Analyzing Signal Transduction Circuits for Adaptation of Activated Ras in a Eukaryotic Chemotaxis Pathway

Kritika Kashyap, Department of Chemical and Biomolecular Engineering, Cornell University, Ithaca, NY 14850, USA

Abstract

The development of network inference methodologies that accurately predict connectivity in signaling pathways may enable us to discover network motifs in transcription networks and understand their functionalities. There are many possible patterns that can appear in a network, only a few are found significantly and are network motifs. The network motifs have defined information processing functions. The benefit of these functions may explain why the same network motifs are rediscovered by evolution again and again in diverse systems. To find significant patterns, we will look at different random three-node patterns and analyze their regulation described by each of its 3 edges. Each of these edges can be an activation or a repression interaction. Living cells integrate extremely robust circuits that exhibit significant heterogeneity, but still respond to external stimuli in predictable ways. Adaptation in signaling systems plays a crucial role in eukaryotic chemotaxis. Mathematically modeling for simple network topologies can provide us better understanding of the mechanism of perfect adaptation in bacterial chemotaxis. Logic-based networks can provide a predictive approximation of the transfer of signals in a network.

Introduction

Boolean logic can be used to simulate signal transfer in a molecular and biological network in a manner similar to signal transfer in a digital circuit. Boolean logic models can be used to explicitly simulate enzyme activation and inhibition in a set of biochemical reactions. Logic-based models are powerful tools for studying complex systems, such as intracellular molecular networks, because they are qualitatively predictive and do not require parameter information or mechanistic details needed for more quantitatively precise kinetic methods. Dynamic signaling models based on logic networks have high predictive power only if the underlying logic network is accurately constructed. Of course, this is also true of highly quantitative kinetic-based differential equation models, which require knowledge of the underlying circuitry as well as a very large number of unknown parameters.

In order to rationally develop precise understanding of Ras signaling pathway (downstream from G protein couples receptors) in bacterial chemotaxis, it is critical to develop methodologies that can help us understand how complex intracellular signaling pathways are wired as an integrated network. Many biological systems exhibit perfect adaptation using an integral control strategy in which a buffering component of the signaling network integrates the difference between the response and the desired basal amount. This difference is then fed back to achieve perfect adaptation via negative regulation. A systematic computational analysis of a three-node network reveals that an incoherent feedforward mechanism can also achieve robust perfect adaptation. In this topology two nodes of the network are activated proportionally by the input stimulus. These two nodes then act on the third node with opposite effects (such that one activates and one inhibits), leading to a transient response that adapts perfectly. To date, no clear examples of biological networks that use the incoherent feed-forward strategy have been identified, which is surprising because networks that contain feedforward loop (FFL) perform better than those that use integral control.

Methodology

Model equations demonstrating the three-node incoherent FFL for Ras signaling pathway have been studied and described ((1) and Appendix A). To analyze the adaptation kinetics of a eukaryotic chemotaxis

pathway, the system of cells in (1) was exposed to sudden uniform increases and decreases in the concentration of the chemoattractant cAMP. This helped in the examination of the dynamics of activated Ras, Ras-GTP, with the Ras binding domain of human Raf1 [RBD-green fluorescent protein(GFP)] as a reporter, which preferentially measures activated Ras. Ras proteins are activated by RasGEF's and are inactivated by RasGAP's.

In incoherent FFL, both RasGEF and RasGAP are activated by the chemo-attractant (cAMP) signal acting through the receptor R (Fig.1A). Because RasGEF and RasGAP activate and inhibit Ras, respectively, this network can have the properties of ultrasensitivity. Several species of receptors with different binding affinities for cAMP exist. Two types of receptors have been included in the model, one with a high affinity (R_1, with a dissociation constant K_{d1}=60nM) and one with a low affinity (R_2, with a dissociation constant K_{d2}=450nM). It is assumed that the two receptors activate downstream components in an identical fashion. These models can be cast in terms of ordinary differential equations (Appendix A) that can be integrated in time to determine the dynamical response for various temporal patterns of the stimulus as a function of changes in the amount of stimulus. The differential equations have been solved on MATLAB, software for ordinary differential equations, using the ode45 solver (Range Kutta 4^{th} order). Initial Value Problem (Cauchy Problem) solving methodology was used. The numerical fit for experimental and simulated data in (1) using 21 data points constrained the nine free model parameters (Table 1) which are used in this paper.

Networks consisting three nodes namely RasGEF, RasGAP and Ras-GTP have been used to analyse adaptation kinetics. The steady-state value of all nodes can be determined explicitly in a logic model if the initial values of all nodes in the network are known a priori. Equations A.4, A.5, A.6 from Appendix A can be equated to zero under steady state assumption to get the initial values for RasGEF, RasGAP, Ras-GTP. It is rarely practical to measure all molecules in realistic biological networks, this simulation approach requires only knowledge of the initial activation state of nodes (e.g., constitutively *ON* or *OFF* proteins). The output of a simulation is a value between 0.0 and 1.0 for each node, indicating the probability a node is activated (*ON*) in the steady state. Table 2 lists the logic regulations for a three-node test network. In a three-node circuit consisting of A, B, and C (Fig. 2), there are 729 interactions and 2,197 logical networks possible (Table 2 and (3)). The ON probabilities were calculated for various test networks (see Fig.2 where each network is encoded as [Ai, Bi, Ci], and i indicates a unique logic rule identifier possible for a given node).

Using a particular chemoattractant concentration ode45 (cAMP = 1000 nM) was used to integrate eq A.3 (Appendix A). This gave a list of 150941 values for R for integrating from t = 0 to t = 30. This list was imported from MATLAB into python using scipy.io. Values of R helped set the initial values for A, B, C in the three-node networks in Fig.1 and Fig.6 which were analysed based on the logic rules in Table 2.

Results and Discussions

The RBD-GFP amount in incoherent feedforward could respond to a wide range of concentrations of chemoattractant stimuli and adapted quickly (Fig.3).

The results reveal that the activation kinetics of RasGEF were faster than that of the RasGAP. This is expected because a positive Ras-GTP response after an increase in cAMP requires that the activation step initially is larger than the deactivation step. Eventually, the RasGAP kinetics catches up, resulting in a steady state Ras-GTP amount that is independent of the stimulus strength. The kinetics of the model is shown in (Fig 4), where Ras-GTP is plot as a function of time, after a sudden increase in chemoattractant concentration. This result validates the experimental results in (1).

Analyzing the ON probabilities for incoherent FFL shown in Fig.1, we can see that certain network topologies give similar results (Fig.5). Two trends can be clearly noticed. Logical networks [0,5,11],

[0,3,11], [0,4,9], [13,0,2] show an increasing ON probability of Ras and then decrease. Logic networks [0,4,10], [14,0,2], [0,5,12], [0,3,12] shows a decreasing ON probability and then a sudden increase. As validated from our results from Fig.4, Ras GTP shows an increment till the time that RasGAP catches up with the kinetics, and starts inhibiting the protein Ras. Hence [0,5,11], [0,3,11], [0,4,9], [13,0,2] can be shortlisted as the biologically plausible networks for Ras signaling pathway in bacterial chemotaxis.

Analyzing the ON probabilities for integral control shown in Fig.6, we can see that certain network topologies give similar results (Fig.7). Logic Networks [0,2,9], [13,4,2], [13,2,2], [0,1,11], [15,1,2], [15,3,2], [16,3,2] show an increase in ON probability of Ras which then reaches a steady value. Logic Networks [0,2,10], [14,4,2], [14,2,2], [0,1,12], [16,1,2] show an increasing trend in Ras ON probability. None of these networks are in agreement with the trend obtained in Fig.4 which is an experimentally validated simulation result (1).

Conclusion

Adaptation in the RasG signaling pathway does not rely on integral control mechanisms that contain negative feedback loops. The integral control mechanism was incompatible with the experimental results in (1). Instead, and unlike any other biological systems analyzed to date, adaptation is achieved through the simultaneous activation of an activator and an inhibitor (Fig.1A).

Scope for Future

To infer the signaling network circuitry inside a cell from experimentally measured western blot data using asynchronous Boolean logic simulation methods and a heuristic search with genetic algorithms. The objective of this approach is to identify the most likely underlying structure of a signaling network by minimizing the difference between simulated model output and experimental data collected from a series of molecular perturbation experiments. Steady-state node probabilities greater than 0.0 but less than 1.0 indicate that, for the conditions tested, (a) the node is oscillating as part of a limit cycle attractor, (b) the node can be in one of two fixed point attractors, or (c) the node can be in both a limit cycle and a fixed point attractor. Heuristic search with genetic algorithms can be used to identify candidate logic networks consistent with observed data collected from a set perturbation experiments. Values of *ON* or *OFF* can be assigned to indicate the phosphorylation state of proteins in the cells of social amoeba Dictyostelium discoideum for each condition tested. In general, faint or absent blots can be assigned *OFF* values (0.0) and prominent blots can be assigned *ON* values (1.0). A predictive score (p-score) formula can be developed to quantify how well simulated model output matches experimentally collected data.

This method will help in matching the simulated results with experimental results and eliminate various network motifs which are biologically infeasible. Hence our search for signaling networks with lower p-score (well matched with experimental data) will bring us closer to discovering the signaling networks inside cells.

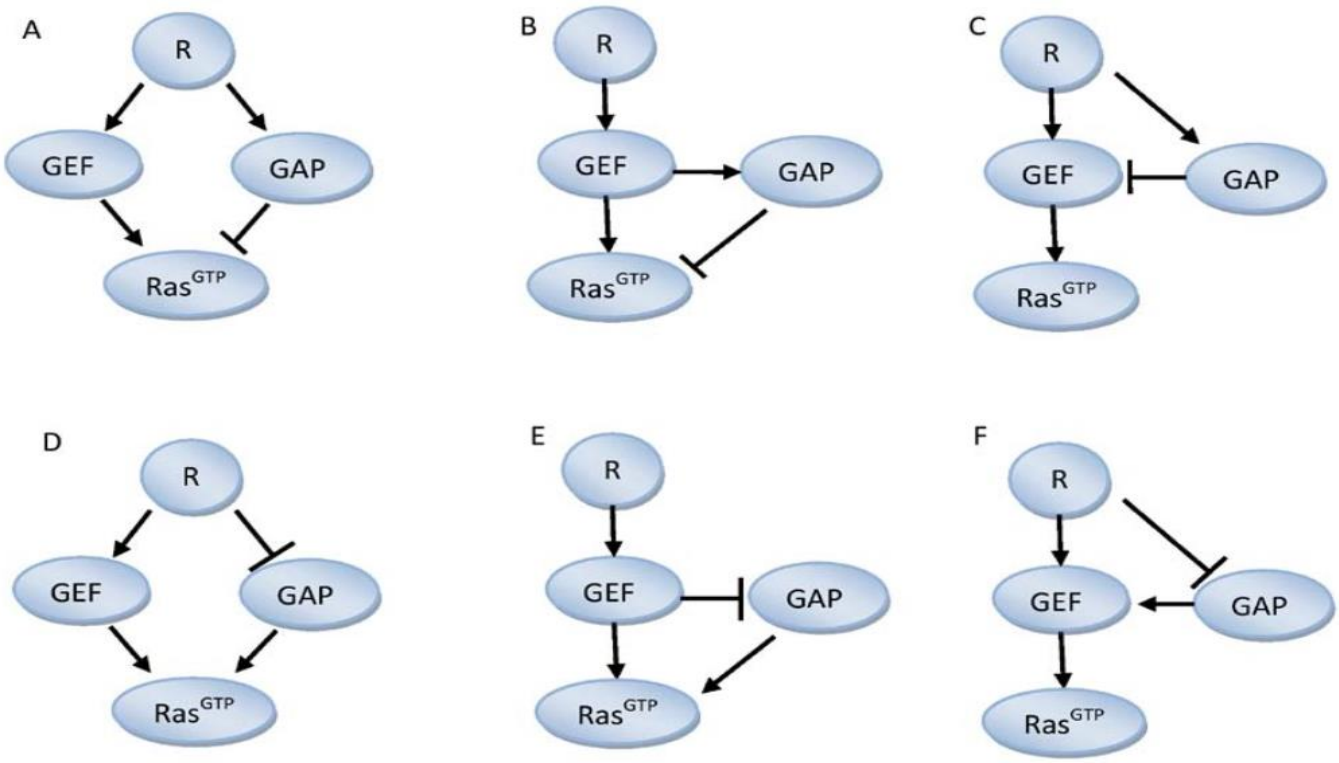


Figure 1: Theoretically possible Incoherent FeedForward Topologies, including ones that are biologically not plausible

Table 1: Model Parameters used in this study.

parameter	value	parameter	value
R_1^{tot}	0.1	k_{-GEF}	0.4 sec^{-1}
R_2^{tot}	0.9	k_{GAP}	0.01 sec^{-1}
k_{R1}	0.00267 $nM^{-1}sec^{-1}$	k_{-GAP}	0.1 sec^{-1}
k_{-R1}	0.16 sec^{-1}	Ras^{tot}	1
k_{R2}	0.00244 $nM^{-1}sec^{-1}$	k_{Ras}	390 sec^{-1}
k_{-R2}	1.1 sec^{-1}	k_{-Ras}	3126 sec^{-1}
r_1	0.012 nM	RBD^{tot}	1
r_2	0.115 nM	k_{RBD}^{off}	0.53 sec^{-1}
k_{GEF}	0.04 sec^{-1}	k_{RBD}^{on}	1.0 sec^{-1}

Table 2: The * after the target node indicates the value the target node will assume in the next time step.

Rule Count id	Node Rule id	Logic Regulation
1	0	A* = A
2	1	A* = not C
3	2	A* = C
4	3	A* = not B
5	4	A* = B
6	5	A* = not B and not C
7	6	A* = not B or not C
8	7	A* = not B and C
9	8	A* = not B or C
10	9	A* = B and not C
11	10	A* = B or not C
12	11	A* = B and C
13	12	A* = B or C
14	13	A* = R and not B
15	14	A* = R or not B
16	15	A* = R and B
17	16	A* = R or B
18	0	C* = C
19	1	C* = not A
20	2	C* = A
21	3	C* = not B
22	4	C* = B
23	5	C* = not A and not B
24	6	C* = not A or not B
25	7	C* = not A and B
26	8	C* = not A or B
27	9	C* = A and not B
28	10	C* = A or not B
29	11	C* = A and B
30	12	C* = A or B
31	0	B* = B
32	1	B* = not C
33	2	B* = C
34	3	B* = not A
35	4	B* = A

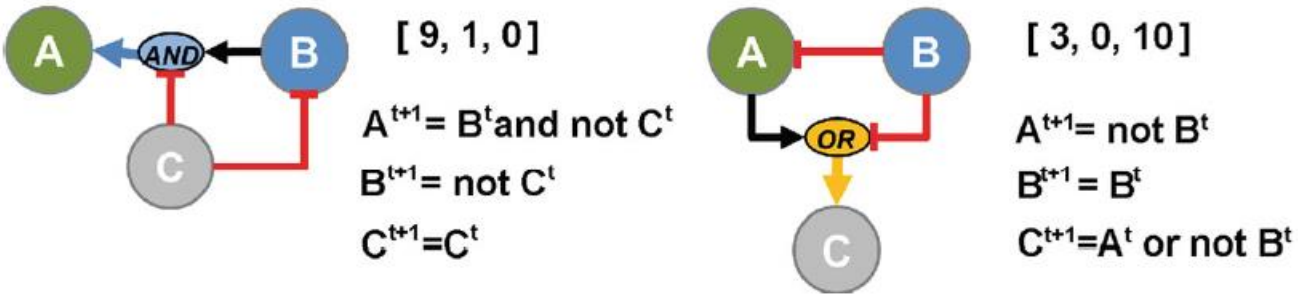


Figure 2: A=GEF B=GAP C=Ras-GTP, each logic network encoded as [Ai, Bi, Ci].

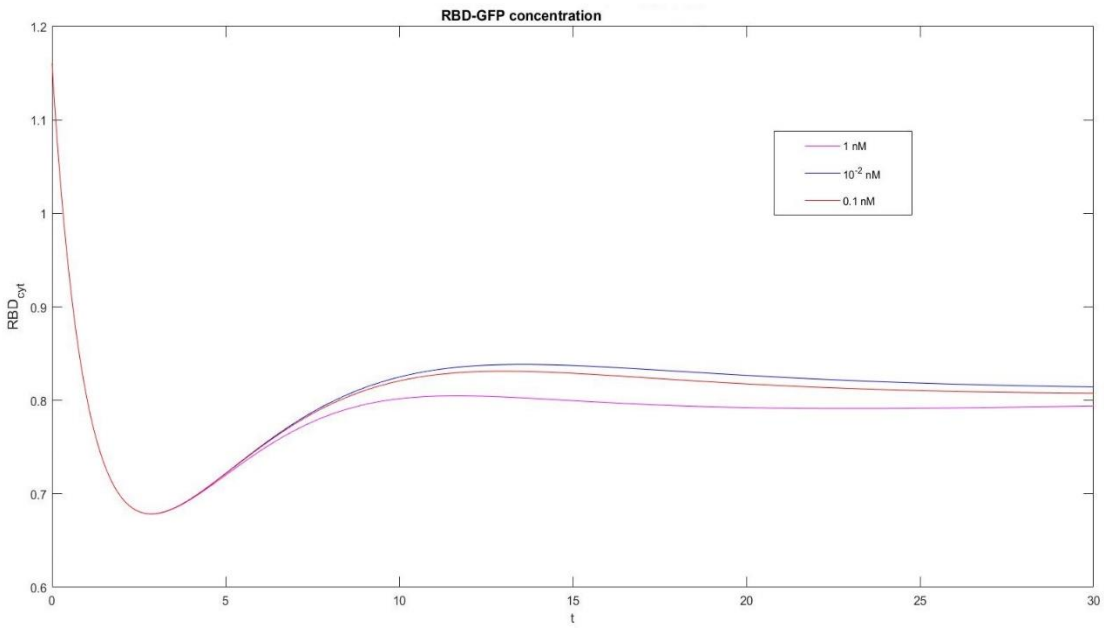


Figure 3: RBD-GFP concentration, for different amounts of chemoattractant stimulation

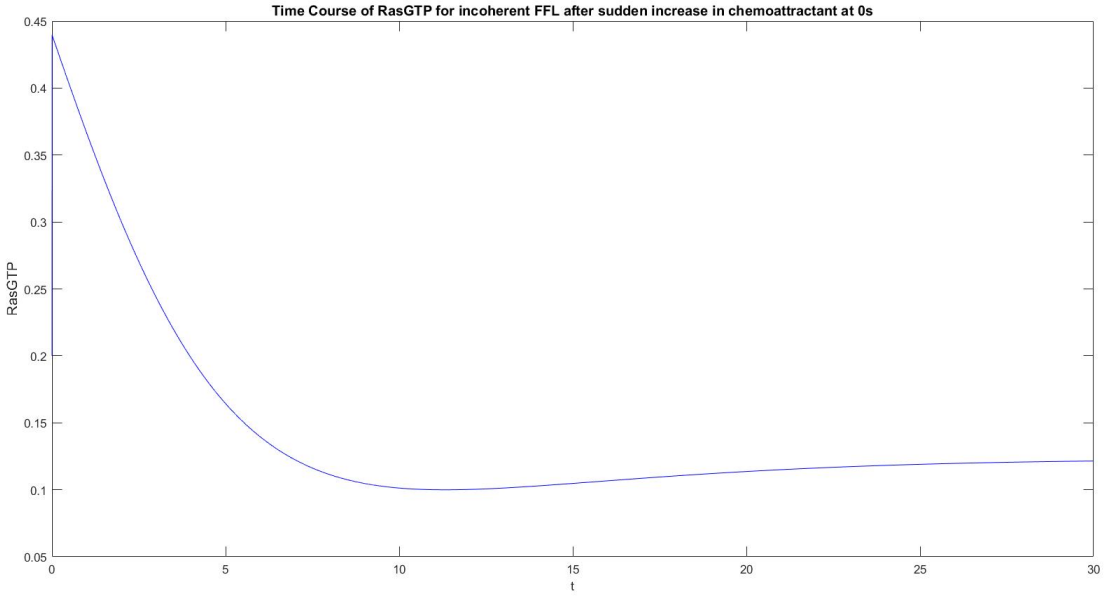


Figure 4: A typical time course of Ras-GTP for the incoherent feedforward model after a sudden increase in chemo-attractant at 0 s.

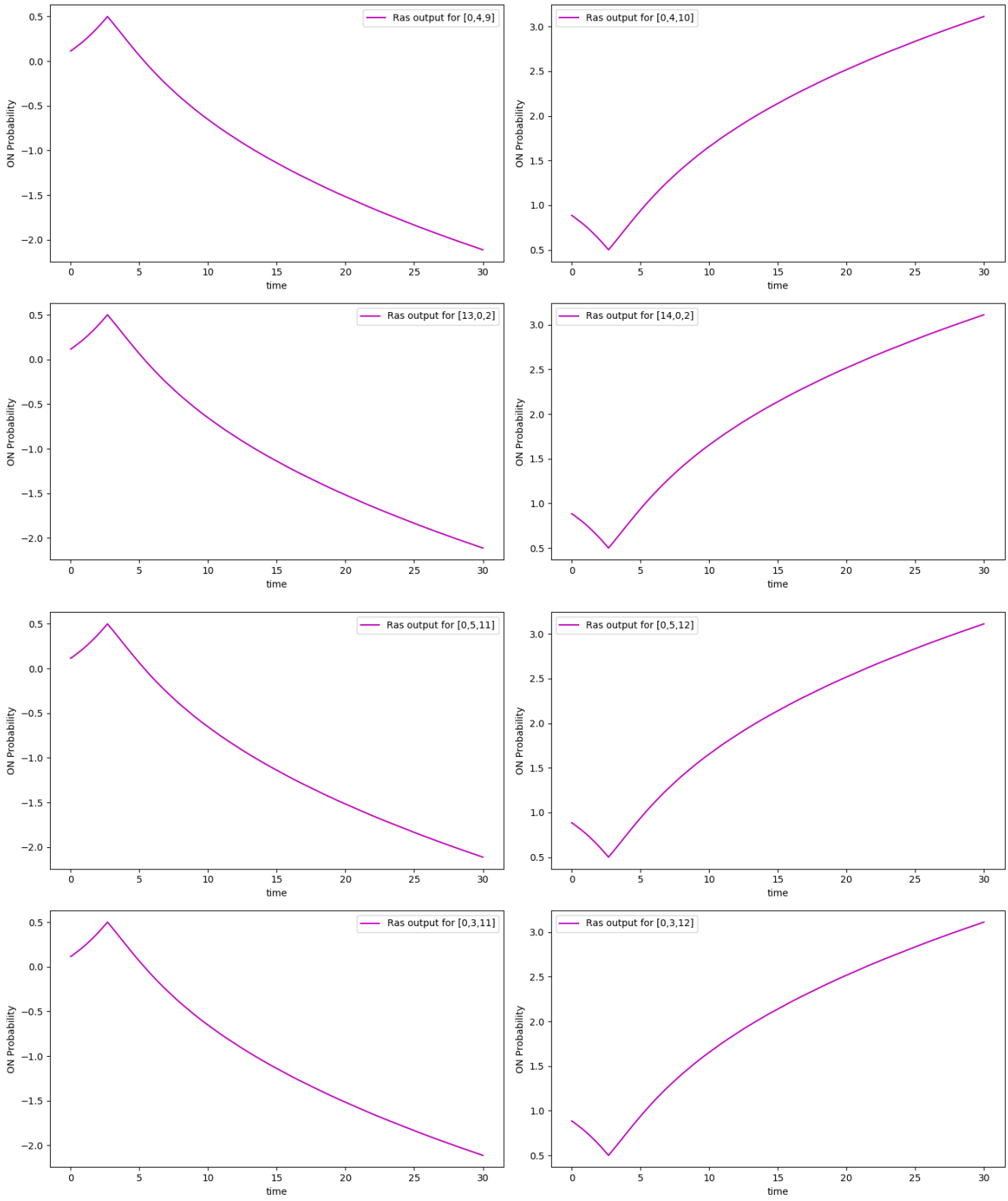


Figure 5: The Ras-GTP ON probability for networks in Fig.1 (B, C, D, E)

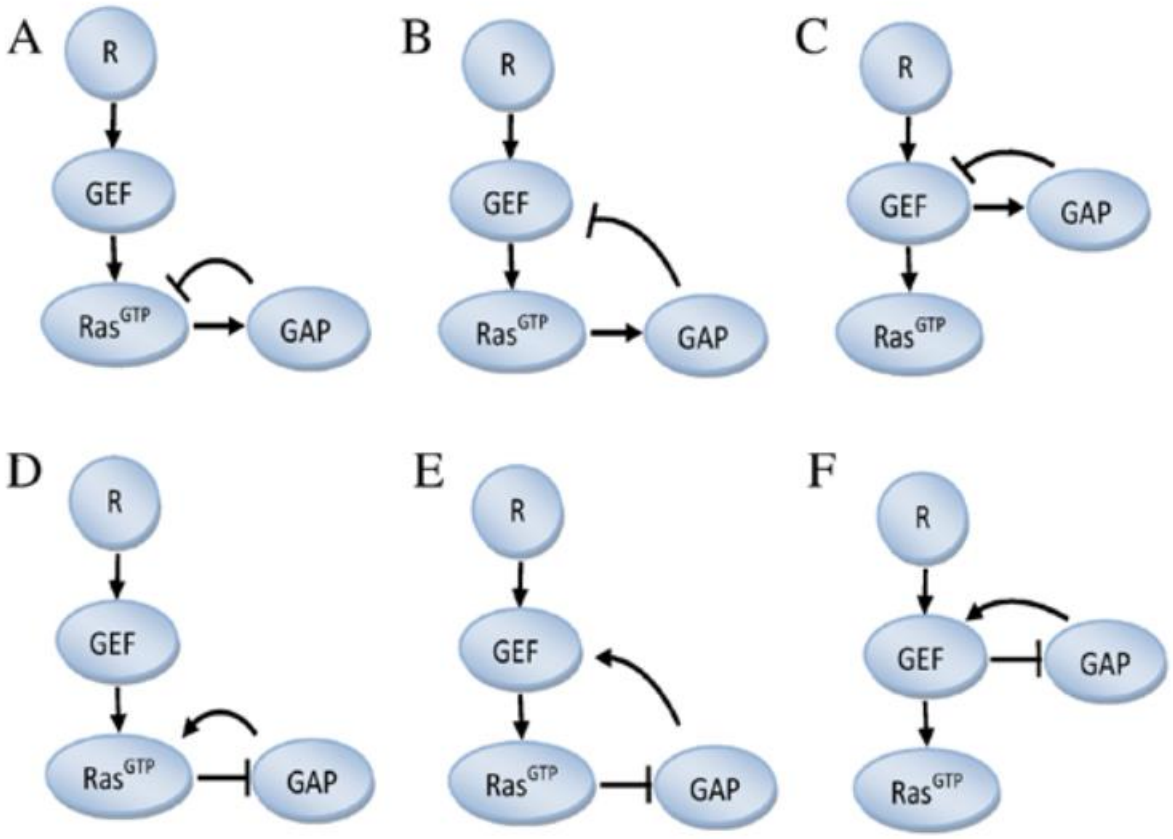
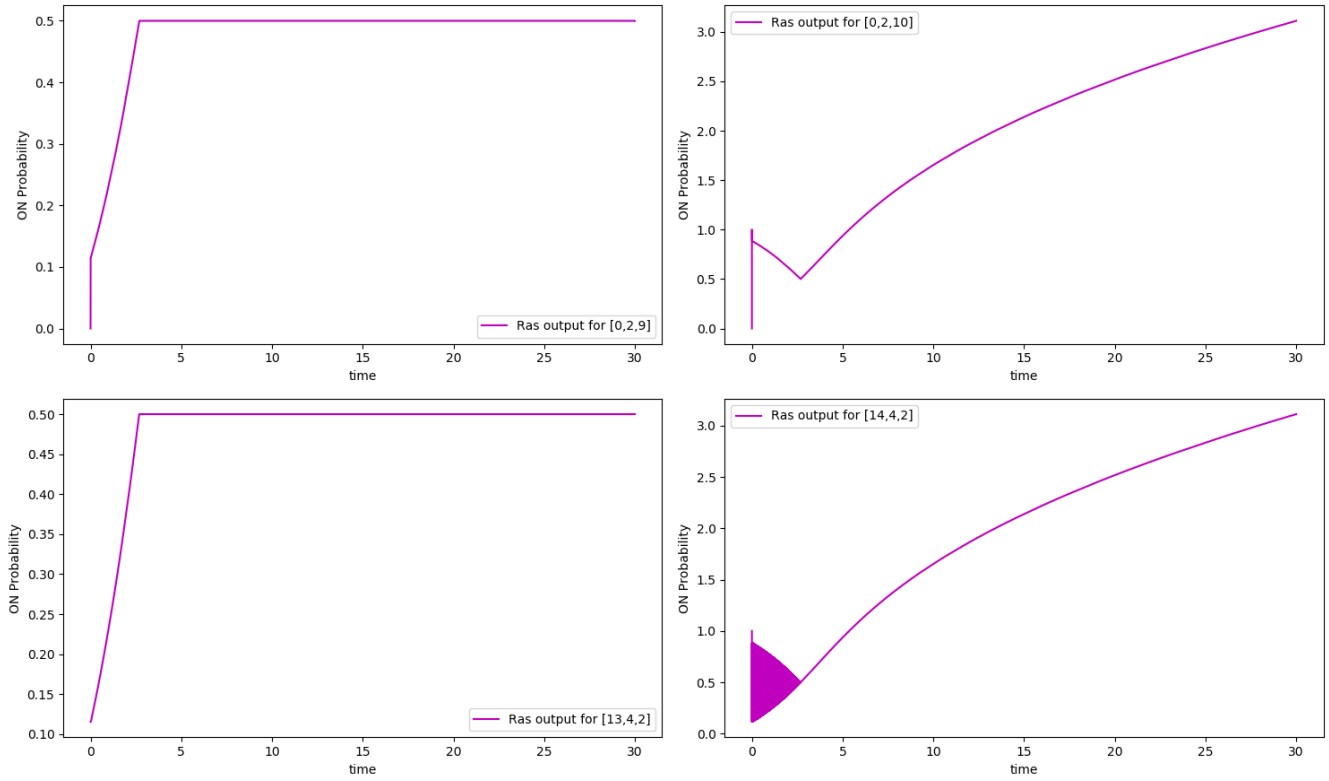


Figure 6: All possible three-node topologies containing a negative feedback loop (integral control topologies), including ones that are biologically not plausible.



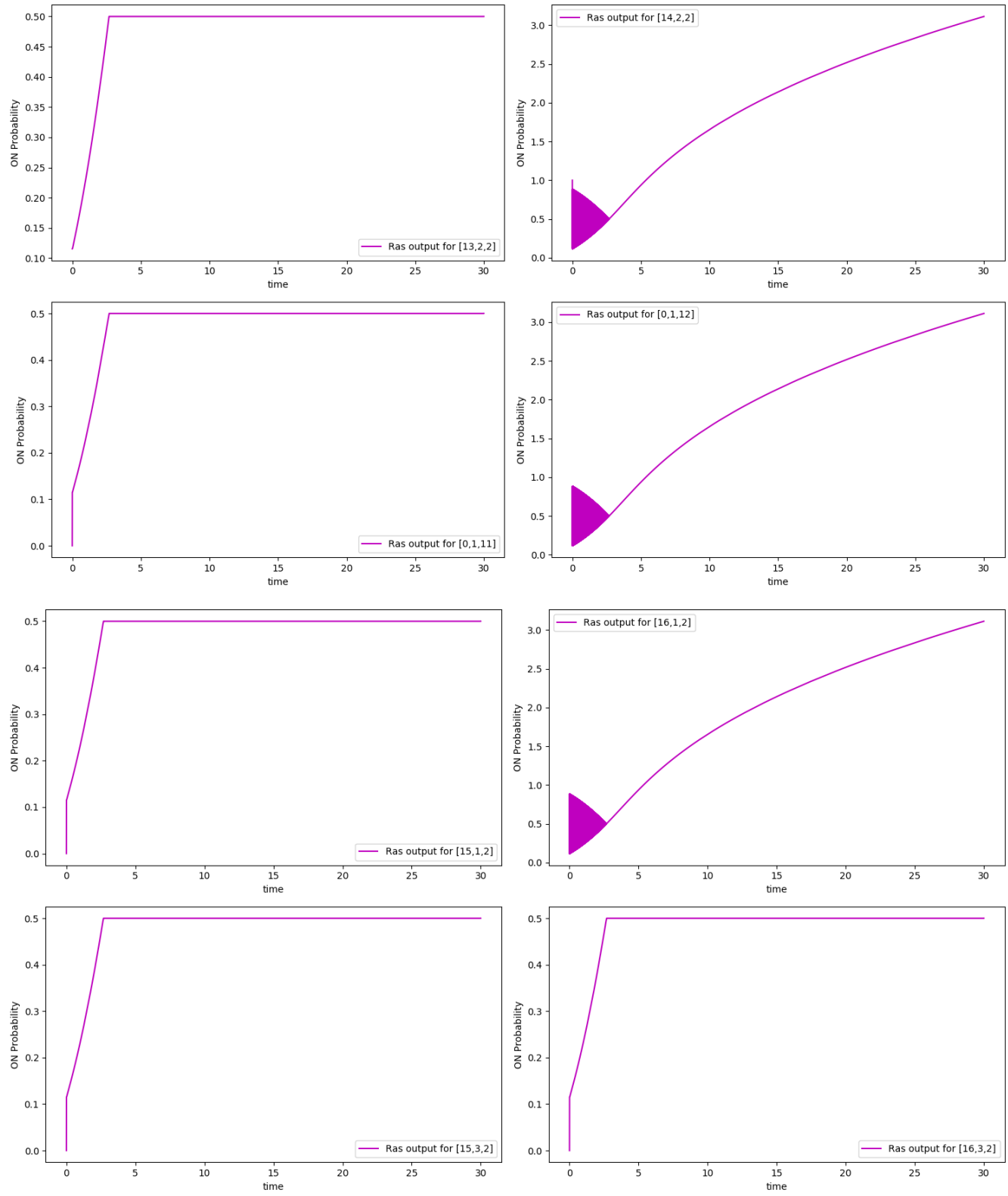


Figure 7: The Ras-GTP ON probability for networks in Fig.6 (A, B, C, D, E, F)

Acknowledgements
I would like to extend my thanks and appreciation to Professor J.Varner and Professor M.Paszek for providing me excellent guidance in the course of Advanced Biomolecular Engineering. Professor Varner provided a motivating, enthusiastic, and critical atmosphere during the discussions we had. Professor Paszek's valuable advice, guidance and encouragement led to the successful completion of the project. I am also thankful to my fellow classmates in the course CHEME7770 for their assistance throughout the project work.

References

1. Kosuke Takeda, Danying Shao, Micha Adler, Pascale G. Charest, William F. Loomis, Herbert Levine, Alex Groisman, Wouter-Jan Rappel, Richard A. Firtel. Incoherent Feedforward Control Governs Adaptation of Activated Ras in a Eukaryotic Chemotaxis Pathway (Feb19, 2014).
2. Michelle L. Wynn, Megan Egbert, Nikita Consul, Jungsoo Chang, Zhi-Fen Wu, Sofia D. Meravjer, Santiago Schnell. Inferring Intracellular Signal Transduction Circuitry from Molecular Perturbation Experiments (June30, 2016).
3. Uri Alon. An Introduction to Systems Biology: Design Principles of Biological Circuits (Chapman & Hall/CRC Mathematical and Computational Biology).

Github Link for codes
https://github.com/kspkritika/Signal-Transduction-in-Bacterial-Chemotaxis

Appendix A

If we assume that the cytosolic concentration of all components is uniform, we can cast the incoherent feed forward model and the integral control model in terms of a set of coupled ordinary differential equations (ODEs). These equations, which describe the dynamics of the concentrations of the components, can then be easily integrated to determine the dynamics of the various components. For the incoherent feedforward model, the equations take on the form:

$$\frac{\mathrm{d}R_1}{\mathrm{d}t} = k_{R1}(cAMP + r_1)(R_1^{tot} - R_1) - k_{-R1}R_1 \tag{A.1}$$

$$\frac{\mathrm{d}R_2}{\mathrm{d}t} = k_{R2}(cAMP + r_2)(R_2^{tot} - R_2) - k_{-R2}R_2 \tag{A.2}$$

$$R = R_1 + R_2 \tag{A.3}$$

$$\frac{\mathrm{d}GEF}{\mathrm{d}t} = k_{GEF}R - k_{-GEF}GEF \tag{A.4}$$

$$\frac{\mathrm{d}GAP}{\mathrm{d}t} = k_{GAP}R - k_{-GAP}GAP \tag{A.5}$$

$$\frac{\mathrm{d}Ras^{GTP}}{\mathrm{d}t} = k_{Ras}GEF(Ras^{tot} - Ras^{GTP}) - k_{-Ras}GAPRas^{GTP} \tag{A.6}$$

$$\frac{\mathrm{d}RBD^{cyt}}{\mathrm{d}t} = k_{RBD}^{off}(RBD^{tot} - RBD^{cyt}) - k_{-RBD}^{on}Ras^{GTP}RBD^{cyt} \tag{A.7}$$

The first two equations describe the binding process of the external chemoattractant, cAMP, to the two receptor populations, R_1 and R_2. One population has a large K_d value and one has a small K_d value. The downstream activity of the bound receptors is assumed to be the same for both populations, such that the effective input in the equations for the downstream components is simply the sum, R. Also, we have allowed for the possibility of constitutive activation, parameterized through r_1 and r_2. The fourth and fifth equations describe the first order activation and deactivation of RasGEF and RasGAP (denoted for brevity by GEF and GAP in Fig.1A), and the sixth equation models the dynamics of activated Ras, denoted here as *RasGTP*. The total concentration of Ras is given by *Rastot*. The final equation describes the cytosolic reporter molecule RBD-GFP, denoted by *RBDcyt*. Its total concentration is *RBDtot* and it binds membrane-bound activated Ras, leading to a removal from the cytosol, and is removed from the membrane with simple first order kinetics.