Engineering genetic circuits in receiver cells for diffusion-based molecular data communications

Merve Gorkem Durmaz,[†] Neval Tulluk,[†] Recep Deniz Aksoy,[†] H. Birkan Yilmaz,[†] Bill Yang,[‡] Anil Wipat,[‡] Ali Emre Pusane,[¶] Göksel Mısırlı,[§] and Tuna Tugcu[*,†]

†*Department of Computer Engineering, NETLAB, Bogazici University, Istanbul, Turkey*

‡*School of Computing, Newcastle University, UK*

¶*Department of Electrical and Electronics Engineering, Bogazici University, Istanbul, Turkey*

§*School of Computer Science and Mathematics, Keele University, Staffordshire, UK*

E-mail: g.misirli@keele.ac.uk,tugcu@boun.edu.tr

Abstract

Developments in bioengineering and nanotechnology have ignited the research on biological and molecular communication systems. Despite potential benefits, engineering communication systems to carry data signals using biological messenger molecules is challenging. Diffusing molecules may fall behind their schedule to arrive at a receiver, interfering with symbols of subsequent time slots and distorting the signal. Theoretical molecular communication models often focus solely on the characteristics of the communication channel and fail to provide an end-to-end system response, since they assume a simple thresholding process for a receiver cell and overlook how the receiver can detect the incoming distorted molecular signal. There is a need to develop viable end-to-end communication models. In this paper, we present a model-based framework for designing diffusion-based molecular communication systems coupled with synthetic

genetic circuits. We describe a novel approach to encode information as a sequence of bits, each transmitted from a sender as a burst of specific number of molecules, control cellular behavior, and minimize cellular signal interference by employing equalization techniques from communication theory. This approach allows the encoding and decoding of data bits efficiently using two different types of molecules that act as the data carrier and the antagonist to cancel out the heavy tail of the former. We also present Period Finder, as a tool to optimize communication parameters, including the number of molecules and symbol duration. This tool facilitates automating the choice of communication parameters and identifying the best communication scenarios that can produce efficient cellular signals.

Keywords

Synthetic biology, molecular communications, genetic circuits, model-driven design, reciever design

1 Introduction

Cells communicate with the environment and each other to maintain life (*1*). Examples include single-cell organisms, such as bacteria that are organized in micro societies and neurons in the human brain. Naturally, biological and signaling molecules are used as information carriers. The communication mechanisms between sender and receiver cells are of importance for the communications engineers to engineer multicellular applications. Synthetic genetic circuits (*2*, *3*) can offer a rewarding tool for molecular communication systems to go beyond the diffusion of biochemical molecules to modulate and carry information between sender and receiver cells by providing the capability to understand and engineer complex multicellular applications. Integrating approaches from molecular communications and synthetic biology and considering intracellular and intercellular dynamics of signaling molecules can

help achieve communication-based and robust biological applications. However, the application of such integrative approaches to develop biological communication systems is limited to date.

Molecular communication systems can become challenging to engineer when living cells are involved. Signaling molecules propagating in the intercellular medium may trigger adverse cellular responses (*4*). One way to facilitate the design of predictable biological molecular communication channels is to apply model-based design approaches, which may involve integrating multi-scale mathematical models (*5*, *6*). This process can be complemented by a holistic approach to simulate the dynamics of, and cellular response to, messenger molecules that carry information for signaling and diffuse across a communication channel.

Incorporating the dynamics of underlying transmission channels has several advantages in developing nano- and micro-scale biological communication systems. Inspired by nature, molecular communication systems are bio-compatible and less invasive than their alternatives, such as utilizing electromagnetic radiation (*7*). These systems are energy efficient since the utilized biochemical reactions consume low energy. In addition, the transmitted molecules propagate freely, and molecular communication via diffusion (MCvD) has no external energy requirements (*8*). A well-known model of a molecular communication channel developed by Shannon and Weaver (*9*) consists of five key elements: an information source that generates the message, a sender or transmitter (Tx) that encodes the message into a communication signal, a channel in which the signal propagates, a receiver (Rx) that decodes or translates the received signal, and a destination node.

Developing signaling channels and encoding information can be affected by the inherent noise and stochastic behavior of diffusing molecules in a crowded medium, such as the intracellular environment full of water, ions, and various macromolecules (*10*). Noise is the undesired element of a communication channel and can be defined as any inference that changes the received signal, typically in a destructive way (*9*). Diffusing molecules may not arrive at the receiver cells on their predestined time slots and may interfere with the subse-

quent data transmissions (*11*). As a result, when molecules released from a sender represent a particular symbol, the receiver may decode this symbol incorrectly. This situation causes significant intersymbol interference (ISI), which is considered one of the major challenges in diffusion-based communication systems that hinders communication at high data rates (*12*, *13*).

Different theoretical approaches have been proposed to handle ISI (*14–16*). Noel and co-workers (*17*) proposed adding enzymes into the propagation channel. Enzymes degrade information molecules and prevent stray molecules from interfering with future transmissions. Tepekule and co-workers (*18*) proposed a molecular transition shift keying technique in which the presence of two different types of molecules (type-A and type-B) is used to encode bit-1. The absence of these molecules represents bit-0. The choice of molecule type to encode bit-1 depends on the following bit-0 or bit-1 symbol. If the next symbol is bit-0, type-B molecules are sent. If the next symbol is bit-1, type-A molecules are sent. This strategy ensures that only type-B molecules are released before bit-0, and the accumulation of molecules and ISI are restrained (*18*). Another proposed solution is the pre-equalization method (*19*), which involves transmitting two different molecule types from a sender: type-A information encoding molecules and type-B destructive molecules. In this approach, type-B molecules eliminate the effect of the stray type-A molecules. The impact of the destructive molecule is imitated by employing a subtraction operation at the receiver. However, the aforementioned methods do not address the issue of how these molecules are converted to intracellular signals that affect the system's response and noise when genetic receivers are employed. This study incorporates and improves the pre-equalization method that has the reduced complexity compared to other methods to address the effects of intercellular and intracellular signaling processes.

Different modulation techniques have been proposed to encode information using molecular communication systems. Some of these techniques are based on the concentration of the transmitted molecules from a Tx sender (*20*), the type of the transmitted molecules

(*21*), the release time of the transmitted molecules within a communication time slot (*22*), and the locations of spatially-separated multiple transmitters (*23*). There are also hybrid techniques that combine more than one aspect of the information carriers (*24*, *25*).

The movement of molecules in a molecular communication channel can be modeled by the diffusion process or the Brownian motion. In a fluidic environment without any flow, molecules move randomly (*26*). When diffusing molecules reach receiver cells, they may activate some processes or yield information bits after a demodulation process. Therefore, evaluating the expected number of received molecules is critical for designing an effective MCvD system that involves receiver cells. Channel properties, absorption rate, and biochemical kinetic rates may affect the activation of a synthetic genetic circuit inside an Rx receiver. The activation of the circuit may also depend on the specific level of intercellular signaling molecules for a certain amount of time. It is desirable to incorporate models of genetic circuits to design efficient molecular communication systems and understand the effect of ISI on cellular response.

Various modeling formalisms exist to analyze the dynamic behavior of genetic circuits. For example, the Systems Biology Markup Language (SBML) (*27*) standardizes the representation of biochemical reactions. Tools such as COPASI (*28*) can simulate SBML models deterministically or stochastically to gain insight into emerging cellular behavior. Such modeling formalisms have already been used to demonstrate the model-driven design of synthetic genetic circuits. The Virtual Parts Repository (VPR) (*29*, *30*) provides reusable, modular, and mathematical models of biological components such as promoters, ribosomal binding sites (RBSs), and coding sequences (CDSs). These models can be joined together to create models of desired systems. The resulting models are hierarchical, based on models of parts and interactions and their aggregation. This model-driven approach is ideal for designing and optimizing genetic circuits and computationally exploring large design spaces of desirable biological systems.

Standardization efforts are essential to exchange information between tools without loss

of data. Synthetic genetic regulatory circuits can be computationally represented using the Synthetic Biology Open Language (SBOL) (*31*, *32*). SBOL designs can include constraints to capture the sensing of external molecules and descriptions of intended biochemical reactions. This qualitative information can consequently be used to create models that can be simulated. The recent version of VPR provides SBOL-to-SBML conversion to automate the generation of computational models of genetic circuits (*30*, *33*).

Here, we present a computational modeling approach to facilitate the design of molecular communication systems that can be coupled with living cells. This approach is workflow based and integrates the modeling efforts in molecular communications and synthetic biology. Computational simulations are used to design different communication aspects to encode and send information using biological molecules. Our modeling framework aims to minimize signal interference and allows computationally optimizing key communication parameters, such as symbol duration, to produce efficient cellular signals.

2 Results

The computational modeling approach presented here was developed to design communication systems using molecular and biological communication channels. This process involves coupling intracellular and extracellular processes with diffusion dynamics and three-dimensional molecular channel propagation (*34*). The information can be encoded as sequential bits, each representing a number of molecules released by a sender. As a result, a response signal is created at a receiver via the accumulation of cellular molecules. We demonstrate this approach computationally using a receiver design based on synthetic and bacterial genetic regulatory networks to decode information. The coupling of intercellular and intracellular mechanisms is implemented as a workflow in which an MCvD system with pre-equalizer (*19*) can eliminate ISI. The workflow incorporates VPR to facilitate model-driven design and optimization of genetic receivers.

A Pre-equalizer for Engineered Receiver Cells. ISI mitigation is especially challenging when living cells act as receivers due to different timescales in intercellular diffusion dynamics and intracellular biochemical reactions. Different biochemical rate parameters may cause received signals to interfere with the subsequent cellular signals. In this work, we propose a cellular pre-equalizer method building upon the literature (*11*). This method involves two input signals emitted at a sender and two additional cellular signals at a receiver.

The input signals together carry a single bit of data to reduce interference. The first signal (type-A) is the actual data carrier, and each bit corresponds to type-A molecules being sent over a specific period. The second input signal (type-B) removes the heavy tail of the type-A signal. The effectiveness of the removal of the interference is calculated by the receiver cells. Hence, A_{out} and B_{out} signals are transformed into intracellular molecules A_{in} and B_{in}, respectively. A_{in} is the observable molecule that relays the signal, and B_{in} is the antagonist of A_{in} to cancel out the right amount of A_{in} to mitigate the adverse effects of ISI.

It is crucial to comply with the processing rates of Rx receiver cells to send sequential information or data bits. This process requires maintaining a specific level of messenger molecule concentration. Moreover, to eliminate the heavy tail of A_{in} at an Rx receiver, B_{out} is emitted t_{shift} seconds after A_{out} is released at a Tx sender. With the appropriate communication parameters, the deteriorating effects of the ISI molecules can be eliminated by considering the difference between the two cellular molecule types ($A_{in} - B_{in}$). Hence, communication performance can be improved significantly. However, determining the parameters of the pre-equalizer, exploring potential communication scenarios, and measuring the performance of each scenario's end-to-end communication performance is not trivial. This design space exploration process is carried out by Period Finder.

Communication Period Finder. A tool called Period Finder has been developed to implement the proposed pre-equalizer method and efficiently design communication channels. Period Finder evaluates the communication performance and optimizes pre-equalizer's parameters to minimize the degrading effects of ISI. These processes are achieved using

computational simulations.

Period Finder generates possible signal propagation scenarios with different signal parameters with an effective A_{out}/B_{out} ratio to mitigate ISI. The resulting scenarios are evaluated to ensure that each bit persists for the desired period or symbol duration. The signal parameters of these scenarios generated by Period Finder (Figure 2) are assessed through performance metrics involving MOL-eye diagrams (*35*). MOL-eye is similar to the 'eye' diagram that is used for measuring the quality of signals in conventional communication schemes and is adapted to molecular communications. As a result, each communication scenario is scored, and the most effective scenario with the highest score in terms of A_{out}/B_{out} is selected. This approach allows adjusting the amplitude and symbol duration ratio of the actual data carrier signal to that of the antagonist signal. Consequently, ISI is minimized.

Table 1: Simulation parameters and values

Parameter	Value
Diameter of the receiver	2 µm (Ref. *36*)
Total number of molecules	3,500,000 (Ref. *37*)
Distance between Tx and Rx (d)	200 µm (Ref. *7*)
Diffusion coefficients for IPTG and aTc	$\{600, 870\}$ µm^2/s (Ref. *38*)
Signal shift between A_{out} and B_{out} (t_{shift})	900 s
Maximum symbol duration	4000 s

The flowchart of the program is shown in Figure 2. Period Finder retrieves the distinct periods (t_s values) that take the system to return to equilibrium, which is the base state of the cell before any symbol is received. A vector of different A_{out}/B_{out} values coupled with a vector of t_{shift} values, the desired maximum t_s to limit the symbol duration retrieved, and the total number of molecules sent per symbol are given as input. The initial state of A_{in}/B_{in} in a receiver cell (R_x) without any inputs is shown in Figure 3a. After the warm-up period (first two bits), the system reaches an equilibrium state where A_{in} and B_{in} decrease until they converge on a steady state. Period Finder then creates diverse scenarios with the vectors of A_{out}/B_{out} values coupled with a vector of t_{shift} values to simulate "00100000000000", a one-shot signal where bit-0 represents nothing is sent, and bit-1 represents that A_{out} and

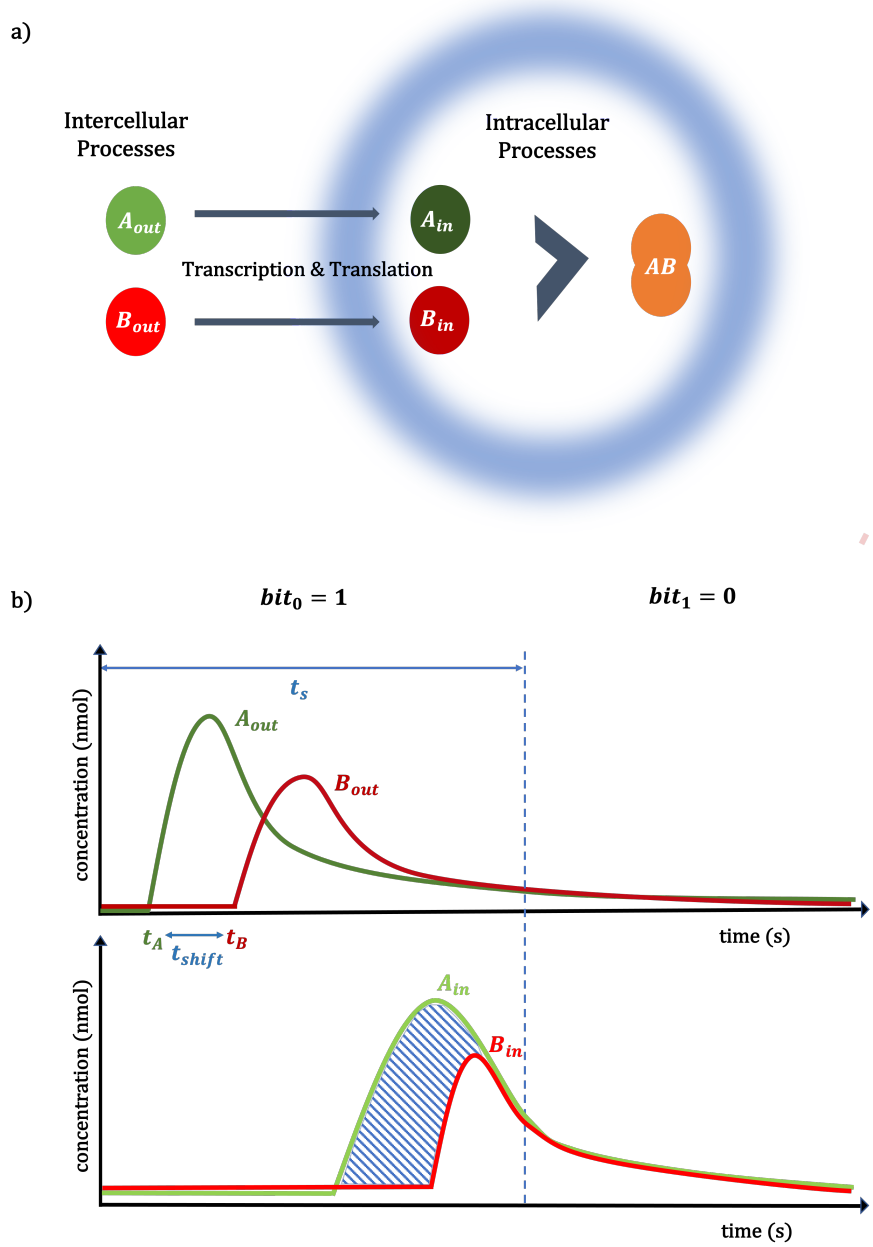


Figure 1: (a) The schematic representation of the general framework for Period Finder, where intercellular signals are converted to intracellular signals. (b) A hypothetical illustration of the signals, where bit-0 represents no transmission and bit-1 represents the transmission of the molecules. In the upper graph, t_A represents the transmission time of A_{out}, and t_B represents the transmission time of B_{out}. t_B-t_A is t_{shift}, and t_s denotes the symbol duration. A_{out} molecules are transmitted into intercellular space. This signal is converted into intracellular A_{in} signal with a delay due to cellular processes such as transcription and translation. B_{in} molecules mitigate the heavy tail of the A_{in} signal. A_{in}-B_{in} is denoted by the shaded region.

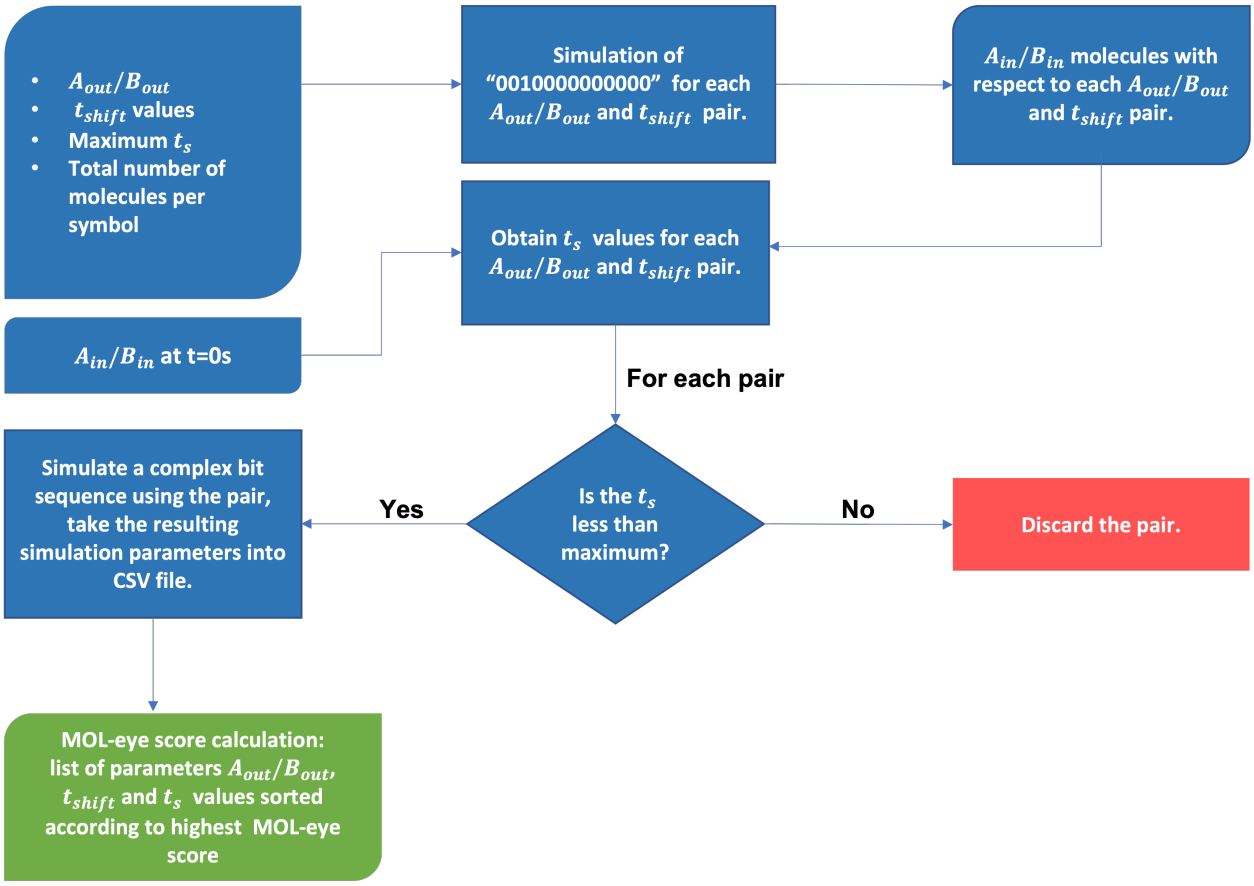


Figure 2: Flowchart of Period Finder. The input of the program is a CSV file containing different A_{out}/B_{out} ratios at different t_{shift} values, a desired maximum t_s value, total number of molecules per symbol, and A_{in}/B_{in} before any transmission (t=0). For each A_{out}/B_{out} and t_{shift} pair, the "0010000000000" bit sequence is simulated, and a ratio of A_{in}/B_{in} is obtained. Values of t_s are obtained for each resulting A_{in}/B_{in}. If the resulting t_s values are larger than the maximum t_s value, they are discarded. Otherwise, a complex bit sequence is simulated using the pair, and the resulting parameters are written into a new CSV file to be evaluated by MOL-eye calculation. After the analysis, A_{out}/B_{out} and t_{shift} pair and t_s values are sorted according to the highest MOL-eye score.

B_{out} molecules are sent. In the one-shot signal, bit-1 is sent only once in the sequence shown in Figure 3b. After A_{in} and B_{in} reach their maximum levels in the simulation, the program obtains the corresponding symbol duration to return to equilibrium. Here, we send enough bits to observe the system's behavior and simulate the scenario in a limited time. If the achieved t_s value is larger than the desired maximum t_s value, then it is discarded. Waiting for the equilibrium state to be achieved ensures the system is on hold until the channel is

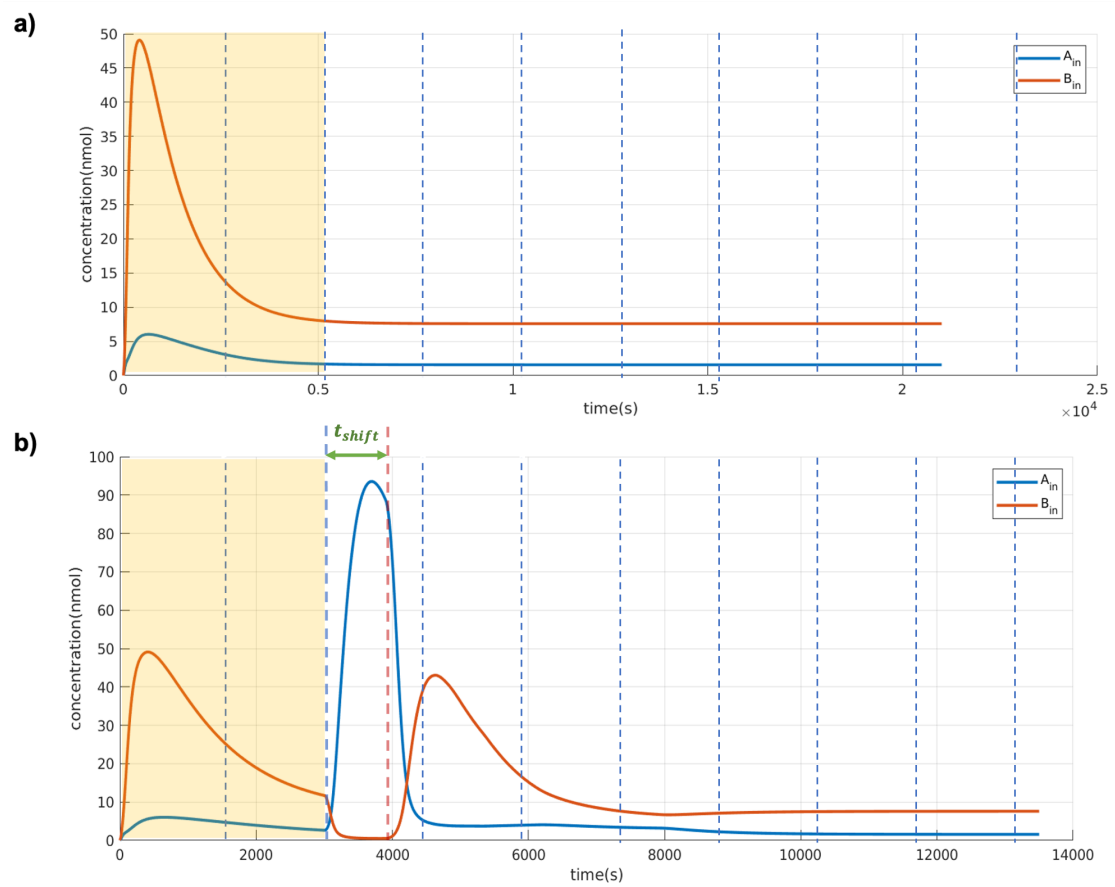


Figure 3: Required inputs for Period Finder (a) Native state of Rx, (b) Rx response for a one-shot signal of both A_{out} and B_{out}. There is a t_{shift} time between the transmission of A_{out} and B_{out} to keep the A_{in} signal as large as possible in a pre-defined t_s. Simulation parameters are given in Table 1. The first two t_s symbol durations are marked with yellow to indicate the warm-up period.

cleared out to send the subsequent symbol, decreasing the effect of ISI.

Simulation results are stored in a CSV table that includes the A_{in}/B_{in} values of one-shot signals for every t_{shift} and A_{out}/B_{out} ratio pair and their corresponding t_s values. The simulation results are evaluated according to MOL-eye diagram performance metrics, and the scenarios with the highest MOL-eye score is reported.

Minimizing Interference in MCvD. A biological subtraction operation can minimize cellular interface when two signals are used. Here, this operation has been implemented using a design pattern involving two molecules that can bind together (*39*). Inside a receiver

cell, A_{out} activates the production of A_{in}, and B_{out} activates the production of B_{in}. The difference between A_{in} and B_{in} molecules is evaluated via a biological subtraction operation. In this design, B_{in} sequesters A_{in}. The remaining A_{in} molecules represent the result of the subtraction operation. Conversion of an intercellular signal to an intracellular signal can be achieved in multiple ways. This process involves designing and incorporating genetic circuits. Here, we adopt transcriptional activation and inhibition processes to control the production of proteins corresponding to different signals. Interactions of these proteins form the basis of our biological program.

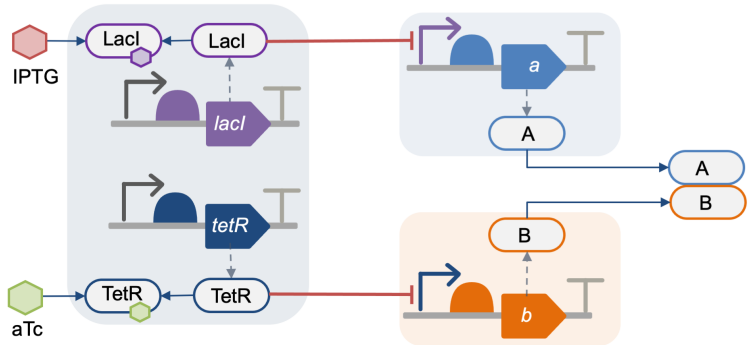


Figure 4: Representation of the system model used in simulations to perform the biological subtraction operation. Remaining A_{in} molecules over a defined period represent a data bit.

In this design, IPTG and aTc diffusing molecules represent the A_{out} and B_{out} signals (Figure 4). LacI and TetR repressors inhibit the production of A_{in} and B_{in}, respectively. Hence, IPTG activates the production of A_{in} by inhibiting LacI, and aTc activates the production of B_{in} by inhibiting TetR. Molecules, such as ExsD and ExsA, that can bind together represent the A_{in} and B_{in} molecules (*40*, *41*). Here, ExsD and ExsA were chosen since they can act as TFs to control subsequent cellular response.

Period Finder relies on VPR to create models of genetic circuits that receive the external A_{out} and B_{out} signals and convert them into cellular A_{in} and B_{in} signals (Figure 1a). Models of genetic parts such as promoters, RBSs, CDSs, and their interactions are joined together to derive SBML models that can be simulated. Each resulting SBML model represents a possible solution and is evaluated by Period Finder. Models include rate parameters for various

biological processes involving genetic production, complex formation, and binding. Period Finder uses the COPASI engine to automate the simulation of each model and determines A_{in} and B_{in} values according to the simulation results.

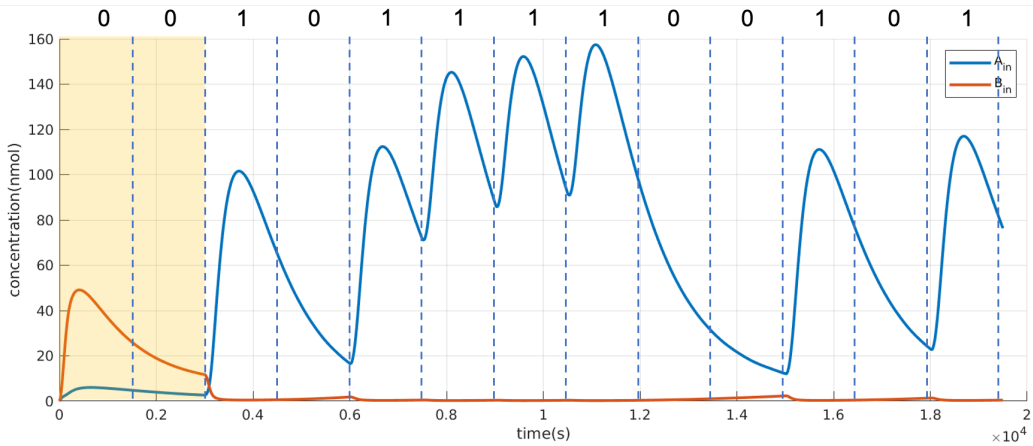


Figure 5: Rx response without using the B_{out} pre-equalizer. The bit sequence of the signal (0010111100101) is shown in the upper part of the graph. The first two t_s symbol duration slots are marked with yellow color to indicate the warm-up period, the y-axis shows the concentration (nmol) of molecules A_{in} (blue line) and B_{in} (red line), and the x-axis shows the time in seconds. After a train of consecutive bit-1 symbols, the signaling molecules accumulate in the channel. The concentration of signaling molecules at the 9th symbol of bit-0 is misleadingly high due to the effect of ISI, which is likely to cause an incorrect detection at the receiver side.

Simulation results when the B_{out} pre-equalizer is not incorporated are shown in Figure 5. The bit sequence for this simulation is 0010111100101, and the simulation parameters from Table 1 are applied. The first two symbol duration slots represent the warm-up period (*42*) for which the system is unstable. Hence, the simulation starts after this warm-up period. The first two bits are always chosen to be 00 to provide enough time for the system's stabilization. The lack of the pre-equalizer causes ISI and the accumulation of stray molecules. As shown on the graph, after the 5th symbol, consecutive bit-1 symbols cause the accumulation of molecules. Consequently, the concentration of signal molecules at the 9th symbol of bit-0 is misleadingly higher than expected.

Period Finder applies the pre-equalizer approach and adjusts the ratio of the amplitudes

of A_{out} and B_{out} signals and their symbol duration to decrease ISI. As a result, B_{in} molecules bind and reduce the number of stray A_{in} molecules remaining after the symbol duration.

Period Finder evaluates the resulting scenarios according to MOL-eye performance metrics and reports the optimum communication scenarios with the highest MOL-eye scores. An example, MOL-eye diagram of the optimum scenario is shown in Figure 6 where the subsequent signals are superimposed to a single composite graph and the area between the minimum of bit-1 symbol and maximum of bit-0 symbol is denoted as eye-height. The noise can be expected to be the highest when eye is in the most closed form, where it becomes challenging to distinguish bit-1 and bit-0 symbols. However, considering the optimum communication scenarios identified by Period Finder, even when the minimum of bit-1 and maximum of bit-0 is used, the eye-opening is still clear as shown in Figure 6b.

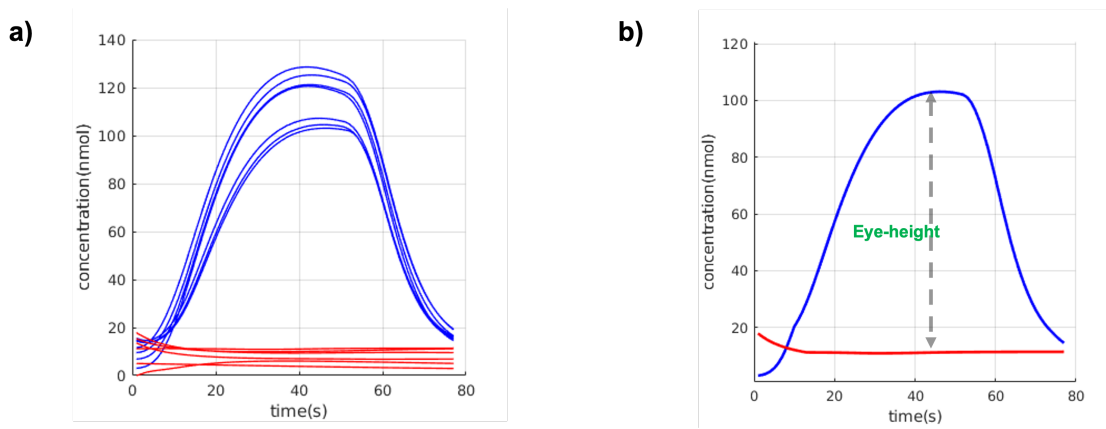


Figure 6: MOL-eye diagram, where successive bit-1 and bit-0s are superimposed, red lines represent bit-0 and blue lines represent bit-1. (a) Overlapped representation of bit-1 and bit-0 symbols (b) The eye-height is the vertical opening between the maximum of bit-0 and minimum of bit-1.

The same communication scenario is shown in Figure 7, including simulation results for both A_{in} and B_{in} signals in more detail. As expected, when the 0010111100101 bit sequence is sent, consecutive bit-1 symbols in the middle do not cause the accumulation of the signaling molecules anymore. Moreover, each bit-1 and bit-0 information is exchanged clearly, demonstrating that the effects of ISI are drastically reduced.

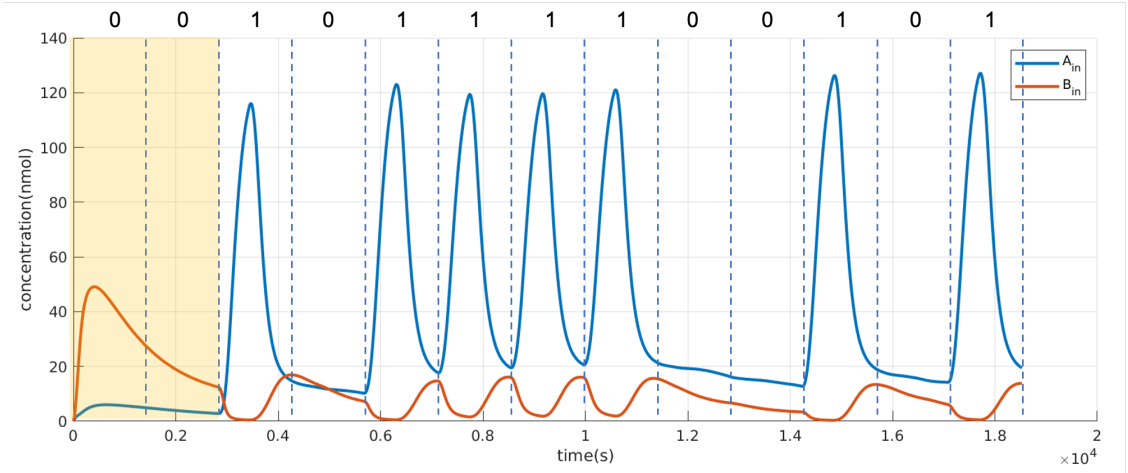


Figure 7: Rx response using the B_{out} pre-equalizer. A_{out}/B_{out} is 0.15. The first two t_s symbol durations are marked with yellow to indicate the warm-up period. The solid blue and red lines show the concentration of A_{in} and B_{in} molecules over time, respectively. As desired, signaling molecules do not start accumulating during the 5th symbol duration. Furthermore, the concentration of signaling molecules at the 9th symbol of bit-0 is as expected since the pre-equalizer (B_{in}) mitigates the effect of ISI.

Discussion. Molecular communication systems have several advantages in developing biological applications. However, it is challenging to develop such applications, which involve encoding and decoding information due to the noisy and uncertain environment of a medium. When these communication systems are coupled with cells that act as receivers, this complexity increases even further due to the large number of genetic parts that can be chosen to decode information and control cellular response. Trial and error-based wet-lab experiments can be costly. Moreover, wet-lab experimentation is out of reach for most researchers in the communication engineering area, since it requires qualified laboratories, equipment, and specialized staff (*43*). Simulation environments can provide valuable insights to overcome these issues and develop and test novel communication models. However, this process is not trivial. Diffusion-based cellular processes can be complex and may have different timescales. Resulting computational models may provide insights with varying details for different processes. The work presented here involves algorithms, design patterns, and a simulation approach to overcome the obstacles in engineering receiver cells that function via

molecular communications and diffusion of molecules to encode and send information.

One of the main challenges in using engineered receiver cells and diffusion-based systems is decoding the information due to inherent noise. Our work extends the previously proposed pre-equalizer approach (*18*), which involves using two different signals to distinguish data bits, by incorporating two additional cellular signals. A biological subtraction operation for these cellular signals has been defined as a genetic circuit design to improve the molecular channel response, reduce cellular noise and control cellular response.

Period Finder can propose successful communication scenarios based on the transmission time difference between the signaling molecules and their concentration ratio. These scenarios are ranked using the MOL-eye performance metrics. Hence, Period Finder can be ideal to automate exploring different communication parameters via computational simulations.

Nature-inspired communication systems can have several advantages for developing biological applications. On the other hand, it may be challenging to implement expectations of generic communication systems. Notably, a communication scenario can be slow. Here, we explored minimizing symbol interface by considering symbol duration, which can be minutes due to the diffusion of molecules and accumulation of cellular molecules via transcription and translation events (*44*). As a result, a communication scenario involving a series of data bits may take hours.

Another challenge in our experiments is establishing the warm-up period. In a discrete-event simulations, the system is initially empty and idle, and the steady state has not been reached. The situation is different in biological systems, due to the leakiness and basal expression of biological molecules, which can cause inaccurate results (*45*). To improve a molecular channel's efficiency, decoding information in receiver cells should start after a sufficient warm-up period for the system to reach the initial steady state. In our simulations, this steady state is reached after the first two symbol durations. Hence, the first two symbols are chosen as bit-0 in our simulations. Results within the warm-up period are ignored to improve the accuracy of overall results.

The modeling approach presented here demonstrates how efforts in molecular communications and synthetic biology can be combined to provide an integrated view of intracellular and intercellular processes to design novel communication systems. Our approach allows validating molecular communication designs *in silico* and identifing suitable system parameters computationally to inform wet-lab experiments.

3 Methodology

Period Finder was implemented in MATLAB and Java to produce efficient data signals and mitigate ISI caused by diffusion and cellular noise. Diffusion models were integrated with cellular models in a multi-scale approach, and simulations were automated. Our models incorporate cellular signaling to integrate the effect of diffusing molecules to trigger a desired cellular activity within an Rx receiver, for example, to convert a molecular signal into a cellular signal. Here, we designed a genetic circuit that builds upon gene regulatory networks involving transcriptional activation and repression processes to control cellular output in response to sensing intercellular (A_{out} and B_{out}) signaling molecules. Molecular interactions between different circuit components are used to evaluate intracellular signals ($A_{in} - B_{in}$).

Diffusion dynamics were linked to cellular genetic circuit models via molecular communication parameters. These parameters were used to derive genetic circuit models, each representing a possible design. Genetic design automation was implemented in Java. VPR2 (*30*) was used to define models of biological parts such as promoters, RBSs, and CDSs and molecular interactions, such as TF-promoter activation and inhibition, complex formation, and degradation. The model construction process was facilitated using the SVPWrite language to specify the order and types of biological parts. For example, the "`prom1:prom;rbs1:rbs;cds1:cds;ter1:ter`" input specifies a single transcriptional unit where "prom1" is a promoter, "rbs1" is an RBS; "cds1" is a CDS and "ter1" is a terminator. The resulting genetic circuit definition was represented in SBOL2(*31*), and information about molecu-

lar interactions was added. Hierarchical SBML models (*46*) representing potential genetic circuits were derived via VPR2's SBOL-to-SBML conversion. The SBML models were simulated using the COPASI Java bindings (*28*). The simulation results were evaluated by Period Finder. The MOL-eye diagram is adopted for the evaluation of potential communication scenarios.

The movement of particles in a three-dimensional space can be represented via three independent displacements, one for each dimension, where each displacement follows a normal distribution with zero mean and σ^2 variance, denoted as

$$\Delta x, \Delta y, \Delta z \sim \mathcal{N}(0, \sigma^2) \tag{1}$$

where t is the time, $\sigma = \sqrt{2D\Delta t}$, and D is the diffusion coefficient that describes the molecule's mobility inside the fluid (*26*).

Assuming a simple MCvD channel without flow, the expected fraction of diffusing type-A molecules which will reach and be absorbed by the Rx during the time frame t_k can be calculated as

$$E[N^{\mathrm{Rx}}_{A_{out}}(t_k)] = N^{\mathrm{Tx}}_{A_{out}}\{F^{\mathrm{Rx}}(t_k^+) - F^{\mathrm{Rx}}(t_k^-)\} \tag{2}$$

where $E[\cdot]$ is the expectation operator, $N_{A^{\mathrm{Tx}}_{out}}$ is the number of emitted molecules, $F^{\mathrm{Rx}}(t)$ is the time-dependent formula for the expected cumulative fraction of arriving molecules (*34*), t_k^- is the start and t_k^+ is the end of time frame t_k (*18*). For a simple and symmetric topology like a point transmitter and a single spherical absorber, $F^{\mathrm{Rx}}(t)$ is known analytically(*34*).

In the general model shown in Figure 1, the resulting proteins inside the Rx (A_{in} and B_{in}, respectively) can bind together. Therefore, B_{in} can eliminate the effect of stray molecules at the receiver. If A_{in} exceeds a certain level of concentration (λ) in time slot t_k, Rx interprets

the received symbol as sym_1 and sym_0 otherwise. This process can be represented as

$$S[t_k] = \begin{cases} sym_1 & N_{A_{in}}^{\mathrm{Rx}}[t_k] \geq \lambda \\ sym_0 & N_{A_{in}}^{\mathrm{Rx}}[t_k] < \lambda \end{cases}$$

where $S[t_k]$ is the received or decoded symbol in the time slot t_k.

4 Acknowledgements

References

1. Perbal, B. (2003) Communication is the key. Cell Commun. Signaling 1, 3.

2. Buddingh', B. C., Elzinga, J., and van Hest, J. C. M. (2020) Intercellular communication between artificial cells by allosteric amplification of a molecular signal. Nat. Commun. 11, 1652.

3. Yokobayashi, Y., Collins, C. H., Leadbetter, J. R., Arnold, F. H., and Weiss, R. (2003) Evolutionary design of genetic circuits and cell-cell communications. Adv. Complex Syst. 6, 37–45.

4. Ellis, R. J. (2001) Macromolecular crowding: an important but neglected aspect of the intracellular environment. Curr. Opin. Struct. Biol. 11, 114–119.

5. Bush, S. F., Paluh, J. L., Piro, G., Rao, V., Prasad, R. V., and Eckford, A. (2015) Defining Communication at the Bottom. IEEE Trans. Mol. Biol. Multiscale Commun. 1, 90–96.

6. Haselmayr, W., Springer, A., Fischer, G., Alexiou, C., Boche, H., Hoeherk, P., Dressler, F., and Schober, R. (2019) Integration of Molecular Communications into Future Generation Wireless Networks. 6G Wireless Summit

7. Nakano, T., Moore, M., Wei, F., Vasilakos, A., and Shuai, J. (2012) Molecular Communication and Networking: Opportunities and Challenges. IEEE Trans. Nanobioscience 11, 135–48.

8. Deng, Y., Guo, W., Noel, A., Elkashlan, M., and Nallanathan, A. (2017) Enabling Energy Efficient Molecular Communication via Molecule Energy Transfer. IEEE Commun. Lett. 21, 254–257.

9. Chandler, D., and Munday, R. A Dictionary of Media and Communication; OUP Oxford, 2011.

10. Popielec, A., Ostrowska, N., Wojciechowska, M., Feig, M., and Trylska, J. (2020) Crowded environment affects the activity and inhibition of the NS3/4A protease. Biochimie 176, 169–180.

11. Tepekule, B., Pusane, A. E., Yilmaz, H. B., and Tugcu, T. Energy efficient ISI mitigation for communication via diffusion. 2014 IEEE International Black Sea Conference on Communications and Networking (BlackSeaCom). 2014; pp 33–37.

12. Kuran, M. Ş., and Tugcu, T. Co-Channel Interference for Communication via Diffusion System in Molecular Communication. Bio-Inspired Models of Networks, Information, and Computing Systems. Berlin, Heidelberg, 2012; pp 199–212.

13. Pierobon, M., and Akyildiz, I. F. Intersymbol and co-channel interference in diffusion-based molecular communication. 2012 IEEE International Conference on Communications (ICC). 2012; pp 6126–6131.

14. Wang, J., Yin, B., and Peng, M. (2017) Diffusion based molecular communication: principle, key technologies, and challenges. China Communications 14, 1–18.

15. Kuscu, M., Dinc, E., Bilgin, B. A., Ramezani, H., and Akan, O. B. (2019) Transmitter and Receiver Architectures for Molecular Communications: A Survey on Physical Design

With Modulation, Coding, and Detection Techniques. Proceedings of the IEEE 107, 1302–1341.

16. Pierobon, M., and Akyildiz, I. F. (2013) Capacity of a Diffusion-Based Molecular Communication System With Channel Memory and Molecular Noise. IEEE Trans. Inf. Theory 59, 942–954.

17. Noel, A., Cheung, K., and Schober, R. Improving Diffusion-Based Molecular Communication with Unanchored Enzymes. In Bio-Inspired Models of Network, Information, and Computing Systems; Di Caro, G. A., and Theraulaz, G., Eds.; Springer International Publishing: Cham, 2014; pp 184–198.

18. Tepekule, B., Pusane, A. E., Yilmaz, H. B., Chae, C.-B., and Tugcu, T. (2015) ISI Mitigation Techniques in Molecular Communication. IEEE Trans. Mol. Biol. Multiscale Commun. 1, 202–216.

19. Tepekule, B., Pusane, A. E., Kuran, M. , and Tugcu, T. (2015) A Novel Pre-Equalization Method for Molecular Communication via Diffusion in Nanonetworks. IEEE Commun. Lett. 19, 1311–1314.

20. Atakan, B., and Akan, O. B. (2010) Deterministic capacity of information flow in molecular nanonetworks. Nano Commun. Netw. 1, 31–42.

21. Aminian, G., Mirmohseni, M., Nasiri Kenari, M., and Fekri, F. On the capacity of level and type modulations in Molecular communication with ligand receptors. 2015 IEEE International Symposium on Information Theory (ISIT). 2015; pp 1951–1955.

22. Srinivas, K. V., Eckford, A. W., and Adve, R. S. (2012) Molecular Communication in Fluid Media: The Additive Inverse Gaussian Noise Channel. IEEE Trans. Inf. Theory 58, 4678–4692.

23. Koo, B. H., Yilmaz, H. B., Chae, C.-B., and Eckford, A. Detection algorithms for molecular MIMO. 2015 IEEE International Conference on Communications (ICC). 2015; pp 1122–1127.

24. Lin, Y.-K., Lin, W.-A., Lee, C.-H., and Yeh, P.-C. (2015) Asynchronous Threshold-Based Detection for Quantity-Type-Modulated Molecular Communication Systems. IEEE Trans. Mol. Biol. Multiscale Commun. 1, 37–49.

25. Kuran, M. S., Yilmaz, H. B., Demirkol, I., Farsad, N., and Goldsmith, A. (2021) A Survey on Modulation Techniques in Molecular Communication via Diffusion. IEEE Commun. Surv. Tutor. 23, 7–28.

26. Acar, B., Akkaya, A., Genc, G., Yilmaz, H. B., Kuran, M. Ş., and Tugcu, T. Understanding Communication via Diffusion: Simulation Design and Intricacies. In Modeling, Methodologies and Tools for Molecular and Nano-scale Communications: Modeling, Methodologies and Tools; Suzuki, J., Nakano, T., and Moore, M. J., Eds.; Springer International Publishing: Cham, 2017; pp 139–163.

27. Hucka, M., Finney, A., Sauro, H. M., Bolouri, H., Doyle, J. C., Kitano, H., Arkin, A. P., Bornstein, B. J., Bray, D., Cornish-Bowden, A., et al. (2003) The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. Bioinformatics 19, 524–531.

28. Hoops, S., Sahle, S., Gauges, R., Lee, C., Pahle, J., Simus, N., Singhal, M., Xu, L., Mendes, P., and Kummer, U. (2006) COPASI—a COmplex PAthway SImulator. Bioinformatics 22, 3067–3074.

29. Mısırlı, G., Hallinan, J., and Wipat, A. (2014) Composable Modular Models for Synthetic Biology. ACM J. Emerg. Technol. Comput. Syst. 11, 1–19.

30. Mısırlı, G., Yang, B., James, K., and Wipat, A. (2021) Virtual Parts Repository 2: Model-Driven Design of Genetic Regulatory Circuits. ACS Synth. Biol. 10, 3304–3315.

31. Madsen, C., Moreno, A. G., P, U., Palchick, Z., Roehner, N., Atallah, C., Bartley, B., Choi, K., Cox, R. S., Gorochowski, T., Grunberg, R., Macklin, C., McLaughlin, J., Meng, X., Nguyen, T., Pocock, M., Samineni, M., Scott-Brown, J., Tarter, Y., Zhang, M., Zhang, Z., Zundel, Z., Beal, J., Bissell, M., Clancy, K., Gennari, J. H., Göksel,, Mısırlı,, Myers, C., Oberortner, E., Sauro, H., and Wipat, A. (2019) Synthetic Biology Open Language (SBOL) Version 2.3. J. Integr. Bioinform. 16, 20190025.

32. Galdzicki, M., Clancy, K. P., Oberortner, E., Pocock, M., Quinn, J. Y., Rodriguez, C. A., Roehner, N., Wilson, M. L., Adam, L., Anderson, J. C., Bartley, B. A., Beal, J., Chandran, D., Chen, J., Densmore, D., Endy, D., Grünberg, R., Hallinan, J., Hillson, N. J., Johnson, J. D., Kuchinsky, A., Lux, M., Mısırlı, G., Peccoud, J., Plahar, H. A., Sirin, E., Stan, G.-B., Villalobos, A., Wipat, A., Gennari, J. H., Myers, C. J., and Sauro, H. M. (2014) The Synthetic Biology Open Language (SBOL) provides a community standard for communicating designs in synthetic biology. Nat. Biotechnol. 32, 545–550.

33. Mısırlı, G., Nguyen, T., McLaughlin, J. A., Vaidyanathan, P., Jones, T. S., Densmore, D., Myers, C., and Wipat, A. (2019) A Computational Workflow for the Automated Generation of Models of Genetic Designs. ACS Synth. Biol. 8, 1548–1559.

34. Yilmaz, H. B., Heren, A. C., Tugcu, T., and Chae, C.-B. (2014) Three-Dimensional Channel Characteristics for Molecular Communications With an Absorbing Receiver. IEEE Commun. Lett. 18, 929–932.

35. Turan, M., Kuran, M. , Yilmaz, H. B., Chae, C.-B., and Tugcu, T. MOL-eye: A new metric for the performance evaluation of a molecular signal. 2018 IEEE Wireless Communications and Networking Conference (WCNC). 2018; pp 1–6.

36. Westoby, M., Nielsen, D. A., Gillings, M. R., Litchman, E., Madin, J. S., Paulsen, I. T., and Tetu, S. G. (2021) Cell size, genome size, and maximum growth rate are near-

independent dimensions of ecological variation across bacteria and archaea. Ecol. Evol. 11, 3956–3976.

37. Morishita, Y., Kobayashi, T. J., and Aihara, K. (2006) An Optimal Number of Molecules for Signal Amplification and Discrimination in a Chemical Cascade. Biophys. J. 91, 2072–2081.

38. Wang, J., and Hou, T. (2011) Application of molecular dynamics simulations in molecular property prediction II: Diffusion coefficient. J. Comput. Chem. 32, 3505–3519.

39. Aoki, S. K., Lillacci, G., Gupta, A., Baumschlager, A., Schweingruber, D., and Khammash, M. (2019) A universal biomolecular integral feedback controller for robust perfect adaptation. Nature 570, 533–537.

40. Shopera, T., Henson, W. R., Ng, A., Lee, Y. J., Ng, K., and Moon, T. S. (2015) Robust, tunable genetic memory from protein sequestration combined with positive feedback. Nucleic Acids Res. 43, 9086–9094.

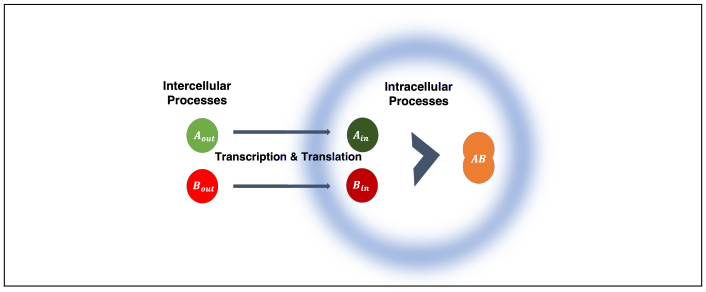
41. Moon, T. S., Lou, C., Tamsir, A., Stanton, B. C., and Voigt, C. A. (2012) Genetic programs constructed from layered logic gates in single cells. Nature 491, 249–253.

42. Kolahi, S. S. Simulation Model, Warm-up Period, and Simulation Length of Cellular Systems. 2011 Second International Conference on Intelligent Systems, Modelling and Simulation. 2011; pp 375–379.

43. Farsad, N., Yilmaz, H. B., Eckford, A., Chae, C.-B., and Guo, W. (2016) A Comprehensive Survey of Recent Advancements in Molecular Communication. IEEE Commun. Surv. Tutor. 18, 1887–1919.

44. Shamir, M., Bar-On, Y., Phillips, R., and Milo, R. (2016) SnapShot: Timescales in Cell Biology. Cell 164, 1302–1302.e1.

45. Grassmann, W. K. (2014) Factors affecting warm-up periods in discrete event simulation. SIMULATION 90, 11–23.

46. Hucka, M., Bergmann, F. T., Hoops, S., Keating, S. M., Sahle, S., Schaff, J. C., Smith, L. P., and Wilkinson, D. J. (2015) The Systems Biology Markup Language (SBML): Language Specification for Level 3 Version 1 Core. J. Integr. Bioinform. 12, 382–549.

Graphical TOC Entry