow rate of the stream of buffer (or low-concentration cAMP solution) from inlet B to the right of the cAMP stream is low, and the volumetric rate of flow of buffer from inlet B to the left of the cAMP stream is high, thus placing the cAMP stream close to the right wall of the gradient channel. The flow rate of the buffer stream on the right linearly increases with time, while the flow rate of the buffer on the left linearly decreases with time, with the total flow rate remaining

nearly unchanged. As a result, the high concentration cAMP stream (which is squeezed between the buffer streams) linearly drifts from right to left, and, at a given y-axis position, a bell-shaped pulse of cAMP is sweeping from right to left (in the negative x-axis direction) at a constant speed. The length of the sweep is l≈1425 μm that for a duration of the sweep T=10 min results in a speed of 2.5 μm/s for the motion of the cAMP pulse across the gradient channel (note that it takes 30s for the wave to be reset, such that the effective duration is 570s). At the end of a sweep, the flow of the concentrated cAMP solution is turned off, making the concentrated cAMP to completely disappear from the gradient channel, and then the flow rates from inlets A and B restored to their initial values. After that, the flow of the concentrated cAMP solution is turned on again, making the stream of concentrated cAMP reappear at its initial location close to the right wall of the gradient channel, and the next cycle begins. Importantly, the cAMP stream is effectively moved from its final to initial location without directing it (even transiently) anywhere else in the chemotaxis channel. Also, because of the relatively large length of the sweep, and fast decay of [cAMP] with the distance from the peak of the bell-shaped profile, there is ∼500 μm wide region near the middle of the gradient channel, where [cAMP] is very close to the background concentration, both in the beginning and end of each cycle. As a result, all cells in this region are exposed to the same periodically repeating spatiotemporal pattern of [cAMP] and can be pooled and analyzed together, after accounting for a time difference within a cycle of $\Delta t = \Delta x/v_x$, where Δx is the distance between cells along the x-axis. Therefore, large sets of data on chemotactic indices (CIs) and velocities of individual cells at different time points (phases) of the repeated cAMP wave cycle can be collected.

Further details on the operation of the microfluidic device and flow setup and on the experimental procedure as well as a detailed mathematical analysis are provided in SI of Ref. [3]. The following modifications were made to the original experimental setup and procedure of Ref. [3]. (1) Because flow in the microfluidic device is driven and controlled by applying differential pressures between the inlets and the outlets, it is necessary to have substantial fluidic resistances between the inlets and the outlet. In the original device, those resistances were

implemented in 100 μm deep and 60 μm wide channels, making the device easy to fabricate (all channels had the same depth), but requiring the resistance channel lines to be long, because of their relatively low fluidic resistance per unit length. In the present device, the resistance lines are $\sim$30 μm deep and $\sim$30 μm wide channels, which have an $\sim$40 larger fluidic resistance per unit length, greatly reducing the footprint of the microchannel network (at the expense of making the fabrication of the device somewhat more involved). (2) The reduced footprint of the microchannel network makes it possible to seal the PDMS chip against the coverglass with the micropattern of PEG-gel stripes using the application of vacuum ($\sim$ -30 kPa) to a deep (100 μm) and wide (4 - 6.5 mm) O-shaped grove around the microchannel network that acts as a vacuum cup. This vacuum-assisted sealing is essential, because it allows keeping the micropatterned substrate wet at all times, thus preserving the structure and functionality of the PEG-gel stripes (that *Dictyostelium* cells cannot adhere to). (3) The possibility of attaching the microfluidic chip to a wet substrate also enables direct plating of *Dictyostelium* cells onto the substrate. The PDMS chip is attached to the substrate only after cells settle and adhere to glass stripes. As a result, cell plating is substantially simplified and streamlined as compared with the original microfluidic device, where cells needed to be delivered to the gradient channel through a dedicated inlet and connecting microchannel that are both absent in the present device.

Microfluidic chips were cast in polydimethylsiloxane (PDMS, Sylgard 184 by Dow corning) using a microfabricated master mold, a 5 inch silicon wafer with a micro-relief of UV-curable SU8 photoresist (by MicroChem) produced with UV photolithography [44]. First, an $\sim$30 μm thick layer of SU8 2015 photoresists was spin-coated onto the wafer, pre-backed, exposed to collimated UV-light through a photomask, and post-baked. Next, the wafer was spin-coated with SU8 2050 photoresist to a total thickness of 100 μm, pre-baked, exposed to UV-light through a second photomask, post-baked, and developed, revealing a micro-relief with 30 μm and 100 μm tall micro-ridges. The 30 μm deep micro-grooves on the PDMS replica of the mold produced the 30$\times$30 μm resistance channels, and 100 μm deep micro-grooves produced all other flow channels and the O-shaped vacuum cups. The PDMS cast was cut into

individual chips, and the inlet and outlet holes in the chips were punched with a Luer stub.

To verify the wave profile computed from numerical simulations of the diffusion process [3], we added Alexa Fluor^{TM} 594 Hydrazide (Invitrogen), a fluorescent dye, to the cAMP solution. The fluorescent intensity profile with $[\text{cAMP}]_{\text{bg}}$=0 was well fitted using a Gaussian

$$\text{cAMP}_{\text{Gauss}}(x) = Ae^{\frac{-(x-x_0)^2}{2\sigma^2}}$$

where the wave peak position was given by x_0, the wave width was given by σ and its amplitude by A. Using the numerical simulations, this amplitude was determined to be 1000 nM at the experimental field of view [3]. An example of a fit is given in Fig. 2.1b-c. From these fits, we determined σ = 115 μm, corresponding to a full width at half maximum of 230 μm. Furthermore, by using a linear fit to x_0 as a function of time, we verified that the wave speed is v = 2.5 μm/s. For non-zero $[\text{cAMP}]_{\text{bg}}$ the cAMP concentration was the Gaussian profile plus a constant equal to the value of $[\text{cAMP}]_{\text{bg}}$ such that the total cAMP profile is given by $\text{cAMP}(x) = \text{cAMP}_{\text{Gauss}}(x) + \text{cAMP}_{\text{bg}}$.

Cell tracking

A machine learning (ML) model was developed for reliably tracking the complex movement of the *Dictyostelium* cells. 20 videos of cells moving in response to chemical gradients were provided for training of the model. In order to avoid having too many repeated or similar images, only one in every thirty frames was used for training and validating the model, resulting in 140 images. Each of the images, of size 256x256 pixels representing 25% of the full field of view, was assumed to be completely independent. Images were converted to grayscale and normalized. The model used was a U-Net convolutional neural network, which can be trained on very few images and outperforms most existing methods [172]. The model consisted of three rounds of convolution and max pooling, followed by a fully connected layer with 512 elements, and finally three rounds of convolution and upsampling. The convolution layers were 3x3 with

32, 64, and 128 filters. Rectified linear unit (Relu) activation was applied after each convolution. The pooling and upsampling layers were 2x2. Data augmentation, including shifting and flipping of the images, was used during model training to increase the size of the dataset. A labeled dataset was obtained by marking the centroid of cells in a small number of images using human input. A U-Net model was then trained on this small sample and then used to aid in marking further images. This process was repeated several times until the entire dataset was marked. The labeled dataset was then split with 80% of all images used for training, and 20% used for validation set. In order to avoid target leakage, the model's parameters and hyperparameters were reset and trained only on the new training dataset. Data augmentation was again used to increase the size of the training dataset. The final model consistently tracked 90% of all cells in the validation dataset. Finally, the model was tested on a new set of videos not seen in the original dataset. A comparison between blinded, manual marking of cells and the ML algorithm showed that approximately 97% of cells were tracked by the automated procedure (5623 vs. 5487).

Once cell positions were determined, we computed cell trajectories by connecting nearest points in consecutive frames. Only cells that were within a 300 μm wide region at the middle of the gradient channel were considered. These cells experienced a full cAMP wave, starting and ending at the background concentration. Furthermore, we only included cells that were continuously recognized for more than 3.5 min, that were moving with an average speed of at least 0.16 μm/min, and that were at least 33 μm from the neighboring cells (these cells are marked by red circles in Fig. 2.1). In addition, although rarely observed, cells moving on the PEG-gel stripes were also excluded. Data from the first cAMP wave was excluded to eliminate any initial differences in cells.

The cell's velocity in the x-direction for a given frame, V_x, was computed as the difference in the x-position 3 frames prior and 3 frames after the given frame, divided by the time interval between these frames (90s). The time with respect to the passage of the cAMP peak (phase) was assigned by measuring the distance of the cell from the location of the peak of the wave

Table 2.1. Model parameters

Parameter (Units)	Value	Parameter (Units)	Value
k_R	0.002463 $nM^{-1} \cdot s^{-1}$	k_{-R}	1.006 s^{-1}
k_E	0.159 s^{-1}	k_{-E}	0.159 s^{-1}
k_I	0.139 s^{-1}	k_{-I}	0.139 s^{-1}
k_{basal}	0.000015 nM	k_{Mem}	0.2550 $nM^{-2} \cdot s^{-1}$
a	0.0239 nM	b	0.3550 nM
k_S	2000 s^{-1}	K_{m1}	0.018 nM
k_{-S}	2048 s^{-1}	K_{m2}	0.001 nM
k_{S2}	149.4 s^{-1}	K_{m3}	0.0366 nM
M_f^{tot}	1	M_b^{tot}	1
R_f^{tot}	1	R_b^{tot}	1
S_f^{tot}	1	S_b^{tot}	1

relative to the extent of the sweep. The CI as a function of time was computed by averaging data points in equally spaced bins of width 0.5 min. The time averaged CI, $\langle \mathrm{CI} \rangle$, was computed by averaging CI in time over the entire wave cycle. The time average of V_x, $\langle V_x \rangle$, was calculated in a similar fashion. Experiments were repeated at least three times for each $[\mathrm{cAMP}]_{\mathrm{bg}}$. For each $[\mathrm{cAMP}]_{\mathrm{bg}}$, a distribution (N = 10,000 trials) of sample mean was generated by bootstrapping over the replicate experiments, and the experimental averages and error bars reported in this study were the mean and standard deviation of this resampled distribution.

Parameter fitting

As mentioned in the main text, the loss function $L = \sum_{n=1}^{N} |x_{exp} - x_{sim}|/(N\sigma_{exp})$ is minimized to find the optimal parameters. This loss function compares N=54 experimental data points to simulation data (see main text). We take 46 discrete points from previous experimental data (detailed in Ref. [3]) and 8 points derived from the current study. As in Ref. [3], σ_{exp} was taken as the standard deviation of the experimental data or 50% of x_{exp} where standard deviations were not available.

Following Ref. [3], the 46 data points using previous studies were chosen as follows: 24 data points were taken from uniform dose response data where cells were exposed to uniform

stimulus switching from 0nM or 100nM to higher concentrations [40]. The peak amplitudes, peak times and the state of memory module M were fitted (see also Fig. 2.4a-b). Additionally, 15 data points were chosen from gradient-to-uniform data obtained from experiments in which cells were first exposed to a gradient (0-100 nM across a 70 μm wide channel), followed by a uniform concentration of 0, 1, 10, 20 or 50nM [3]. These experiments, detailed in Ref. [3], quantified the localization of activated Ras, a directional sensing marker, using the Ras-binding domain of Raf tagged to GFP (RBD-GFP). The times for disappearance and reappearance as well as the ratios of fluorescence intensity before and after the switching from gradient to uniform concentration were fitted. As indicated from the experiments, we required the memory M_f to stay high (i.e., $M_f = 1$) for a uniform concentration of 10, 20 and 50 nM and to become low ($M_f = 0$) for 0 nM . The final 7 data points were chosen from gradient reversal data where the initial gradient was switched to a reversed gradient of 100-0 nM, 75-25 nM or 60-40 nM across the gradient chamber. The ratios of fluorescent RBD-GFP intensity before and after the switching were fitted. The memory was required to reverse in the 100 to 0nM profile and remain unchanged in the 60 to 40nM profile. Further details of these data points can be found in Ref. [3].

The "new" 10 data points were chosen from the current background experiments with $[\text{cAMP}]_{\text{bg}}$= 0, 0.01, 0.5, 60 and 150 nM. To match with the experimental results, which showed that cells have a high CI during the front of the wave for $[\text{cAMP}]_{\text{bg}}$= 0 and 0.01 nM, persistent high CI for $[\text{cAMP}]_{\text{bg}}$= 0.5 and 60 nM, and significantly reduced CI for $[\text{cAMP}]_{\text{bg}}$= 150nM, we required the front memory, M_f, at 2.5 minutes before the peak of the wave, when cells experienced a positive gradient, to be high for $[\text{cAMP}]_{\text{bg}}$= 0, 0.01, 0.5 and 60nM and low for 150nM. At 5 minutes after the peak, at a near-zero cAMP concentration, M_f was required to be low for $[\text{cAMP}]_{\text{bg}}$= 0, 0.01 and 150nM and high for $[\text{cAMP}]_{\text{bg}}$= 0.5 and 60nM.

We fitted the five parameters (a, b, k_{Mem}, k_{S2} and Km_3) related to the memory module using the lower bound constraints of [0.001, 0.001, 0.005, 20, 0.0001] and upper bound constraints of [1, 1, 5, 200000, 1] in units of [nM, nM, $\text{nM}^{-2}\text{s}^{-1}$, s^{-1}, nM]. To minimize L we used simulated annealing, which avoids trapping in local minima of the parameter space, and

uses an artificial temperature that is decreased during the parameter search such that sampling becomes more and more selective. In each iteration, a new parameter set was generated from the previous one with a step size based on this artificial temperature. The difference in the value of the loss function between the new and previous iteration was then computed. For a negative difference the new parameter set was automatically accepted while for a positive difference the new parameter set was accepted with a chance that ranged between 0 to 50% and which was based on the loss function difference and the artificial temperature. The search was carried out in log space to cover the wide range of parameter values and the algorithm was concluded when the difference in loss function became smaller than 0.01 for more than 500 iterations. Multiple runs of simulated annealing were conducted with randomized initial parameter values. At the end of each run, pattern search was applied to obtain more precise local minima, using Matlab routine patternsearch. In each iteration, this algorithm attempts parameters values at a fixed step from the previous parameters. If the loss function was reduced, the step size would be doubled in order to explore a wider space. If the loss function did not reduce, the step size would be halved. The parameter search terminated when both the change in loss function became smaller than 0.01 and the absolute step size became smaller than 0.0001 in log10 space.

The final parameters are listed in Table 2.1 and correspond to $L = 0.556$. As in Ref. [3] (Fig. S5b in that study), we can compare the simulation results with experimental results of Takeda *et al.* [40], which recorded the intensity of RBD-GFP following a sudden change in uniform cAMP concentration. In Fig. 2.4a-b, we plot the numerically computed intensity peak, I_{peak}, and the time corresponding to the maximum peak amplitude, T_{peak} as lines. These results are in good agreement with the experimental results, plotted as symbols. When we consider the loss function L' that includes only the 46 data points from the previous experiments, the current parameter set gives $L' = 0.655$ while the original parameter set results in $L' = 0.731$.

Chapter 2, in full, is a reprint of the material as it appears in the Physical Review E 2021. Karmakar, R., Tang, M.H., Yue, H., Lombardo, D., Karanam, A., Camley, B.A., Groisman, A. and Rappel, W.J., American Physical Society, 2021. The dissertation author was a co-author of

this paper.

Chapter 3

Eukaryotic chemotaxis under periodic stimulation shows temporal gradient dependence

Introduction

Chemotaxis, the chemically guided motion of cells, is critical to several biological processes such as foraging, wound healing, embryonic developement, and cancer metastasis [23, 24, 25, 26, 27]. The social amoeba *Dictyostelium discoideum* is a well-characterized model organism to study chemotaxis. It displays a unicellular to multicellular transition in its life cycle when starved of nutrients, by aggregating through chemotaxis [173, 149]. Chemotaxis of *Dictyostelium* cells is driven by the chemoattractant cyclic Adenosine Mono Phosphate (cAMP), a small molecule that is internally synthesized and secreted by the cells [174]. The multicellular aggregate then forms a stalk and a fruiting body, which contains spores that can later be dispersed [175, 176].

During the aggregation process, *Dictyostelium* cells not only produce the chemoattractant but also relay the signal, resulting in cAMP waves that sweep through the population [42]. This signal-relay process is also present in other chemotactic systems [177, 178] and ensures the recruitment of cells over large distances. Initially, the cAMP waves arise spontaneously from many locations in the population. As development continues, sources with high frequencies dominate and become stable centers to which cells aggregate [179]. A large number of models

of the aggregation process have been proposed, ranging from qualitative excitable system models to more biochemically oriented ones [165, 180, 181, 150, 182]. In addition, a significant number of modeling studies have been published which attempt to describe the chemotactic response of a single cell to chemoattractant gradients [156, 31, 158].

Traditionally, most experimental and modeling chemotaxis studies have focused on the response of cells to static gradients [183, 184]. More recently, using microfluidic devices, it has become possible to expose cells to carefully controlled complex and time-varying gradients [185, 167, 186, 39, 187, 4]. For example, experiments that generated chemoattractant waves of controlled speed have shown that cells can be more sensitive to the positive gradient in the incoming half of a wave and less sensitive to the negative gradient in the back half of the wave [3, 37]. The extent of this bias of cell motion has been shown to depend on the period of the chemoattractant waves [3, 37] as well as the background concentration of the chemoattractant [4]. These studies also support the involvement of a Local Excitation Global Inhibition (LEGI) module in chemotaxis [3, 37]. In this module, the stimulus produces both a membrane bound localized activator and a globally diffusible inhibitor, with fast and slow kinetics, respectively [32, 156].

These previous studies did not determine whether cell motion depends on temporal sensing. Furthermore, it did not address the question of the exposure to multiple waves, a question relevant given the aggregation process of *Dictyostelium.* In this study, we investigate the effect of exposure to multiple waves of identical amplitude and frequency on the chemotactic ability of *Dictyostelium* cells. Our results show that this ability is enhanced upon exposure to multiple waves of a short period (6 min) and remains low and unchanged when the period is long (15 and 20 min). Furthermore, using modeling, we show that our result supports a mechanism for temporal gradient sensing and rectification, working in parallel with a mechanism for spatial gradient sensing in the form of a LEGI model, and explains the observed trend in the chemotactic ability of the cells.

We exposed cells of the axenic strain AX4 [168] that are developed for 5 hours to multiple

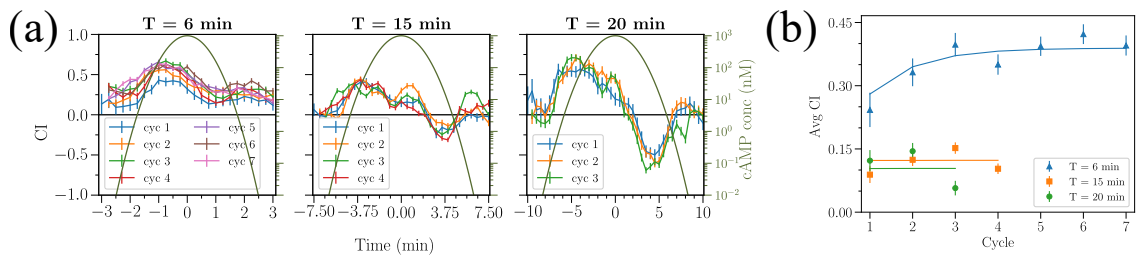


Figure 3.1. (a) Chemotactic Index (CI) as a function of time for time periods $T = 6$ min, 15 min, and 20 min, each with several cycles. Each data point represents the average CI of all the cells in the time bin-width Δt. Error bars represent the standard error of the mean. Δt is taken to be the same as the frame rate of the image capture, i.e., 15 s for $T = 6$ min and 30 s for $T = 15$ min and 20 min. Also shown is the cAMP profile concentration for each wave period. (b) Average CI plotted against cycle number for all periods. Markers represent an average over all tracked cells and error bars represent the standard error of the mean. Lines correspond to the fits from the model.

identical cycles of cAMP waves with a uniform speed. These traveling waves were generated using a microfluidic device, detailed in an earlier study [4]. In short, a stream of cAMP is swept across an observation channel, resulting in a bell-shaped wave profile that is similar to the one measured for natural waves of cAMP [165, 166]. The speed of the wave can be controlled such that each wave sweeps through the substrate in a fixed time period T. The period was set to either 6 minutes (hereafter referred to as short period), corresponding to a physiologically relevant value [34], or to higher values, namely 15 and 20 minutes (hereafter referred to as long periods). Cells were plated on a micropatterned substrate consisting of one dimensional patterns, which was also used in an earlier work [44]. As a result of this patterning, cells were confined to move along only one dimension, either along or opposite the direction of wave propagation, facilitating cell tracking. For each wave period, we carried out experiments on at least three different days.

The motion of the cells was captured by Differential Interference Contrast (DIC) microscopy. The cells in the resulting images were smoothened, segmented through edge detection, and binarized. Then, cell tracks were constructed by identifying the nearest neighbor for each cell in the subsequent frame. Only isolated cells were considered for tracking to preclude the effects of crowding and cell contacts on chemotaxis. Furthermore, we only analyzed cells that

were tracked for at least two successive waves. The total number of cells tracked for each cycle and period ranged from 125 to 648. To keep the total time experimentation time equivalent, we exposed cells to 7, 4, and 3 cycles of T = 6, 15, and 20 minutes, respectively.

Once the tracks of isolated cells were constructed, we computed a measure of the chemotactic ability of the cells called the Chemotactic Index (CI), defined as the ratio of the velocity of the cell in the direction of the source to the magnitude of its speed, averaged over a moving time-window of a certain length (2τ) [3]. It follows, then, that CI takes on values between -1 (motion exactly away from the source) and $+1$ (motion exactly towards the source). In the current one-dimensional context, where cells move only in the x-direction, the CI of a cell at time t is the ratio of its displacement to its path length over the duration 2τ.

$$\mathrm{CI}(t) = \frac{x(t+\tau) - x(t-\tau)}{\int_{t-\tau}^{t+\tau} |dx|} \tag{3.1}$$

In this study, we have taken the moving window duration 2τ to be 2 minutes. After obtaining the time-series, we computed the average chemotactic index, $\langle \mathrm{CI} \rangle$, for each cycle, which allowed us to determine the dependence of $\langle \mathrm{CI} \rangle$ on the number of waves experienced by the cells.

Figure 3.1-a shows the CI vs time data for all the wave periods and cycles used in the study. For the $T = 6$ minute waves, the CI increases uniformly during the first few cycles, and then saturates. In contrast, the CI for the longer wave periods remains almost unchanged for each cycle. This is also evident in Figure 3.1-b, which plots the average CI as a function of the wave cycle. From this figure, it is clear that $\langle \mathrm{CI} \rangle$ increases substantially (by about 50%) for $T = 6\,\mathrm{min}$ and it remains roughly the same for the longer wave periods. We have verified that the increase in CI is due to an increase in the directionality of the cell motion: the (undirected) cell speed is roughly the same $(\sim 3-4\,\mu\mathrm{m\,min}^{-1})$ for all cycles and periods (see Fig 3.4).

To determine whether the increase in the chemotactic ability was due to an increase in development time, we repeated the experiments using cells that were developed for only four hours. These cells also exhibited a significant increase in $\langle \mathrm{CI} \rangle$ for $T = 6\,\mathrm{min}$ waves, similar to 5 h

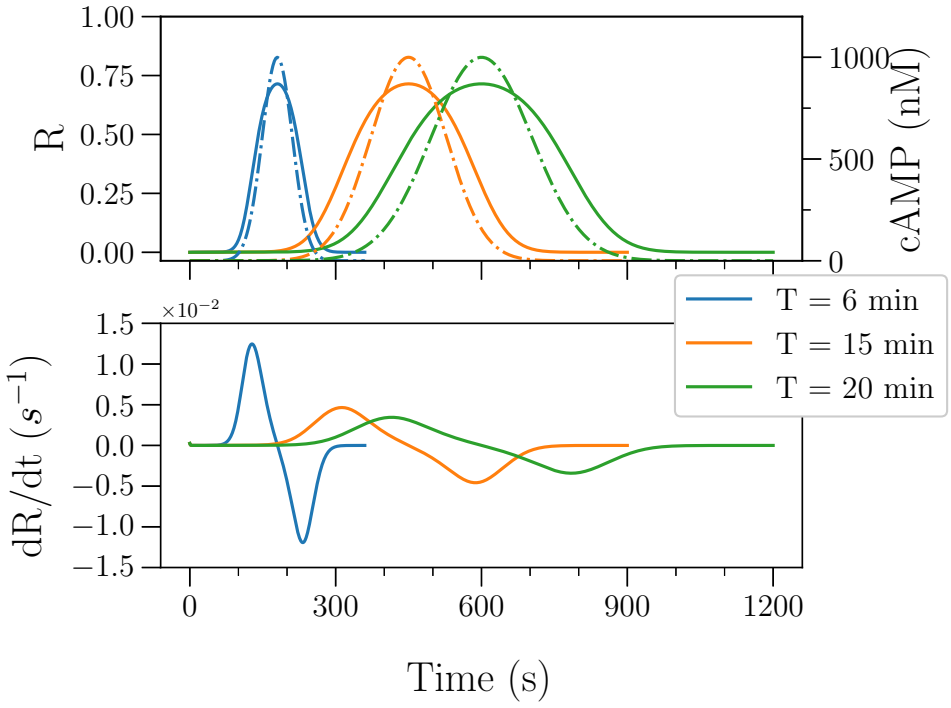


Figure 3.2. Variation with time of the concentration and temporal gradient of R for different periods, for one cycle. cAMP concentration is shown in the top panel with dashed lines. For all wave periods, cAMP and R have the same amplitude, but since the wave takes different times to sweep through the substrate, their temporal gradients are different. A wave with a short period has a higher temporal gradient, as shown in the bottom panel. This plot shows the values at the front of the cell; those at the back take on similar values with a small time lag.

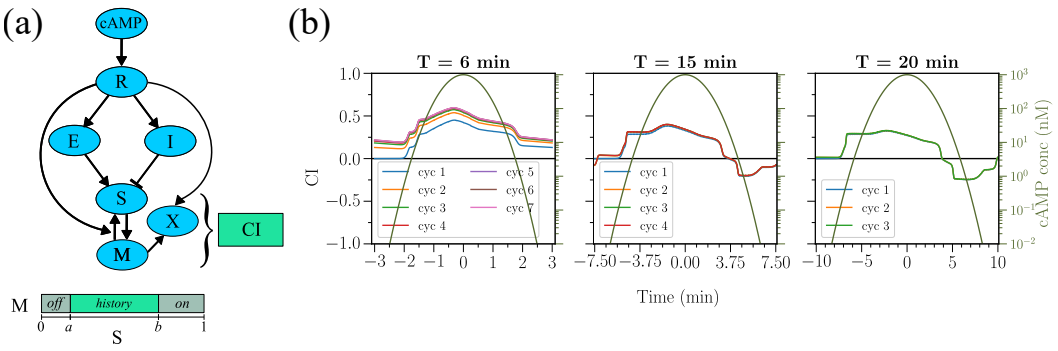


Figure 3.3. (a) Schematic of the model, by extending the LEGI+M model to include X that is activated by the temporal gradient of cAMP with a threshold. See Section 4 in the SI for equations. (b) CI vs time for multiple periods and all cycles as predicted from the model ($\alpha = \beta = \gamma = 0.15$). The plots overlap for $T = 15$ min and 20 min.

developed cells (see Fig. 3.5). We should note that carrying out experiments for cells that were developed for 6 hours or more was challenging since these cells displayed increased adhesion and started to clump into small aggregates. In summary, our experiments show that for short periods, the chemotactic ability of cells markedly improved as the number of waves increased, independent of the development time. In contrast, for longer wave periods, the chemotactic ability was found to be independent of the wave cycle.

We next attempted to address our experimental findings within a modeling framework. The starting point for this is our previous model, which consists of a LEGI module, together with a bistable memory module M [3] (see Figure 3.3-a). In this LEGI+M model, external cAMP binds to the receptor R, which then activates both a membrane-bound activator, E, and a global inhibitor I. The membrane-bound response element S of the LEGI module is activated by E and inhibited by I and feeds into the memory module. The component of this module feeds back to S and this feedback depends on E. This model was able to explain how cells are able to chemotax toward the wave source, even though the spatial gradient reverses direction in the back of the wave. It was shown that for short period waves the memory at the front, but not the back, is activated, resulting in a continued response in the direction of the original wave [3]. Specifically, the CI was assumed to be determined by a linear combination of the difference in front and back

value of M and S. Importantly, in this study, the parameter values of the LEGI and M modules are identical to the ones used in our previous study and are listed in Table 3.1.

Here we extend the model with a new, local signaling component, X, whose dynamics is assumed to be a function of the fraction of bound receptors R and the memory module M. (Figure 3.3-a, see Section 3.1 for explicit equations). The model is implemented in a one-dimensional geometry, consisting of a line, representing the interior of the cell, and the two ends corresponding to the back (b) and front (f) of the cell. Thus, all components, including X, are solved for both at the front and at the back, except for the global inhibitor I. This global component diffuses within the interior of the cell and we will assume that the diffusion coefficient for I is large, so that its concentration can be taken as uniform within the cell. Finally, the definition of CI is extended to include a contribution from X in addition to M and S: $\mathrm{CI} = \alpha(M_f - M_b) + \beta(S_f - S_b) + \gamma(X_f - X_b)$, where α, β, and γ are parameters that were adjusted to fit the experimental data.

We propose the kinetics of X in order to increase the chemotactic response for cells that are exposed to multiple short period waves but not for long period waves. This will be accomplished if $X_f - X_b$ grows as a function of cycle number only for $T = 6$ minute waves. A simple way to arrest the growth of X for large wave periods is to introduce a time scale for the decay of X that is larger than the short period but smaller than the long period. This would readily work if the activation terms for X for both short and long period waves lead to similar increases in the level of X after each cycle. A closer look at the wave profile reveals that such an activation cannot depend on the magnitude of the cAMP concentration or any other quantity that directly depends on it, such as the receptor fraction R. Figure 3.2 (top panel) shows the concentration of cAMP (dashed lines) and R (bold lines) a cell experiences as a function of time for all three wave periods. Since the difference between the short and long period waves is the speed with which the wave moves over the cells, the profiles of cAMP and R have the same amplitude but different widths (in time). Thus, cells exposed to long period waves experience high values of cAMP concentration for a prolonged period of time. Therefore, if the activation of X is proportional to

the cAMP concentration or depend on a threshold value, than it will reach higher values for long periods than for short periods. Since the chemotactic response is proportional to X, this increase will then lead to an elevated CI, in disagreement with our experimental findings. Further details of models that fail to predict the observed trend are described in Section 3.1.

Cells sense the cAMP concentration through receptor binding, and examining the profiles of R experienced by cells reveals that the temporal gradient of R is greater for a short wave period than for a long period (see Figure 3.2, bottom panel). This suggests that the response of cells may involve the temporal gradient of the R. Therefore, we propose that X is activated only by a temporally increasing cAMP. Specifically, we assume that the dynamics of X at the front of the cell can be written as

$$\frac{\mathrm{d}X_f}{\mathrm{d}t} = \frac{k_1 M_f}{1+\exp\Big(-p(\frac{\mathrm{d}R_f}{\mathrm{d}t}-\theta)\Big)} - k_2 X_f \tag{3.2}$$

with a similar equation for X_b, the component at the back. Here, M, X, and R are the respective de-dimensionalized concentrations of the respective components in the model. They take values between 0 and 1 at the front as well as at the back. The first term describes the activation of X through a sigmoidal function of the temporal concentration $\frac{\mathrm{d}R}{\mathrm{d}t}$ of the activated receptor with a positive threshold θ, which acts as a rectifier and places a lower bound on the temporal gradient for the activation of X_f to occur. The parameter p is the multiplicative factor that controls the steepness of the sigmoidal function. The second term describes the decay of X_f following first-order kinetics.

We simulated the model shown in Figure 3.3-a with the above expression for the dynamics of X for the three different wave periods. For the threshold parameter θ we chose a value such that X only gets activated for the short wave period ($\theta = 5\times10^{-3}\,\mathrm{s}^{-1}$). The parameter k_1 determines the magnitude of X while the value of k_2 determines the timescale of the decay of X in one cycle, and therefore of the slow increase of X over several cycles. For the chosen value of θ, the activation term of X is appreciably different from 0 for approximately 50 s, equating to roughly 14% of the total cycle duration for $T = 6\,\mathrm{min}$. It is then easy to determine that a value

of $k_2 = 2.5 \times 10^{-3}\,\mathrm{s}^{-1}$ results in an increase in average CI that saturates after 6 cycles. Since the values of $\frac{dR}{dt}$ and θ are of the order of $1 \times 10^{-3}\,\mathrm{s}^{-1}$, p is chosen to be equal to $1 \times 10^{3}\,\mathrm{s}$ to make the argument of the sigmoid of the order of unity. To determine the values of α, β, and γ, we chose the simplest case of giving M, S, and X equal weightage in the definition of CI. The values $\alpha = \beta = \gamma = 0.15$ resulted in the best fit with the experimental data. We should note that the value of γ is dependent on k_1 which determines the amplitude of the activation of X.

The resulting CI for the different wave periods is shown in Fig. 3.3-b as a function of time. Consistent with the experimental results, the CI for the short wave period shows in increase for the first several cycles and then saturates. In contrast, the CI for the long periods is unchanged for each cycle. The difference between the short and long period responses can be understood when examining the dynamics of $X_f - X_b$ as a function of time (Figure 3, SI). For the short period waves, the time derivative of R exceeds the threshold value θ. Since the memory M at the front of the cell is non-zero and is zero at the back, only the front value of X increases. The decay of X_f is not rapid enough to reset it to its original value, resulting in an increase of X_f for the first few wave cycles, after which its mean value no longer changes. For the long periods, the time derivative does not exceed the threshold value and X does not accumulate appreciably. Thus, for these periods, $X_f - X_b$ does not increase and does not contribute significantly to CI. This is also evident from the computed average CI for each cycle, shown as solid lines in Fig. 3.3-b. The model is able to replicate both the experimentally observed slow increase in CI data as well as the cycle-independent response for long periods.

Our experimentally observed enhancement of chemotactic ability of cells due to repeated exposure to waves has clear relevance to the aggregation of *Dictyostelium*. When waves of cAMP initially arise spontaneously from random locations, they pass through the cells in all directions; the period of these waves is usually large. Over time, the wave frequency and the cAMP concentration rise [34, 182], and the center with the highest frequency or the lowest period (generally $T \sim 6\,\mathrm{min}$), dominates the rest [179]. Then, waves with a fixed direction and frequency pass through cells. At this point, enhanced chemotaxis under periodic stimulation

would help cells to reach the wave center sooner. This, combined with our earlier observations that the chemotactic ability improves as the background cAMP concentration increases [4], then leads to more optimal aggregation and, thus, better chances of survival.

Our proposed model has modules for both spatial and temporal gradient sensing that operate in parallel and independently. Temporal gradient sensing is the determinant player for bacteria, which are too small and move too fast to employ spatial sensing [188]. Since eukaryotes are large enough to sense and respond to spatial gradients [189, 190], most studies have focused on their response to spatial gradients. Our study uniquely provides evidence for the dependence of chemotaxis in *Dictyostelium* on the temporal dependence of chemoattractant gradient. These results are consistent with a recent study, which shows that migrating myeloid cells can also sense temporal dynamics of chemoattractant concentrations [45].

Our temporal sensing is formulated in terms of abstract variables, without specific identification of biochemical components. Such identification is challenging since a large number of components play a role in chemotaxis [191, 192] and additional studies are required to determine the exact biochemical components. Furthermore, we have chosen the simplest functional forms for activation thresholding and decay kinetics in Equation 3.2. More elaborate schemes may be possible and future work is needed to investigate this. Finally, it would be possible to further probe the temporal sensing module experimentally. Specifically, if the width of the Gaussian wave profile is decreased, keeping all else equal, the magnitude of the temporal gradient increases. Our model predicts that for these narrower profiles the CI will increase as a function of cycle number more rapidly for the short wave periods and may even increase for longer wave periods.

3.1 Supplementary Information

Experimental Setup

We used cells of the axenic *Dictyostelium discoideum* strain AX4, transformed with LimE-GFP (delta coil LimE-GFP) and Coronin-RFP (LimE GFP/corA RFP). Cells were grown in submerged shaking culture in HL5 medium (35.5 g HL5 powder (Formedium, Norfolk, UK) and 10 ml penicillin-streptomycin (10,000 U/ml; Gibco, Thermo Fisher Scientific, USA) per liter of DI water. When cells reached their exponential growth phase ($3-4\times10^6$ cells/mL), they were harvested by centrifugation at 3000 rpm for 5 minutes, resuspended in KN_2/Ca^{2+} buffer (14.6 mM KH_2PO_4, 5.4 mM Na_2HPO_4, 100 µM $CaCl_2$, pH 6.4), and washed once with the same buffer. The cells were resuspended in KN_2/Ca^{2+} at 10^7 cells/mL and developed for 4 and 5 hours with pulses of 50 nM cAMP added every 6 minutes.

We exposed cells to repeated waves of cAMP and recorded their movement. We performed the experiments at three different wave speeds, corresponding to a wave period of 6, 15, and 20 minutes. The microfludic wave generator is identical to the one used in an earlier study [4].

Differential Interference Contrast (DIC) images were taken every 15 seconds (for 6 minute waves) or 30 seconds (for 15 and 20 minute waves) in four fields of view spanning the width of the chemotaxis channel, 2800 µm away from the cAMP inlet, on a spinning-disk confocal Zeiss Axio Observer inverted microscope using a 10X objective and a Roper Cascade QuantEM 512SC camera. Images were captured by using Slidebook 6 (Intelligent Imaging Innovations).

Cells were plated on a glass substrate that was micropatterned with ~ 1.5 µm thick stripes of cell adhesion-blocking polyethylene glycol (PEG) gel, as detailed in Ref.[44]. The pattern consists of 4 narrow (~ 10 µm) and 1 wide (~ 25 µm) untreated glass stripes separated by 30 µm wide nonadhesive PEG-gel stripes. These substrates limit the adhesion and migration of *Dictyostelium* cells to $\sim 6-25$ µm wide stripes of non-PEG treated glass oriented in the

x-direction, along the gradient and perpendicular to the flow. Thus, cell migration was effectively one-dimensional (1D), either up or down the gradient (positive or negative x-direction). This greatly simplified the collection and analysis of data as compared to 2D chemotaxis on a standard glass substrate.

Cell and Wave Tracking

The image analysis pipeline from DIC movies to computing chemotactic indices for the tracked cells is as follows: For each image in the movie, we applied the Scharr filter from Python's Scikit-Image package. The filtered image is further smoothened using Gaussian blurring and was binarized to separate the cells from the background. This method of pre-processing DIC images is similar to that used in Ref. [193]. To segment the cells, we started at a pixel labeled as a cell but not segmented thus far and built a region around it by recursively including other hitherto uncounted cell-pixels that were contiguous with the current region. The region stopped growing once all contiguous pixels were identified. This process was repeated until all the cell-pixels are counted and numbered. To avoid objects that were too small or too large, we chose a lower and upper cutoff for the area of individual and isolated cells, based on a histogram of areas of all identified cells.

In order to construct tracks of individual cells, the region of a given cell in the current frame is separately used as a filter on the next frame. The region in the next frame with the highest overlap is considered to correspond to the given cell. Among the constructed cell tracks, only tracks that last at least two periods and at least three quarters of the duration of each period were considered for the next step. This was done so that we tracked the same set of cells over a prolonged period of time. Furthermore, if at any point in the track, a cell gets closer than a cutoff distance of 12 μm to another cell, that time point is excluded from the track. This is done to preclude any effects of cAMP produced by neighboring cells on chemotaxis. Once the cell tracks were constructed and selected, the chemotactic index (CI) of each cell as a function of time was calculated using Equation 1 in the main text. The average CI of the population for a

given wave was then computed as the area under the CI-time curve divided by the number of bins, averaged over all the tracks in that cycle.

To visualize the cAMP wave, a fluorescent dye, Alexa Fluor 594 Hydrazide (Invitrogen), was added to the central stream along with cAMP at a concentration of 1000 nM. The fluorescence intensity in each frame was fitted to a univariate Gaussian profile $I(x) = A\exp\left(-\frac{(x-b)^2}{2\sigma^2}\right) + d$, where x is the distance from the peak of the wave along direction of propagation, and $A,\ b,\ \sigma,\ d$ are, respectively, the amplitude, position of the peak, width of the Gaussian, and the background intensity. $I(x)$ is the intensity along the direction of the propagation after averaging in the perpendicular direction. While simulating the wave, we used the function

$$\text{cAMP}(x,t) = A\exp\left(-\frac{(x-vt)^2}{2\sigma^2}\right) \tag{3.3}$$

where $A = 1000\,\text{nM}$, $v = 1400\,\mu\text{m}/T$ where T is the wave period in seconds ($1400\,\mu\text{m}$ is the distance a wave moves in each period from start to finish), $\sigma = 120\,\mu\text{m}$.

Trends in Chemotactic Index

The mean value of Chemotactic Index (defined in Equation 1 in the main text) in a cycle increases upon exposure to multiple waves if the wave period T is short and remains low and unchanged if T is large. This is due to an increase in the directionality of cells when T is small; the (undirected) cell speed remains roughly the same ($\sim 3-4\ \mu\text{m}\,\text{min}^{-1}$) for all cycles and periods, as shown in Figure 3.4.

When the chemotaxis experiments were repeated with cells that were developed for four hours, the cells still showed a significant increase in the mean chemotactic index for the $T = 6\,\text{min}$ waves, similar to the cells developed for five hours, suggesting that the increase in the chemotactic ability is in response to exposure to the chemoattractant waves, and not merely the passage of time. See Figure 3.5.

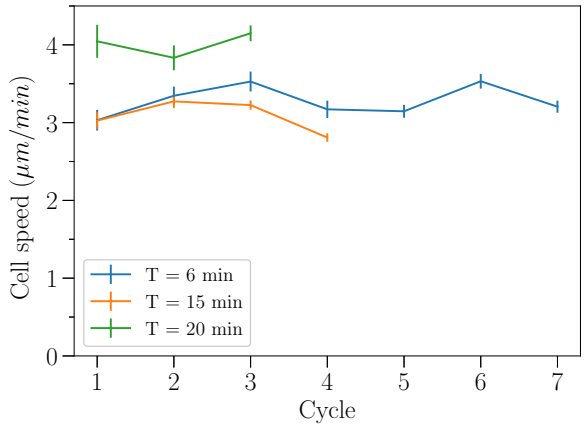


Figure 3.4. The average undirected cell speeds remain roughly the same for all cycles and wave periods. Error bars in the plot denote the standard error of the mean.

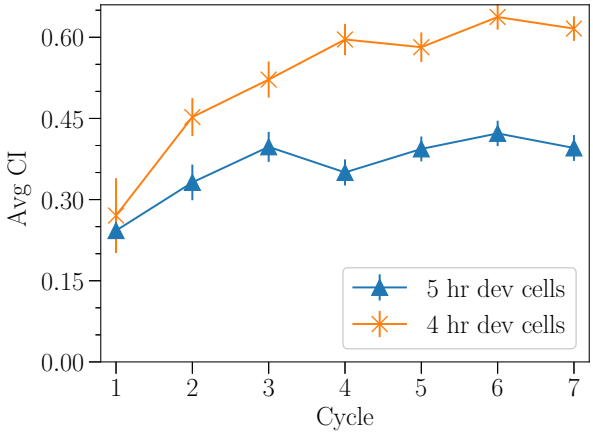


Figure 3.5. Average CI increases as a function of cycle number when cells were developed for four or five hours prior to the exposure of cAMP waves.

Model Description

The complete model used is an extension of the LEGI+M model along with Equation 2 of the main text. We used the LEGI+M model as detailed in an earlier study [3]. The entire set of equations and the parameters are reproduced here for completeness.

External cAMP activates the receptors on the surface of the cell. Two kinds of receptors R_1 and R_2, differing in their binding affinities and constitutive activations are considered. The two types of receptors have the same downstream effect, so their sum $R = R_1 + R_2$ is considered in the rest of the model. R activates E and I, which respectively excite and inhibit the response element S following zeroth-order ultrasensitivity kinetics, so that a weak gradient in R and E is amplified to a strong gradient in S. S also promotes the switching of the bistable memory module M from 0 (low) to 1 (high), which feeds back to activate S, particularly on the back half of the wave. The new component X is activated when M is high and when the temporal gradient of R is above a positive threshold θ. All the components of the model except I are localized at the membrane. Thus they take two values, one at the front and one at the back, denoted by the subscripts f and b respectively. Equation 3.4 shows the equations governing the kinetics of (global) I, the model components at the front of the cell, and the definition of the chemotactic index (CI). A corresponding set of equations exists for the components at the back of the cell.

$$
\begin{aligned}
\frac{\mathrm{d}R_{1f}}{\mathrm{d}t} &= k_{R1}(\mathrm{cAMP}_f + r_1)(R_1^{\mathrm{tot}} - R_{1f}) - k_{-R1}R_{1f} \\
\frac{\mathrm{d}R_{2f}}{\mathrm{d}t} &= k_{R2}(\mathrm{cAMP}_f + r_2)(R_2^{\mathrm{tot}} - R_{2f}) - k_{-R2}R_{2f} \\
R_f &= R_{1f} + R_{2f} \\
\frac{\mathrm{d}E_f}{\mathrm{d}t} &= k_E R_f - k_{-E}E_f \\
\frac{\mathrm{d}I}{\mathrm{d}t} &= k_I\left(k_{\mathrm{basal}} + \frac{R_f + R_b}{2}\right) - k_{-I}I \\
\frac{\mathrm{d}M_f}{\mathrm{d}t} &= -k_{\mathrm{mem}}M_f(M_f - M^{\mathrm{tot}})\left(M_f - M^{\mathrm{tot}}\frac{b - S_f}{b - a}\right) \\
\frac{\mathrm{d}S_f}{\mathrm{d}t} &= k_S E\frac{S^{\mathrm{tot}} - S_f}{K_{m1} + S^{\mathrm{tot}} - S_f} - k_{-S}I\frac{S_f}{K_{m2} + S_f} + k_{S2}M_f R_f\frac{S^{\mathrm{tot}} - S_f}{K_{m3} + S^{\mathrm{tot}} - S_f} \\
\frac{\mathrm{d}X_f}{\mathrm{d}t} &= \frac{k_1 M_f}{1 + \exp\left(-p(\frac{\mathrm{d}R_f}{\mathrm{d}t} - \theta)\right)} - k_2 X_f \\
\mathrm{CI}(t) &= \alpha(M_f - M_b) + \beta(S_f - S_b) + \gamma(X_f - X_b)
\end{aligned}
\tag{3.4}
$$

The parameters of the model are shown in Table 3.1. The concentration of cAMP is measured in nM and the rest of the components are de-dimensionlized or unitless.

The model equations are integrated using MATLAB's ODE solver `ode23tb` [194], which is an adaptive Runge-Kutta method with a variable step size. The cell is assumed to be one-dimensional with a length of $L = 10\,\mu\mathrm{m}$. The cAMP stimulus is provided according to Equation 3.3. The center of the cell is assumed to be at $x = 0$ and t ranges from $-T/2$ to $T/2$ for each cycle. To prevent M from getting stuck at the extrema $M = 0$ and $M = 1$, a "kick" is given to M every 15 s in the simulation: M is set to 0.05 if $M = 0$ and to 0.95 if $M = 1$.

Spatial versus Temporal Sensing

We have argued in the main text, along with Figure 2 therein, why a model for the activation of X that depends on the *concentration* of cAMP or any other component downstream of it fails to reproduce the experimental trend in $\langle \mathrm{CI} \rangle$. Here we explicitly show the results of

Table 3.1. Parameters of the extended LEGI+MX model. The values of M^{tot}, S^{tot} are set to 1 at the front as well as at the back.

Parameter (Units)	**Value**	**Parameter (Units)**	**Value**
k_{R1} ($\text{nM}^{-1}\text{s}^{-1}$)	2.67×10^{-3}	k_{-R1} (s^{-1})	0.16
k_{R2} ($\text{nM}^{-1}\text{s}^{-1}$)	2.44×10^{-3}	k_{-R2} (s^{-1})	1.1
r_1 (nM)	0.012	r_2 (nM)	0.115
R_1^{tot}	0.1	R_2^{tot}	0.9
k_E (s^{-1})	0.159	k_{-E} (s^{-1})	0.159
k_I (s^{-1})	0.139	k_{-I} (s^{-1})	0.139
k_{basal}	1.5×10^{-5}	k_M (s^{-1})	0.360
a	0.058	b	0.5
k_S (s^{-1})	2000	k_{m1}	1.8×10^{-2}
k_{-S} (s^{-1})	2048	k_{m2}	1×10^{-3}
k_{S2} (s^{-1})	116.7	k_{m3}	0.127
k_1 (s^{-1})	2.5×10^{-2}	k_2 (s^{-1})	2.5×10^{-3}
α	0.15	β	0.15
γ	0.15	θ (s^{-1})	5×10^{-3}
p (s)	1×10^{3}		

integrating a model where the rate equation for X obeys simple mass action kinetics dependent on concentrations. Consider the model with an equation for X_f given by

$$\frac{\mathrm{d}X_f}{\mathrm{d}t} = k_1 M_f (X^{\text{tot}} - X_f) - k_2 X_f \tag{3.5}$$

with a similar equation for X_b. Equation 3.4 is integrated along with this equation, and the resulting time series of $X_f - X_b$ is plotted as a function of time in Figure 3.6 (top panel). The levels of $X_f - X_b$ for long periods exceed that for the short period. As explained in the main text, this happens because in the case of long periods, the cells perceive high levels of M for a prolonged amount of time, and that leads to stronger activation. This is contrary to our purpose of introducing X as a means to explain rising $\langle \text{CI} \rangle$ for short period only. On the other hand, the temporal model shows, in Figure 3.6 (bottom panel), that the values of $X_f - X_b$ stay higher for the short period than for the long periods.

Chapter 3, in full, is a reprint of the material as it is submitted for publication and as it

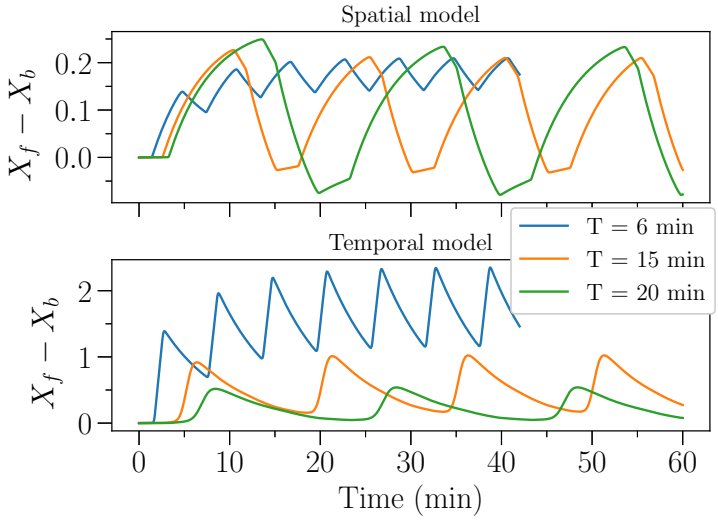


Figure 3.6. Comparison of time series of $X_f - X_b$ when the kinetics of X is given by equation 3.5 of the spatial model (above) and the temporal model (below). For the spatial model, $k_1 = 1.0 \times 10^{-3}\,\mathrm{s}^{-1}$, $k_2 = 2.5 \times 10^{-3}\,\mathrm{s}^{-1}$, and $X^{\mathrm{tot}} = 1$. The parameters for the temporal model are as defined in Table 3.1.

may appear in the Physical Review Letters, 2024. Karmakar, R., Karanam, A.R., Tang, M.H. and Rappel, W.J., American Physical Society, 2024. The dissertation author was the primary investigator and author of this paper.

Chapter 4

Conclusion

In the preceeding chapters, I described new studies on guard cell dynamics leading to stomatal closure in plants [19], and on gradient sensing mechanisms in *Dictyostelium* leading to chemotaxis [4, 41]. Through a combination of modeling and experiment, we advanced our knowledge of the mechanisms by which these processes are driven. The choice of model in each study is informed, among other considerations, by the size of the reaction pathway, availability of quantitative biochemical information, and the desired precision from the model predictions. On the modeling front, software for analyzing cell motion and wave progression in chemotaxis experiments, and for the design and simulation of Boolean networks was developed. The latter was packaged in a user-friendly manner and shared with the community.

Large signaling networks with incomplete information are common in biology in general, and plant physiology. As an example, the published ABA signaling network discussed in chapter 1 contains more than 80 components [18]. The vast majority of the interaction strengths and kinetic parameters between these components is not known, making it difficult to formulate mathematical models of these networks. Motivated by the simplicity and utility of Boolean networks[6, 195] and the challenges associated with formulating detailed rate equation-based models for these large networks, we have presented here Boolink [19], a software package with a graphical user interface (GUI) that can simulate, visualize, and plot the results of a user-defined Boolean network. Boolink is free to use and distribute, and is built from free and open-source

software. The Supplementary Text contains detailed instructions on downloading, installing, and running Boolink on Windows, MacOS, and Linux-based machines. In addition to this open-source version, we also packaged the software in a *Docker container*, which allows execution of Boolink in an even more facile and direct computer operating system-independent fashion.

Using Boolink, we reproduced the findings of the model published in Ref. [18] and then extended the network to include the influence of CO_2 on stomatal closure. Our proposed additions to the existing ABA network, which illustrate the potential use of Boolink[19], are meant as a starting point for further explorations and further research is needed to determine the precise mechanism by which CO_2 signaling merges with abscisic acid signal transduction. Nevertheless, several improvements can be suggested. First, it is conceivable that CO_2 affects yet unknown mechanisms. Second, the CO_2 pathway may contain feedback loops, which can be easily implemented within Boolink. Finally, we should point out that Boolean networks do not incorporate explicit rate constants and contain nodes that can only take one of two values (0 or 1). Therefore, these networks are not able to address the time-dependence of responses nor how they respond to graded inputs.

In chapter 2, we investigated the effect on the chemotactic ability of *Dictyostelium* cells upon adding a non-zero background concentration of the chemoattractant in addition to a wave stimulus[4]. We find that the background concentration of the chemoattractant has a profound effect on the cellular memory of chemotaxing cells. As the background concentration of cAMP is raised, the average of the chemotactic index (CI), which is a measure of the orientation of the cell towards the wave source, undergoes two qualitative changes. Firstly, at intermediate cAMP concentrations, $\langle \mathrm{CI} \rangle$ increases from its value at zero background. At very high cAMP concentrations, chemotaxis, and thus $\langle \mathrm{CI} \rangle$ shuts down altogether. In the model, the rise in $\langle \mathrm{CI} \rangle$ is seen as a result of sustained activation of the bistable memory M at intermediate cAMP concentrations. At very high background cAMP levels, the response element S decays to zero after a transient rise and does not activate M at all. The two switch-like transitions in $\langle \mathrm{CI} \rangle$ predicted by the experiments are in very good agreement with the experiments. It is worth noting

that the experimentally estimated value of $[\text{cAMP}]_{\text{bg}}$ during the natural aggregation process of *Dictyostelium* cells is $\sim$10nM [42]. This value is within the intermediate range, suggesting that aggregation may be facilitated by increased cellular memory due to accumulating cAMP. Future work will be required to identify the precise biochemical components that are responsible for the observed behavior.

In chapter 3, the effect on chemotaxis of exposure to multiple identical chemoattractant waves was studied. This study uncovered the presence of a temporal gradient sensing mechanism that raises the chemotactic ability of cells, measured as CI, through a rise in the concentration of a component X in the model over multiple waves [41]. The increase in the chemotactic ability of cells upon exposure to multiple waves was observed when the wave moves faster, i.e., the period T is small, but not when the wave moves slower, i.e., the period T, and so, the time of exposure to a gradient is large. This strongly hints at the role of a rate of change instead of exposure time in building long-term memory. Any mechanism for the development of long-term memory that is based on *concentrations* of components upstream of X will act stronger for the large wave periods, as we have shown.

A temporal gradient sensing ability has not been reported in *Dictyostelium* before to the best of our knowledge. A study on migrating myeloid cells has shown that those cells too can sense temporal dynamics of chemoattractant concentrations [45]. Further studies are required to understand the temporal gradient sensing mechanism in greater detail. We proposed potential tests to the hypothesis: if the width of the cAMP wave is reduced keeping the same amplitude the model predicts a a greater rise in the levels of X – a narrower wave implies a faster rise in concentration, which implies a greater rise in the levels of X. Secondly, the mechanism and the molecular players involved in temporal gradient sensing are unexplored.

Taken together, the three studies demonstrate that quantitative modeling of signaling processes in biological systems can be successfully employed to uncover the topology of signaling pathways and the nature of interactions among the components in the reaction network. The models can be improved through an iterative process of generating predictions from models and

testing them by experiments, which inform the next steps in modeling.

In the studies on eukaryotic chemotaxis, the predictions also show quantitative agreement with the experiments. It is important to highlight the progress that is achieved in understanding biology by going across the boundaries of traditionally defined disciplines, by approaching it through concepts and tools developed in physics, mathematics, and computation.

Bibliography

[1] Ludovic Ouellet, Keith J Laidler, and Manuel F Morales. Molecular kinetics of muscle adenosine triphosphatase. *Archives of Biochemistry and Biophysics*, 39(1):37–50, 1952.

[2] Thomas D Pollard. Rate constants for the reactions of ATP-and ADP-actin with the ends of actin filaments. *The Journal of cell biology*, 103(6):2747–2754, 1986.

[3] Monica Skoge, Haicen Yue, Michael Erickstad, Albert Bae, Herbert Levine, Alex Groisman, William F Loomis, and Wouter-Jan Rappel. Cellular memory in eukaryotic chemotaxis. *Proceedings of the National Academy of Sciences*, 111(40):14448–14453, 2014.

[4] Richa Karmakar, Man-Ho Tang, Haicen Yue, Daniel Lombardo, Aravind Karanam, Brian A Camley, Alex Groisman, and Wouter-Jan Rappel. Cellular memory in eukaryotic chemotaxis depends on the background chemoattractant concentration. *Physical Review E*, 103(1):012402, 2021.

[5] Ryan N Gutenkunst, Joshua J Waterfall, Fergal P Casey, Kevin S Brown, Christopher R Myers, and James P Sethna. Universally sloppy parameter sensitivities in systems biology models. *PLoS computational biology*, 3(10):e189, 2007.

[6] Julian D Schwab, Silke D Kühlwein, Nensi Ikonomi, Michael Kühl, and Hans A Kestler. Concepts in Boolean network modeling: What do they all mean? *Computational and Structural Biotechnology Journal*, 2020.

[7] Stuart A Kauffman. Metabolic stability and epigenesis in randomly constructed genetic nets. *Journal of theoretical biology*, 22(3):437–467, 1969.

[8] René Thomas. Boolean formalization of genetic control circuits. *Journal of theoretical biology*, 42(3):563–585, 1973.

[9] Ilya Shmulevich, Edward R Dougherty, and Wei Zhang. From Boolean to probabilistic Boolean networks as models of genetic regulatory networks. *Proceedings of the IEEE*, 90(11):1778–1792, 2002.

[10] Franziska Herrmann, Alexander Groß, Dao Zhou, Hans A Kestler, and Michael Kühl. A boolean model of the cardiac gene regulatory network determining first and second heart field identity. *Plos one*, 7(10):e46798, 2012.

[11] Stefan Bornholdt. Boolean network models of cellular regulation: prospects and limitations. *Journal of the Royal Society Interface*, 5(suppl_1):S85–S94, 2008.

[12] Fangting Li, Tao Long, Ying Lu, Qi Ouyang, and Chao Tang. The yeast cell-cycle network is robustly designed. *Proceedings of the National Academy of Sciences*, 101(14):4781–4786, 2004.

[13] Kai-Yeung Lau, Surya Ganguli, and Chao Tang. Function constrains network architecture and dynamics: A case study on the yeast cell cycle Boolean network. *Physical Review E*, 75(5):051907, 2007.

[14] Shintaro Munemasa, Felix Hauser, Jiyoung Park, Rainer Waadt, Benjamin Brandt, and Julian I Schroeder. Mechanisms of abscisic acid-mediated control of stomatal aperture. *Current opinion in plant biology*, 28:154–162, 2015.

[15] Sarah M Assmann and Timothy Jegla. Guard cell sensory systems: recent insights on stomatal responses to light, abscisic acid, and CO2. *Current opinion in plant biology*, 33:157–167, 2016.

[16] Agepati S Raghavendra, Vijay K Gonugunta, Alexander Christmann, and Erwin Grill. ABA perception and signalling. *Trends in plant science*, 15(7):395–401, 2010.

[17] Po-Kai Hsu, Guillaume Dubeaux, Yohei Takahashi, and Julian I Schroeder. Signaling mechanisms in abscisic acid-mediated stomatal closure. *The Plant Journal*, 105(2):307–321, 2021.

[18] Réka Albert, Biswa R Acharya, Byeong Wook Jeon, Jorge G T Zañudo, Mengmeng Zhu, Karim Osman, and Sarah M Assmann. A new discrete dynamic model of ABA-induced stomatal closure predicts key feedback loops. *PLoS biology*, 15(9):e2003451, 2017.

[19] Aravind Karanam, David He, Po-Kai Hsu, Sebastian Schulze, Guillaume Dubeaux, Richa Karmakar, Julian I Schroeder, and Wouter-Jan Rappel. Boolink: a graphical interface for open access Boolean network simulations and use in guard cell CO2 signaling. *Plant Physiology*, 187(4):2311–2322, 2021.

[20] Ebe Merilo, Pirko Jalakas, Hannes Kollist, and Mikael Brosché. The role of ABA recycling and transporter proteins in rapid stomatal responses to reduced air humidity, elevated CO2, and exogenous ABA. *Molecular Plant*, 8(4):657–659, 2015.

[21] Po-Kai Hsu, Yohei Takahashi, Shintaro Munemasa, Ebe Merilo, Kristiina Laanemets, Rainer Waadt, Dianne Pater, Hannes Kollist, and Julian I Schroeder. Abscisic acid-independent stomatal CO2 signal transduction pathway and convergence of CO2 and ABA signaling downstream of OST1 kinase. *Proceedings of the National Academy of Sciences*, 115(42):E9971–E9980, 2018.

[22] Li Zhang, Yohei Takahashi, Po-Kai Hsu, Hannes Kollist, Ebe Merilo, Patrick J Krysan, and Julian I Schroeder. FRET kinase sensor development reveals SnRK2/OST1 activation by ABA but not by MeJA and high CO2 during stomatal closure. *Elife*, 9:e56351, 2020.

[23] Shuvasree SenGupta, Carole A Parent, and James E Bear. The principles of directed cell migration. *Nature Reviews Molecular Cell Biology*, 22(8):529–547, 2021.

[24] Federica Sallusto and Marco Baggiolini. Chemokines and leukocyte traffic. *Nature immunology*, 9(9):949–952, 2008.

[25] Elena Scarpa and Roberto Mayor. Collective cell migration in development. *Journal of Cell Biology*, 212(2):143–155, 2016.

[26] Evanthia T Roussos, John S Condeelis, and Antonia Patsialou. Chemotaxis in cancer. *Nature Reviews Cancer*, 11(8):573, 2011.

[27] W. J. Rappel and W. F. Loomis. Eukaryotic chemotaxis. *Wiley Interdiscip Rev Syst Biol Med*, 1(1):141–149, 2009.

[28] H. C. Berg and E. M. Purcell. Physics of chemoreception. *Biophysical Journal*, 20:193–219, 1977.

[29] Wouter-Jan Rappel and Herbert Levine. The role of receptor occupancy noise in eukaryotic chemotaxis. In *Applications of Nonlinear Dynamics: Model and Design of Complex Systems*, pages 61–69. Springer, 2009.

[30] Uri Alon, Michael G Surette, Naama Barkai, and Stanislas Leibler. Robustness in bacterial chemotaxis. *Nature*, 397(6715):168–171, 1999.

[31] H. Levine, D. A. Kessler, and W. J. Rappel. Directional sensing in eukaryotic chemotaxis: a balanced inactivation model. *PNAS*, 103(26):9761–9766, 2006.

[32] C. A. Parent and P. N. Devreotes. A cell's sense of direction. *Science*, 284:765–770, 1999.

[33] Raymond E Goldstein. Traveling-wave chemotaxis. *Physical review letters*, 77(4):775, 1996.

[34] Thomas Gregor, Koichi Fujimoto, Noritaka Masaki, and Satoshi Sawai. The onset of collective behavior in social amoebae. *Science*, 328(5981):1021–1025, 2010.

[35] W Roos, V Nanjundiah, D Malchow, and G Gerisch. Amplification of cyclic-AMP signals in aggregating cells of dictyostelium discoideum. *FEBS letters*, 53(2):139–142, 1975.

[36] BM Shaffer. Secretion of cyclic AMP induced by cyclic AMP in the cellular slime mould Dictyostelium discoideum. *Nature*, 255(5509):549, 1975.

[37] Akihiko Nakajima, Shuji Ishihara, Daisuke Imoto, and Satoshi Sawai. Rectified directional sensing in long-range cell migration. *Nature communications*, 5:5367, 2014.

[38] Tim Lämmermann, Philippe V Afonso, Bastian R Angermann, Ji Ming Wang, Wolfgang Kastenmüller, Carole A Parent, and Ronald N Germain. Neutrophil swarms require LTB4 and integrins at sites of cell death in vivo. *Nature*, 498(7454):371–375, 2013.

[39] Bashar Hamza, Elisabeth Wong, Sachin Patel, Hansang Cho, Joseph Martel, and Daniel Irimia. Retrotaxis of human neutrophils during mechanical confinement inside microfluidic channels. *Integrative Biology*, 6(2):175–183, 2014.

[40] Kosuke Takeda, Danying Shao, Micha Adler, Pascale G Charest, William F Loomis, Herbert Levine, Alex Groisman, Wouter-Jan Rappel, and Richard A Firtel. Incoherent feedforward control governs adaptation of activated ras in a eukaryotic chemotaxis pathway. *Sci. Signal.*, 5(205):ra2–ra2, 2012.

[41] Richa Karmakar, Aravind Rao Karanam, Man-Ho Tang, and Wouter-Jan Rappel. Eukaryotic chemotaxis under periodic stimulation shows temporal gradient dependence. *bioRxiv*, pages 2023–10, 2023.

[42] KJ Tomchik and Peter N Devreotes. Adenosine 3', 5'-monophosphate waves in Dictyostelium discoideum: a demonstration by isotope dilution–fluorography. *Science*, 212(4493):443–446, 1981.

[43] Sonya Bader, Arjan Kortholt, and Peter JM Van Haastert. Seven Dictyostelium discoideum phosphodiesterases degrade three pools of cAMP and cGMP. *Biochemical Journal*, 402(1):153–161, 2007.

[44] Richa Karmakar, Christoph Schich, Nadine Kamprad, Vanessa Scheller, Edgar Gutierrez, Alex Groisman, Wouter-Jan Rappel, and Marco Tarantola. Novel micropatterning technique reveals dependence of cell-substrate adhesion and migration of social amoebas on parental strain, development, and fluorescent markers. *PloS one*, 15(7):e0236171, 2020.

[45] Caren E Petrie Aronin, Yun M Zhao, Justine S Yoon, Nicole Y Morgan, Thorsten Prüstel, Ronald N Germain, and Martin Meier-Schellersheim. Migrating myeloid cells sense temporal dynamics of chemoattractant concentrations. *Immunity*, 47(5):862–874, 2017.

[46] Hiroaki Kitano. Computational systems biology. *Nature*, 420(6912):206–210, 2002.

[47] International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature*, 409(6822):860–921, 2001.

[48] Mike Tyers and Matthias Mann. From genomics to proteomics. *Nature*, 422(6928):193–197, 2003.

[49] Guy Karlebach and Ron Shamir. Modelling and analysis of gene regulatory networks. *Nature reviews Molecular cell biology*, 9(10):770–780, 2008.

[50] Parul Maheshwari, Hao Du, Jen Sheen, Sarah M Assmann, and Reka Albert. Model-driven discovery of calcium-related protein-phosphatase inhibition in plant guard cell signaling. *PLoS computational biology*, 15(10):e1007429, 2019.

[51] M Morris Mano and Charles R Kime. *Logic and computer design fundamentals*. Prentice-Hall, Inc., 1997.

[52] Abhishek Garg, Alessandro Di Cara, Ioannis Xenarios, Luis Mendoza, and Giovanni De Micheli. Synchronous versus asynchronous modeling of gene regulatory networks. *Bioinformatics*, 24(17):1917–1925, 2008.

[53] Adrien Fauré, Aurélien Naldi, Claudine Chaouiya, and Denis Thieffry. Dynamical analysis of a generic Boolean model for the control of the mammalian cell cycle. *Bioinformatics*, 22(14):e124–e131, 2006.

[54] Elisabeth Remy, Paul Ruet, Luis Mendoza, Denis Thieffry, and Claudine Chaouiya. From logical regulatory graphs to standard petri nets: Dynamical roles and functionality of feedback circuits. In *Transactions on Computational Systems Biology VII*, pages 56–72. Springer, 2006.

[55] Carlos Espinosa-Soto, Pablo Padilla-Longoria, and Elena R Alvarez-Buylla. A gene regulatory network model for cell-fate determination during Arabidopsis thaliana flower development that is robust and recovers experimental gene expression profiles. *The Plant Cell*, 16(11):2923–2939, 2004.

[56] Assieh Saadatpour, István Albert, and Réka Albert. Attractor analysis of asynchronous Boolean models of signal transduction networks. *Journal of theoretical biology*, 266(4):641–656, 2010.

[57] Nicola Bonzanni, Abhishek Garg, K Anton Feenstra, Judith Schütte, Sarah Kinston, Diego Miranda-Saavedra, Jaap Heringa, Ioannis Xenarios, and Berthold Göttgens. Hard-wired heterogeneity in blood stem cells revealed using a dynamic regulatory network model. *Bioinformatics*, 29(13):i80–i88, 2013.

[58] René Thomas. Regulatory networks seen as asynchronous automata: a logical description. *Journal of theoretical biology*, 153(1):1–23, 1991.

[59] Julio Aracena, Eric Goles, Andrés Moreira, and Luis Salinas. On the robustness of update schedules in Boolean networks. *Biosystems*, 97(1):1–8, 2009.

[60] Henning Mortveit and Christian Reidys. *An introduction to sequential dynamical systems*. Springer Science & Business Media, 2007.

[61] Ilya Shmulevich, Edward R Dougherty, Seungchan Kim, and Wei Zhang. Probabilistic Boolean networks: a rule-based uncertainty model for gene regulatory networks. *Bioinformatics*, 18(2):261–274, 2002.

[62] Panuwat Trairatphisan, Andrzej Mizera, Jun Pang, Alexandru Adrian Tantar, Jochen Schneider, and Thomas Sauter. Recent development and biomedical applications of probabilistic Boolean networks. *Cell communication and signaling*, 11(1):1–25, 2013.

[63] Ranadip Pal, Aniruddha Datta, Michael L Bittner, and Edward R Dougherty. Intervention in context-sensitive probabilistic Boolean networks. *Bioinformatics*, 21(7):1211–1218, 2005.

[64] Béla Novák and John J Tyson. Design principles of biochemical oscillators. *Nature reviews Molecular cell biology*, 9(12):981–991, 2008.

[65] Maria Davidich and Stefan Bornholdt. The transition from differential equations to Boolean networks: a case study in simplifying a regulatory network model. *Journal of Theoretical Biology*, 255(3):269–277, 2008.

[66] Claudia Stötzel, Susanna Röblitz, and Heike Siebert. Complementing ODE-based system analysis using Boolean networks derived from an Euler-like transformation. *PloS one*, 10(10):e0140954, 2015.

[67] Dominik M Wittmann, Jan Krumsiek, Julio Saez-Rodriguez, Douglas A Lauffenburger, Steffen Klamt, and Fabian J Theis. Transforming Boolean models to continuous models: methodology and application to T-cell receptor signaling. *BMC systems biology*, 3(1):1–21, 2009.

[68] Jacques Demongeot, Adrien Elena, and Sylvain Sené. Robustness in regulatory networks: a multi-disciplinary approach. *Acta Biotheoretica*, 56(1):27–49, 2008.

[69] Shoudan Liang, Stefanie Fuhrman, and Roland Somogyi. Reveal, a general reverse engineering algorithm for inference of genetic network architectures. In *Biocomputing*, volume 3, 1998.

[70] Thomas M Cover. *Elements of information theory*. John Wiley & Sons, 1999.

[71] Harri Lähdesmäki, Ilya Shmulevich, and Olli Yli-Harja. On learning gene regulatory networks under the Boolean network model. *Machine learning*, 52(1):147–167, 2003.

[72] Ilya Shmulevich, Olli Yli-Harja, Jaakko Astola, and CG Core. Inference of genetic regulatory networks under the best-fit extension paradigm. In *IEEE-Eurasip Workshop on Nonlinear Signal and Image Processing*. Citeseer, 2001.

[73] Song Li, Sarah M Assmann, and Réka Albert. Predicting essential components of signal transduction networks: a dynamic model of guard cell abscisic acid signaling. *PLoS Biol*, 4(10):e312, 2006.

[74] Alfred V. Aho, Michael R Garey, and Jeffrey D. Ullman. The transitive reduction of a directed graph. *SIAM Journal on Computing*, 1(2):131–137, 1972.

[75] Julio Aracena, Adrien Richard, and Lilian Salinas. Number of fixed points and disjoint cycles in monotone Boolean networks. *SIAM journal on Discrete mathematics*, 31(3):1702–1725, 2017.

[76] Alan Veliz-Cuba and Reinhard Laubenbacher. On the computation of fixed points in Boolean networks. *Journal of Applied Mathematics and Computing*, 39(1):145–153, 2012.

[77] Peter Krawitz and Ilya Shmulevich. Basin entropy in Boolean network ensembles. *Physical review letters*, 98(15):158701, 2007.

[78] Fumito Mori and Atsushi Mochizuki. Expected number of fixed points in Boolean networks with arbitrary topology. *Physical review letters*, 119(2):028301, 2017.

[79] Sheldon M Ross. *Introduction to probability models*. Academic press, 2014.

[80] Jordan C Rozum, Jorge Gómez Tejeda Zañudo, Xiao Gan, Dávid Deritei, and Réka Albert. Parity and time reversal elucidate both decision-making in empirical models and attractor scaling in critical Boolean networks. *Science Advances*, 7(29):eabf8124, 2021.

[81] Jordan C Rozum, Dávid Deritei, Kyu Hyong Park, Jorge Gómez Tejeda Zañudo, and Réka Albert. pystablemotifs: Python library for attractor identification and control in Boolean networks. *Bioinformatics*, 38(5):1465–1466, 2022.

[82] Steffen Klamt, Julio Saez-Rodriguez, and Ernst D Gilles. Structural and functional analysis of cellular networks with CellNetAnalyzer. *BMC systems biology*, 1(1):1–13, 2007.

[83] Abhishek Garg, Alessandro Di Cara, Ioannis Xenarios, Luis Mendoza, and Giovanni De Micheli. Synchronous versus asynchronous modeling of gene regulatory networks. *Bioinformatics*, 24(17):1917–1925, 2008.

[84] István Albert, Juilee Thakar, Song Li, Ranran Zhang, and Reka Albert. Boolean network simulations for life scientists. *Source code for biology and medicine*, 3(1):1–8, 2008.

[85] Tomáš Helikar and Jim A Rogers. ChemChains: a platform for simulation and analysis of biochemical networks aimed to laboratory scientists. *BMC systems biology*, 3(1):1–15, 2009.

[86] Gautier Stoll, Eric Viara, Emmanuel Barillot, and Laurence Calzone. Continuous time Boolean modeling for biological signaling: application of Gillespie algorithm. *BMC systems biology*, 6(1):1–18, 2012.

[87] Loïc Paulevé. Pint: A static analyzer for transient dynamics of qualitative networks with ipython interface. In *International Conference on Computational Methods in Systems Biology*, pages 309–316. Springer, 2017.

[88] Christoph Müssel, Martin Hopfensitz, and Hans A Kestler. BoolNet—an R package for generation, reconstruction and analysis of Boolean networks. *Bioinformatics*, 26(10):1378–1380, 2010.

[89] Hannes Klarner, Adam Streck, and Heike Siebert. PyBoolNet: a python package for the generation, analysis and visualization of boolean networks. *Bioinformatics*, 33(5):770–772, 2017.

[90] Matthias Bock, Till Scharp, Chaitanya Talnikar, and Edda Klipp. BooleSim: an interactive Boolean network simulator. *Bioinformatics*, 30(1):131–132, 2014.

[91] Camille Terfve, Thomas Cokelaer, David Henriques, Aidan MacNamara, Emanuel Goncalves, Melody K Morris, Martijn van Iersel, Douglas A Lauffenburger, and Julio Saez-Rodriguez. CellNOptR: a flexible toolkit to train protein signaling networks to data using multiple logic formalisms. *BMC systems biology*, 6(1):1–14, 2012.

[92] Julian D Schwab and Hans A Kestler. Automatic screening for perturbations in Boolean networks. *Frontiers in physiology*, 9:431, 2018.

[93] A Gonzalez Gonzalez, Aurélien Naldi, Lucas Sanchez, Denis Thieffry, and Claudine Chaouiya. GINsim: a software suite for the qualitative modelling, simulation and analysis of regulatory networks. *Biosystems*, 84(2):91–100, 2006.

[94] Dirk Merkel. Docker: lightweight linux containers for consistent development and deployment. *Linux journal*, 2014(239):2, 2014.

[95] Frank Emmert-Streib, Matthias Dehmer, and Benjamin Haibe-Kains. Gene regulatory networks and their applications: understanding biological and medical problems in terms of networks. *Frontiers in cell and developmental biology*, 2:38, 2014.

[96] Mengyuan Zhao, Wenying He, Jijun Tang, Quan Zou, and Fei Guo. A comprehensive overview and critical evaluation of gene regulatory network inference technologies. *Briefings in Bioinformatics*, 22(5):bbab009, 2021.

[97] Fernando M Delgado and Francisco Gómez-Vela. Computational methods for gene regulatory networks reconstruction and analysis: A review. *Artificial intelligence in medicine*, 95:133–145, 2019.

[98] Mark WEJ Fiers, Liesbeth Minnoye, Sara Aibar, Carmen Bravo González-Blas, Zeynep Kalender Atak, and Stein Aerts. Mapping gene regulatory networks from single-cell omics data. *Briefings in functional genomics*, 17(4):246–254, 2018.

[99] Lian En Chai, Swee Kuan Loh, Swee Thing Low, Mohd Saberi Mohamad, Safaai Deris, and Zalmiyah Zakaria. A review on the computational approaches for gene regulatory network construction. *Computers in biology and medicine*, 48:55–65, 2014.

[100] Stuart Kauffman. Homeostasis and differentiation in random genetic control networks. *Nature*, 224(5215):177–178, 1969.

[101] Luis Mendoza and Elena R Alvarez-Buylla. Dynamics of the genetic regulatory network for Arabidopsis thaliana flower morphogenesis. *Journal of theoretical biology*, 193(2):307–319, 1998.

[102] Jacques Demongeot and Sylvain Sené. About block-parallel Boolean networks: a position paper. *Natural Computing*, 19(1):5–13, 2020.

[103] Jacques Demongeot, Eric Goles, Michel Morvan, Mathilde Noual, and Sylvain Sené. Attraction basins as gauges of robustness against boundary conditions in biological complex systems. *PloS one*, 5(8):e11793, 2010.

[104] Gonzalo A Ruz, Eric Goles, and Sylvain Sené. Reconstruction of Boolean regulatory models of flower development exploiting an evolution strategy. In *2018 IEEE Congress on Evolutionary Computation (CEC)*, pages 1–8. IEEE, 2018.

[105] Tania Timmermann, Bernardo González, and Gonzalo A Ruz. Reconstruction of a gene regulatory network of the induced systemic resistance defense response in Arabidopsis using boolean networks. *BMC bioinformatics*, 21(1):1–16, 2020.

[106] Corné MJ Pieterse, Christos Zamioudis, Roeland L Berendsen, David M Weller, Saskia CM Van Wees, and Peter AHM Bakker. Induced systemic resistance by beneficial microbes. *Annual review of phytopathology*, 52:347–375, 2014.

[107] Tania Timmermann, Grace Armijo, Raúl Donoso, Aldo Seguel, Loreto Holuigue, and Bernardo González. Paraburkholderia phytofirmans PsJN protects Arabidopsis thaliana against a virulent strain of Pseudomonas syringae through the activation of induced resistance. *Molecular Plant-Microbe Interactions*, 30(3):215–230, 2017.

[108] Tania Timmermann, María Josefina Poupin, Andrea Vega, Cristóbal Urrutia, Gonzalo A Ruz, and Bernardo González. Gene networks underlying the early regulation of Paraburkholderia phytofirmans PsJN induced systemic resistance in Arabidopsis. *PLoS One*, 14(8):e0221358, 2019.

[109] Rainer Storn and Kenneth Price. Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces. *Journal of global optimization*, 11(4):341–359, 1997.

[110] Gonzalo A Ruz, Tania Timmermann, and Eric Goles. Reconstruction of a grn model of salt stress response in Arabidopsis using genetic algorithms. In *2015 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*, pages 1–8. IEEE, 2015.

[111] Eugenio Azpeitia, Mariana Benítez, Iliusi Vega, Carlos Villarreal, and Elena R Alvarez-Buylla. Single-cell and coupled GRN models of cell patterning in the Arabidopsis thaliana root stem cell niche. *BMC systems biology*, 4(1):1–19, 2010.

[112] José Dávila Velderraín, Juan Carlos Martínez-García, and Elena R Álvarez-Buylla. Boolean dynamic modeling approaches to study plant gene regulatory networks: integration, validation, and prediction. In *Plant Gene Regulatory Networks*, pages 297–315. Springer, 2017.

[113] Liam Dolan, Kees Janmaat, Viola Willemsen, Paul Linstead, Scott Poethig, Keith Roberts, and Ben Scheres. Cellular organisation of the Arabidopsis thaliana root. *Development*, 119(1):71–84, 1993.

[114] Renan Pardal and Renze Heidstra. Root stem cell niche networks: it's complexed! Insights from Arabidopsis. *Journal of Experimental Botany*, 72(19):6727–6738, 2021.

[115] Yvonne Stahl, René H Wink, Gwyneth C Ingram, and Rüdiger Simon. A signaling module controlling the stem cell niche in Arabidopsis root meristems. *Current biology*, 19(11):909–914, 2009.

[116] Eugenio Azpeitia, Nathan Weinstein, Mariana Benítez, Luis Mendoza, and Elena R Alvarez-Buylla. Finding missing interactions of the Arabidopsis thaliana root stem cell niche gene regulatory network. *Frontiers in plant science*, 4:110, 2013.

[117] Julio Saez-Rodriguez, Luca Simeoni, Jonathan A Lindquist, Rebecca Hemenway, Ursula Bommhardt, Boerge Arndt, Utz-Uwe Haus, Robert Weismantel, Ernst D Gilles, Steffen Klamt, et al. A logical model provides insights into T cell receptor signaling. *PLoS computational biology*, 3(8):e163, 2007.

[118] Simone Gupta, Siddharth S Bisht, Ritushree Kukreti, Sanjeev Jain, and Samir K Brahmachari. Boolean network analysis of a neurotransmitter signaling pathway. *Journal of theoretical biology*, 244(3):463–469, 2007.

[119] Shubhank Sherekar and Ganesh A Viswanathan. Boolean dynamic modeling of cancer signaling networks: Prognosis, progression, and therapeutics. *Computational and Systems Oncology*, 1(2):e1017, 2021.

[120] Herman F Fumia and Marcelo L Martins. Boolean network model for cancer pathways: predicting carcinogenesis and targeted therapy outcomes. *PloS one*, 8(7):e69008, 2013.

[121] Hanna Hõrak, Maija Sierla, Kadri Tõldsepp, Cun Wang, Yuh-Shuh Wang, Maris Nuhkat, Ervin Valk, Priit Pechter, Ebe Merilo, Jarkko Salojärvi, and Kirk Overmyer. A dominant mutation in the ht1 kinase uncovers roles of map kinases and ghr1 in co2-induced stomatal closure. *The Plant Cell*, 28(10):2493–2509, 2016.

[122] Liina Jakobson, Lauri Vaahtera, Kadri Tõldsepp, Maris Nuhkat, Cun Wang, Yuh-Shuh Wang, Hanna Hõrak, Ervin Valk, Priit Pechter, Yana Sindarovska, Jing Tang, Chuanlei Xiao, Yang Xu, Uli Talas, Alfonso T Garcia-Sosa, Saijaliisa Kangasjarvi, Udo Maran, Maido Remm, M Rob G Roelfsema, Honghong Hu, Jaakko Kangasjarvi, Mart Loog, Julian I Schroeder, Hannes Kollist, and Mikael Brosche. Natural variation in arabidopsis cvi-0 accession reveals an important role of mpk12 in guard cell co2 signaling. *PLoS biology*, 14(12):e2000322, 2016.

[123] Honghong Hu, Aurélien Boisson-Dernier, Maria Israelsson-Nordström, Maik Böhmer, Shaowu Xue, Amber Ries, Jan Godoski, Josef M Kuhn, and Julian I Schroeder. Carbonic anhydrases are upstream regulators of co 2-controlled stomatal movements in guard cells. *Nature cell biology*, 12(1):87–93, 2010.

[124] Honghong Hu, Wouter-Jan Rappel, Rossana Occhipinti, Amber Ries, Maik Böhmer, Lei You, Chuanlei Xiao, Cawas B Engineer, Walter F Boron, and Julian I Schroeder. Distinct

cellular locations of carbonic anhydrases mediate carbon dioxide control of stomatal movements. *Plant physiology*, 169(2):1168–1178, 2015.

[125] Mimi Hashimoto, Juntaro Negi, Jared Young, Maria Israelsson, Julian I Schroeder, and Koh Iba. Arabidopsis ht1 kinase controls stomatal movements in response to co 2. *Nature cell biology*, 8(4):391–397, 2006.

[126] Asami Hiyama, Atsushi Takemiya, Shintaro Munemasa, Eiji Okuma, Naoyuki Sugiyama, Yasuomi Tada, Yoshiyuki Murata, and Ken-ichiro Shimazaki. Blue light and co 2 signals converge to regulate light-induced stomatal opening. *Nature communications*, 8(1):1–13, 2017.

[127] Jingbo Zhang, Paulo De-oliveira Ceciliato, Yohei Takahashi, Sebastian Schulze, Guillaume Dubeaux, Felix Hauser, Tamar Azoulay-Shemer, Kadri Tõldsepp, Hannes Kollist, Wouter-Jan Rappel, and Julian I Schroeder. Insights into the molecular mechanisms of CO2-mediated regulation of stomatal movements. *Current Biology*, 28(23):R1356–R1363, 2018.

[128] Paulo H O Ceciliato, Jingbo Zhang, Qing Liu, Xin Shen, Honghong Hu, Chen Liu, Anton R Schäffner, and Julian I Schroeder. Intact leaf gas exchange provides a robust method for measuring the kinetics of stomatal conductance responses to abscisic acid and other small molecules in Arabidopsis and grasses. *Plant methods*, 15(1):1–10, 2019.

[129] Amnon Schwartz, Nitza Ilan, and David A Grantz. Calcium effects on stomatal movement in commelina communis L.: use of EGTA to modulate stomatal response to light, KCl and CO2. *Plant Physiology*, 87(3):583–587, 1988.

[130] Alex A R Webb, Martin R McAinsh, Terry A Mansfield, and Alistair M Hetherington. Carbon dioxide induces increases in guard cell cytosolic free calcium. *The Plant Journal*, 9(3):297–304, 1996.

[131] Sebastian Schulze, Guillaume Dubeaux, Paulo H O Ceciliato, Shintaro Munemasa, Maris Nuhkat, Dmitry Yarmolinsky, Jaimee Aguilar, Renee Diaz, Tamar Azoulay-Shemer, Leonie Steinhorst, Jan Niklas Offenborn, Jorg Kudla, Hannes Kollist, and Julian I Schroeder. A role for calcium dependent protein kinases in differential CO2 and ABA controlled stomatal closing and low CO2 induced stomatal opening in Arabidopsis. *New Phytologist*, 229(5):2765–2779, 2021.

[132] Yana Qu, Ping Song, Yanwei Hu, Xin Jin, Qianru Jia, Xuedong Zhang, Long Chen, and Qun Zhang. Regulation of stomatal movement by cortical microtubule organization in response to darkness and aba signaling in arabidopsis. *Plant Growth Regulation*, 84(3):467–479, 2018.

[133] Yue Rui and Charles T Anderson. Functional analysis of cellulose and xyloglucan in the walls of stomatal guard cells of arabidopsis. *Plant Physiology*, 170(3):1398–1419, 2016.

[134] Julian D Schwab, Silke D Kühlwein, Nensi Ikonomi, Michael Kühl, and Hans A Kestler. Concepts in Boolean network modeling: What do they all mean? *Computational and structural biotechnology journal*, 18:571–582, 2020.

[135] Alessandro Di Cara, Abhishek Garg, Giovanni De Micheli, Ioannis Xenarios, and Luis Mendoza. Dynamic simulation of regulatory networks using squad. *BMC bioinformatics*, 8(1):1–10, 2007.

[136] Cristal Zuniga, Jennifer Levering, Maciek R Antoniewicz, Michael T Guarnieri, Michael J Betenbaugh, and Karsten Zengler. Predicting dynamic metabolic demands in the photosynthetic eukaryote chlorella vulgaris. *Plant physiology*, 176:450–462, 2018.

[137] Tomáš Helikar, Bryan Kowal, Sean McClenathan, Mitchell Bruckner, Thaine Rowley, Alex Madrahimov, Ben Wicks, Manish Shrestha, Kahani Limbu, and Jim A Rogers. The cell collective: toward an open and collaborative approach to systems biology. *BMC systems biology*, 6(1):1–14, 2012.

[138] Lutz Prechelt. An empirical comparison of seven programming languages. *Computer*, 33(10):23–29, 2000.

[139] Robert Hundt. Loop recognition in c++/java/go/scala. 2011.

[140] Shaowu Xue, Honghong Hu, Amber Ries, Ebe Merilo, Hannes Kollist, and Julian I Schroeder. Central functions of bicarbonate in s-type anion channel activation and ost1 protein kinase in co2 signal transduction in guard cell. *The EMBO journal*, 30(8):1645–1658, 2011.

[141] Ebe Merilo, Kristiina Laanemets, Honghong Hu, Shaowu Xue, Liina Jakobson, Ingmar Tulva, Miguel Gonzalez-Guzman, Pedro L Rodriguez, Julian I Schroeder, Mikael Brosche, and Hannes Kollist. Pyr/rcar receptors contribute to ozone-, reduced air humidity-, darkness-, and co2-induced stomatal regulation. *Plant Physiology*, 162(3):1652–1668, 2013.

[142] Deping Hua, Cun Wang, Junna He, Hui Liao, Ying Duan, Ziqiang Zhu, Yan Guo, Zhizhong Chen, and Zhizhong Gong. A plasma membrane receptor kinase, ghr1, mediates abscisic acid-and hydrogen peroxide-regulated stomatal movement in arabidopsis. *The Plant Cell*, 24(6):2546–2561, 2012.

[143] Maija Sierla, Hanna Hõrak, Kirk Overmyer, Cezary Waszczak, Dmitry Yarmolinsky, Tobias Maierhofer, Julia P Vainonen, Jarkko Salojärvi, Konstantin Denessiouk, Kristiina Laanemets, Kadri Toldsepp, Triin Vahisalu, Adrien Gauthier, Tuomas Puuko, Lars Paulin, Petri Auvinen, Dietmar Geiger, Rainer Hedrich, Hannes Kollist, and Jaakko Kangasjarvi. The receptor-like pseudokinase ghr1 is required for stomatal closure. *The Plant Cell*, 30(11):2813–2837, 2018.

[144] Dean R Dubbe, Graham D Farquhar, and Klaus Raschke. Effect of abscisic acid on the gain of the feedback loop involving carbon dioxide and stomata. *Plant Physiology*, 62(3):413–417, 1978.

[145] Klaus Raschke. Simultaneous requirement of carbon dioxide and abscisic acid for stomatal closing in xanthium strumarium l. *Planta*, 125(3):243–259, 1975.

[146] Marco Baggiolini. Chemokines and leukocyte traffic. *Nature*, 392(6676):565, 1998.

[147] A. J. Ridley, M. A. Schwartz, K. Burridge, R. A. Firtel, M. H. Ginsberg, G. Borisy, J. T. Parsons, and A. R. Horwitz. Cell migration: integrating signals from front to back. *Science*, 302(5651):1704–1709, 2003.

[148] Denise J Montell. Border-cell migration: the race is on. *Nature Reviews Molecular Cell Biology*, 4(1):13–24, 2003.

[149] Herbert Levine and Wouter-Jan Rappel. The physics of eukaryotic chemotaxis. *Phys Today*, 66(2), Feb 2013.

[150] Herbert Levine, Igor Aranson, Lev Tsimring, and Thai Viet Truong. Positive genetic feedback governs cAMP spiral wave formation in dictyostelium. *Proceedings of the National Academy of Sciences*, 93(13):6382–6386, 1996.

[151] Giovanna De Palo, Darvin Yi, and Robert G Endres. A critical-like collective state leads to long-range cell communication in Dictyostelium discoideum aggregation. *PLoS biology*, 15(4), 2017.

[152] Danying Shao, Wouter-Jan Rappel, and Herbert Levine. Computational model for cell morphodynamics. *Physical Review Letters*, 105(10), 2010.

[153] D. Shao, H. Levine, and W. J. Rappel. Coupling actin flow, adhesion, and morphology in a computational cell motility model. *Proc Natl Acad Sci U S A*, Apr 2012.

[154] Wieland Marth and Axel Voigt. Signaling networks and cell motility: a computational approach using a phase field description. *J Math Biol*, 69(1):91–112, Jul 2014.

[155] Carole A Parent and Peter N Devreotes. A cell's sense of direction. *Science*, 284(5415):765–770, 1999.

[156] A. Levchenko and P. A. Iglesias. Models of eukaryotic gradient sensing: application to chemotaxis of amoebae and neutrophils. *Biophys J*, 82:50–63, 2002.

[157] W. J. Rappel and H. Levine. Receptor noise limitations on chemotactic sensing. *PNAS*, 105(49):19270–19275, 2008.

[158] P. A. Iglesias and P. N. Devreotes. Navigating through models of chemotaxis. *Curr Opin Cell Biol*, 20(1):35–40, 2008.

[159] W. J. Rappel and H. Levine. Receptor noise and directional sensing in eukaryotic chemotaxis. *Phys Rev Lett*, 100(22):228101, 2008.

[160] Jose Negrete Jr, Alain Pumir, Hsin-Fang Hsu, Christian Westendorf, Marco Tarantola, Carsten Beta, and Eberhard Bodenschatz. Noisy oscillations in the actin cytoskeleton of chemotactic amoeba. *Physical review letters*, 117(14):148102, 2016.

[161] Hsin-Fang Hsu, Eberhard Bodenschatz, Christian Westendorf, Azam Gholami, Alain Pumir, Marco Tarantola, and Carsten Beta. Variability and order in cytoskeletal dynamics of motile amoeboid cells. *Physical review letters*, 119(14):148101, 2017.

[162] Akihiko Nakajima, Shuji Ishihara, Daisuke Imoto, and Satoshi Sawai. Rectified directional sensing in long-range cell migration. *Nature communications*, 5, 2014.

[163] Paul Herzmark, Kyle Campbell, Fei Wang, Kit Wong, Hana El-Samad, Alex Groisman, and Henry R. Bourne. Bound attractant at the leading vs. the trailing edge determines chemotactic prowess. *Proceedings of the National Academy of Sciences*, 104(33):13349–13354, 2007.

[164] Yuhai Tu and Wouter-Jan Rappel. Adaptation in living systems. *Annual review of condensed matter physics*, 9:183–205, 2018.

[165] Peter N Devreotes, Michael J Potel, and Stephen A MacKay. Quantitative analysis of cyclic AMP waves mediating aggregation in dictyostelium discoideum. *Developmental biology*, 96(2):405–415, 1983.

[166] Marten Postma and Peter JM van Haastert. Mathematics of experimentally generated chemoattractant gradients. In *Chemotaxis*, pages 473–488. Springer, 2009.

[167] D. Fuller, W. Chen, M. Adler, A. Groisman, H. Levine, W. J. Rappel, and W. F. Loomis. External and internal constraints on eukaryotic chemotaxis. *PNAS*, 107(21):9656–9659, 2010.

[168] Maurice Sussman. Cultivation and synchronous morphogenesis of Dictyostelium under controlled experimental conditions. In *Methods in cell biology*, volume 28, pages 9–29. Elsevier, 1987.

[169] L. Bosgraaf and P. J. Van Haastert. Navigation of chemotactic cells by parallel signaling to pseudopod persistence and orientation. *PLoS ONE*, 4(8):e6842, 2009.

[170] Monica Skoge, Micha Adler, Alex Groisman, Herbert Levine, William F Loomis, and Wouter-Jan Rappel. Gradient sensing in defined chemotactic fields. *Integrative Biology*, 2(11-12):659–668, 2010.

[171] Arpan Bhowmik, Wouter-Jan Rappel, and Herbert Levine. Excitable waves and direction-sensing in dictyostelium discoideum: steps towards a chemotaxis model. *Physical biology*, 13(1):016002, 2016.

[172] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.

[173] P. N. Devreotes. *Dictyostelium discoideum*: a model system for cell-cell interactions in development. *Science*, 245:1054–1058, 1989.

[174] W Roos and G Gerisch. Receptor-mediated adenylate cyclase activation in dictyostelium discoideum. *FEBS letters*, 68(2):170–172, 1976.

[175] Richard H. Kessin. *Dictyostelium: The Evolution, Cell Biology, and Development of a Social Organism*. Cambridge University Press, Cambridge, 2001.

[176] William F Loomis. Cell signaling during development of dictyostelium. *Developmental biology*, 391(1):1–16, 2014.

[177] Philippe V Afonso, Mirkka Janka-Junttila, Young Jong Lee, Colin P McCann, Charlotte M Oliver, Khaled A Aamer, Wolfgang Losert, Marcus T Cicerone, and Carole A Parent. Ltb4 is a signal-relay molecule during neutrophil chemotaxis. *Developmental cell*, 22(5):1079–1091, 2012.

[178] Tim Lämmermann, Philippe V Afonso, Bastian R Angermann, Ji Ming Wang, Wolfgang Kastenmüller, Carole A Parent, and Ronald N Germain. Neutrophil swarms require ltb4 and integrins at sites of cell death in vivo. *Nature*, 498(7454):371–375, 2013.

[179] Kyoung J Lee, Edward C Cox, and Raymond E Goldstein. Competing patterns of signaling activity in dictyostelium discoideum. *Physical review letters*, 76(7):1174, 1996.

[180] John J Tyson, Kevin A Alexander, VS Manoranjan, and JD Murray. Spiral waves of cyclic amp in a model of slime mold aggregation. *Physica D: Nonlinear Phenomena*, 34(1-2):193–207, 1989.

[181] Thomas Höfer, Jonathan A Sherratt, and Philip K Maini. Cellular pattern formation during Dictyostelium aggregation. *Physica D: Nonlinear Phenomena*, 85(3):425–444, 1995.

[182] Javad Noorbakhsh, David J Schwab, Allyson E Sgro, Thomas Gregor, and Pankaj Mehta. Modeling oscillations and spiral waves in Dictyostelium populations. *Physical Review E*, 91(6):062711, 2015.

[183] P. J. Van Haastert and P. N. Devreotes. Chemotaxis: signalling the way forward. *Nat Rev Mol Cell Biol*, 5(8):626–634, 2004.

[184] K. F. Swaney, C. H. Huang, and P. N. Devreotes. Eukaryotic chemotaxis: a network of signaling pathways controls motility, directional sensing, and polarity. *Annu Rev Biophys*, 39:265–289, 2010.

[185] S. K. Dertinger, D. T. Chiu, S. K. Dertinger, and G. M. Whitesides. Generation of gradients having complex shapes using microfluidic networks. *Anal Chem*, 73:1240–1246, 2001.

[186] Carsten Beta and Eberhard Bodenschatz. Microfluidic tools for quantitative studies of eukaryotic chemotaxis. *European journal of cell biology*, 90(10):811–816, 2011.

[187] Mohammad A Qasaimeh, Michal Pyzik, Mélina Astolfi, Silvia M Vidal, and David Juncker. Neutrophil chemotaxis in moving gradients. *Advanced Biosystems*, 2(7):1700243, 2018.

[188] Victor Sourjik and Ned S Wingreen. Responding to chemical gradients: bacterial chemotaxis. *Current opinion in cell biology*, 24(2):262–268, 2012.

[189] L. Song, S. M. Nadkarni, H. U. Bödeker, C. Beta, A. Bae, C. Franck, W.-J. Rappel, W. F. Loomis, and E. Bodenschatz. *Dictyostelium discoideum* chemotaxis: threshold for directed motion. *Eur J Cell Biol*, 85:981–989, 2006.

[190] P. J. van Haastert and M. Postma. Biased random walk by stochastic fluctuations of chemoattractant-receptor interactions at the lower limit of detection. *Biophys J*, 93(5):1787–1796, 2007.

[191] Peter N Devreotes, Sayak Bhattacharya, Marc Edwards, Pablo A Iglesias, Thomas Lampert, and Yuchuan Miao. Excitable signal transduction networks in directed cell migration. *Annual review of cell and developmental biology*, 33:103–125, 2017.

[192] Xiaoguang Li, Yuchuan Miao, Dhiman Sankar Pal, and Peter N Devreotes. Excitable networks controlling cell migration during development and disease. In *Seminars in cell & developmental biology*, volume 100, pages 133–142. Elsevier, 2020.

[193] Richa Karmakar, Timothy Tyree, Richard H Gomer, and Wouter-Jan Rappel. Cell dispersal by localized degradation of a chemoattractant. *Proceedings of the National Academy of Sciences*, 118(6), 2021.

[194] MATLAB. *MATLAB version: 9.14.0 (R2023a)*. The MathWorks Inc., Natick, Massachusetts, United States, 2023.

[195] Aravind Karanam and Wouter-Jan Rappel. Boolean modelling in plant biology. *Quantitative Plant Biology*, 3:e29, 2022.