Mini-Lecture 1

Computational Modeling of Physiological Systems: An overview

What is physiology?

- The study of the *function* of biological systems.
  - Contrasts with anatomy: the study of the *structure* of biological systems.
- "Biological Systems" span a tremendous range!

System	**Studies of function**
Low-level chemistry of life	Biochemistry
Individual cells	**Cellular Physiology**
A single tissue or organ system	**[System] Physiology (e.g. Renal, Cardiac, Neural, etc).**
A whole organism	**Organismal Physiology**
Populations of organisms	Ecology

Modeling Physiological Systems

- Physiology is a huge field; in this class we will barely touch the surface.
- A common feature is that we are dealing with complex systems with multiple interacting components.
- Mathematical models can provide deep insight into how these systems work.
- All but the very simplest models typically need to be implemented computationally.

Course goals

- To learn some fundamental concepts in cellular and neural physiology.
- To learn how to develop mathematical models that describe biological systems in varying level of detail.
- To learn how to to use the scientific programming tool Matlab to implement such models.
- To learn how to perform computational experiments that can give insight into the underlying biological system.

A multidisciplinary field

- The big-picture problems come from biology.
- The models we use are based on ideas from physics and chemistry.
- The models are written in the language of mathematics.
- The models are implemented using ideas from computational science.
- The interpretation of computational experiments takes us back to biology.
- These fields all have different ways of thinking and different “languages” for discussing problems.

Conversational Physics?

- In order to do meaningful work in a multidisciplinary field requires a comfort with the various languages that contribute.
- The language of math, of biology, of physics, of computer science …
- You need not be an expert in all these fields, but you should aim to gain a “conversational understanding” in them.

A multidisciplinary class

- Every year, this class has students from many different academic backgrounds:
  - Applied Math & Statistics and (Pure) Mathematics
  - Biology and Biochemistry
  - Biomedical and Chemical & Molecular Engineering
  - Chemistry and Physics
  - Computer Science
  - Mechanical, Computer, and Electrical Engineeering
  - Psychology, Economics, Music, and more
- Take advantage of the diversity of your peers!

Summary

- Computational modeling of physiological processes involves ideas from many different disciplines.
- This can create unique challenges for students, but also makes the material very interesting!
- In this course, we will work on developing the multidisciplinary perspective needed in this field.
  - You are not expected to already have it!

Mini-Lecture 2

Cell-Fate Decisions:

A Grand-Challenge Problem in Cellular Physiology

Cellular Physiology is Foundational

- The cell is the fundamental unit from which all branches of the tree of life derive.
- While multicellular organisms exhibit many emergent behaviors that transcend the capabilities of their individual cells, these behaviors must somehow derive from individual cellular behaviors.
- We can not understand organismal biology without understanding cellular biology.

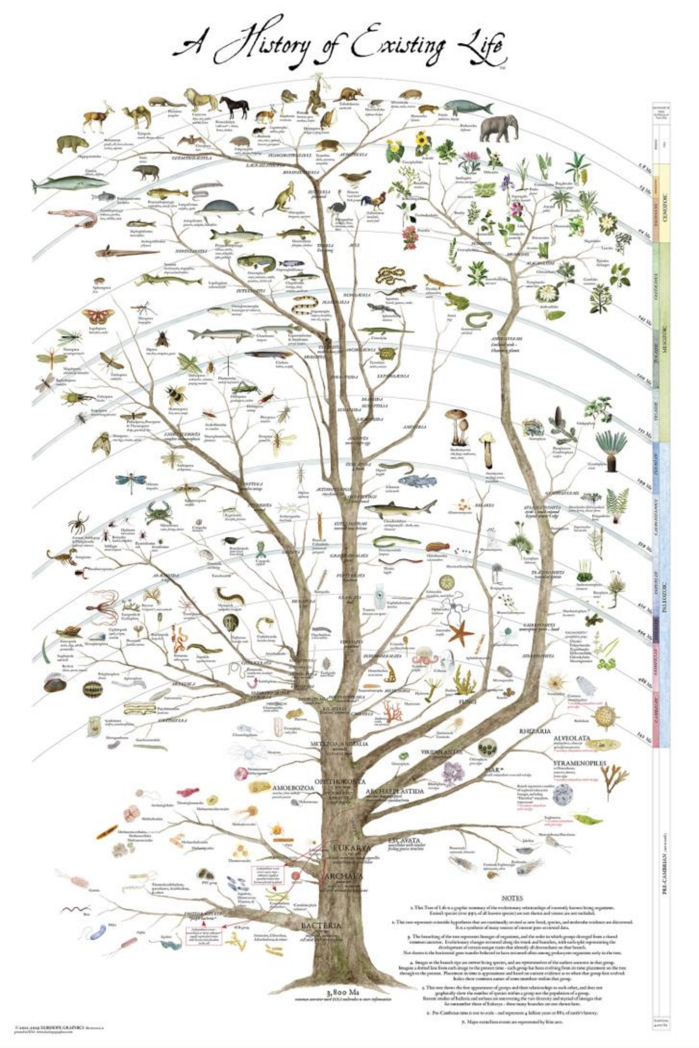


Image from: https://fairhopegraphics.com/ahistoryofexistinglife/a-history-of-existing-life

From Cellular Physiological Processes …

- Catabolism: Processes that break down larger molecules into smaller ones, to produce energy and/or building blocks.
- Anabolism: Processes that build up the molecules of life from smaller components (usually using energy).
- Transport: Processes that move material in and out of the cell and/or from one part of the cell to another.
- Signal transduction: Processes that transmit information from one part of a cell to another.
- Gene regulation: Processes that control which genes in a cell are expressed (“turned-on”) and at what level.

To Fundamental Cellular Behaviors …

The diverse fundamental physiological processes ultimately drive observable cellular behaviors:

- Cell movement: does a cell move or not; if so, how fast and in what direction?
- Cell growth: does a cell grow over time or remain a constant size.
- Cell division: does an individual cell divide into two or remain a single cell?
- Cell differentiation: does a cell remain of the same “type,” or does it transform into a different type?

To Essential Biology.

These "cell-fate decisions" underly many essential processes in biology.

- Development of multicellular organisms from a single cell.
- Maturation of blood cells in healthy adults.
- Wound healing.
- Development of tumors, both benign and malignant.
- Selection of virulent or non-virulent states in microbial infections.
- Microbial biofilm formation.
- Selection of lytic vs lysogenic states of viruses.

The Cell-Fate Grand Challenge

Fundamental scientific understanding:

- Can we understand why, how, and under what conditions a given cell will make a particular cell-fate decision?

Predictive biology and medicine:

- Can this understanding be used to make robust predictions about the outcome of potential cellular behaviors?

Bioengineering, therapeutic development and personalized medicine:

- Can we manipulate a system so that a specific cell makes a particular desired cell-fate decision?

Summary

- Low-level physiological processes (biochemistry) underly observable cellular behaviors.
- Cell-fate decisions are a class of cellular behaviors that are particularly interesting and important.
- A deep, predictive understanding of how cell-fate decisions are made (and how to affect them) would have profound impacts on biology and medicine.

Mini-Lecture 3

Mechanistic foundations:
Protein Function, The Central Dogma, and Gene Regulation

Proteins are the engines of biology

The vast majority of biological functions are driven, at a molecular level, by proteins.

- **Enzymes**: catalyze specific chemical reactions
- **Receptors and binding proteins**: physically associate with other molecules (including other proteins or DNA)
- **Structural proteins**: have shapes and/or mechanical properties that play a functional role
- Many proteins combine several functions!

Enzymes drive catabolism

- Glycolysis and the citric acid cycle are two essential pathways in aerobic respiration; how organisms obtain energy by “burning” glucose.
- A set of enzymes sequentially catalyze the necessary reactions.

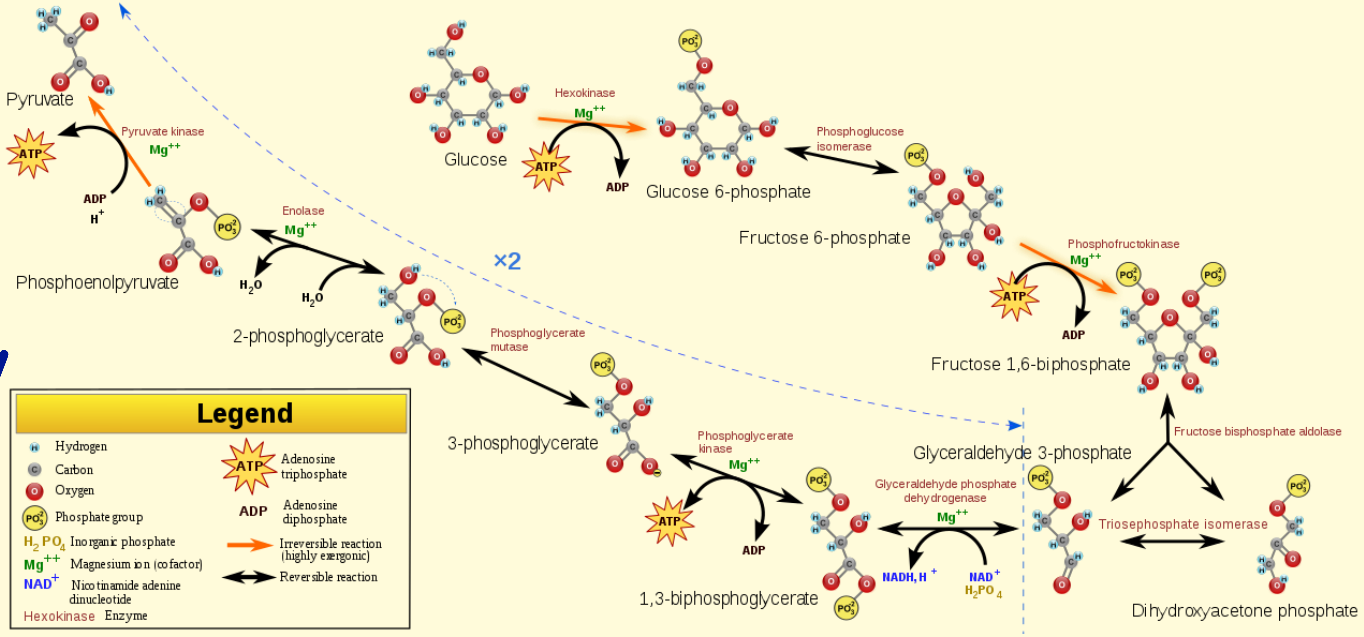


Image from: https://en.wikipedia.org/wiki/Glycolysis

Enzymes drive catabolism

- Glycolysis and the citric acid cycle are two essential pathways in aerobic respiration; how organisms obtain energy by “burning” glucose.
- A set of enzymes sequentially catalyze the necessary reactions.

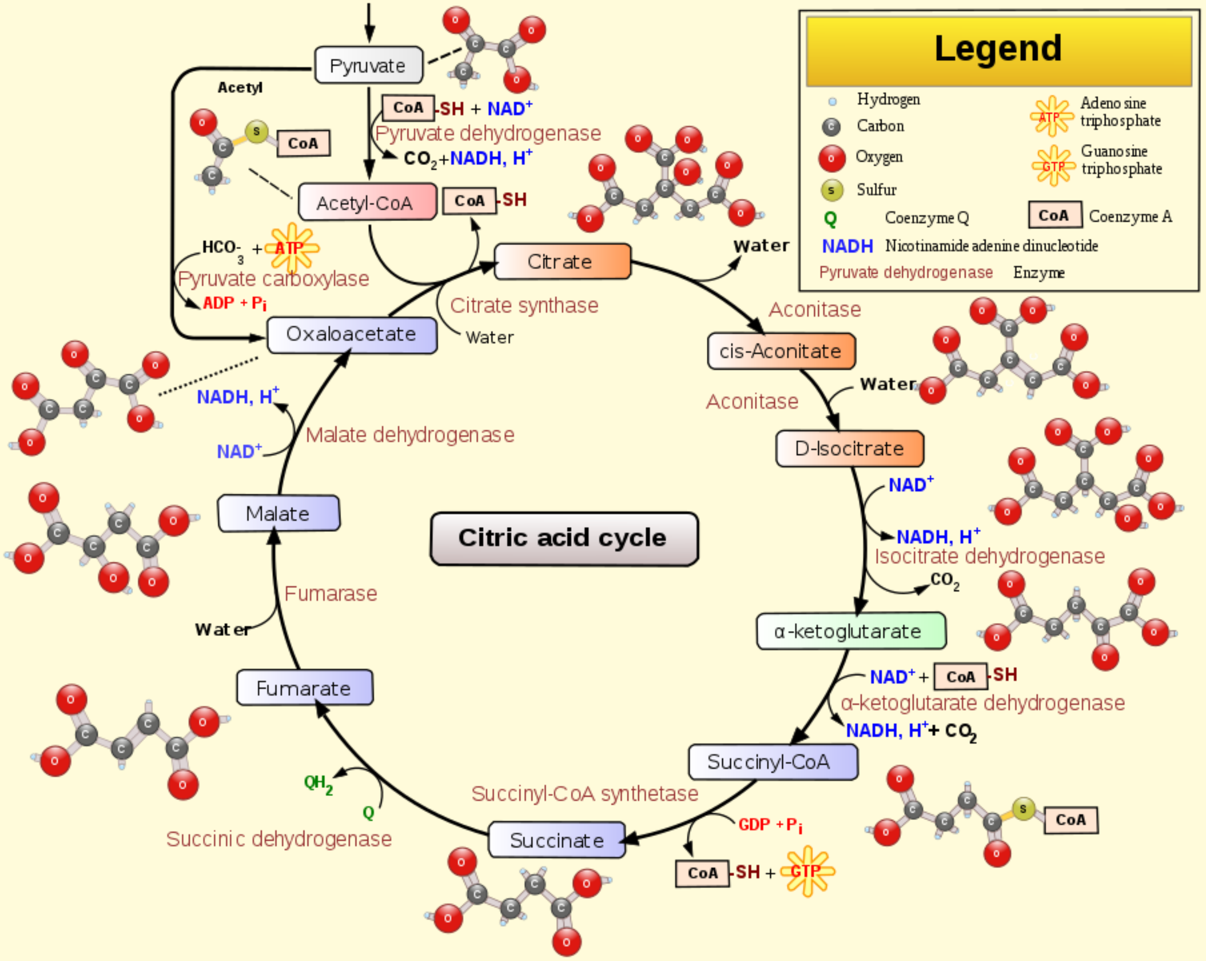
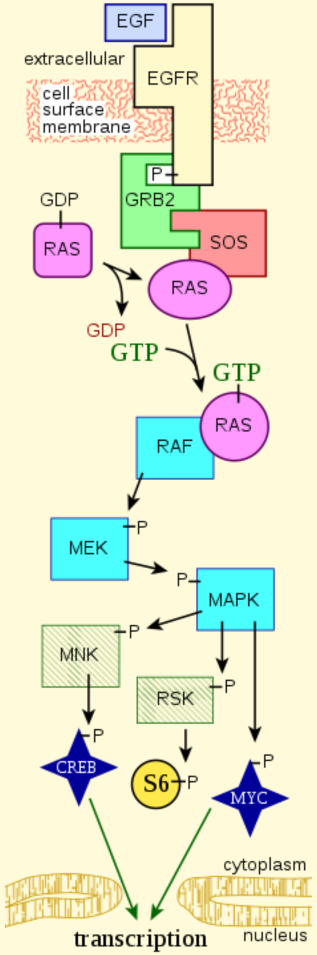


Image from: https://en.wikipedia.org/wiki/Citric_acid_cycle

Receptors and enzymes together drive signal transduction.

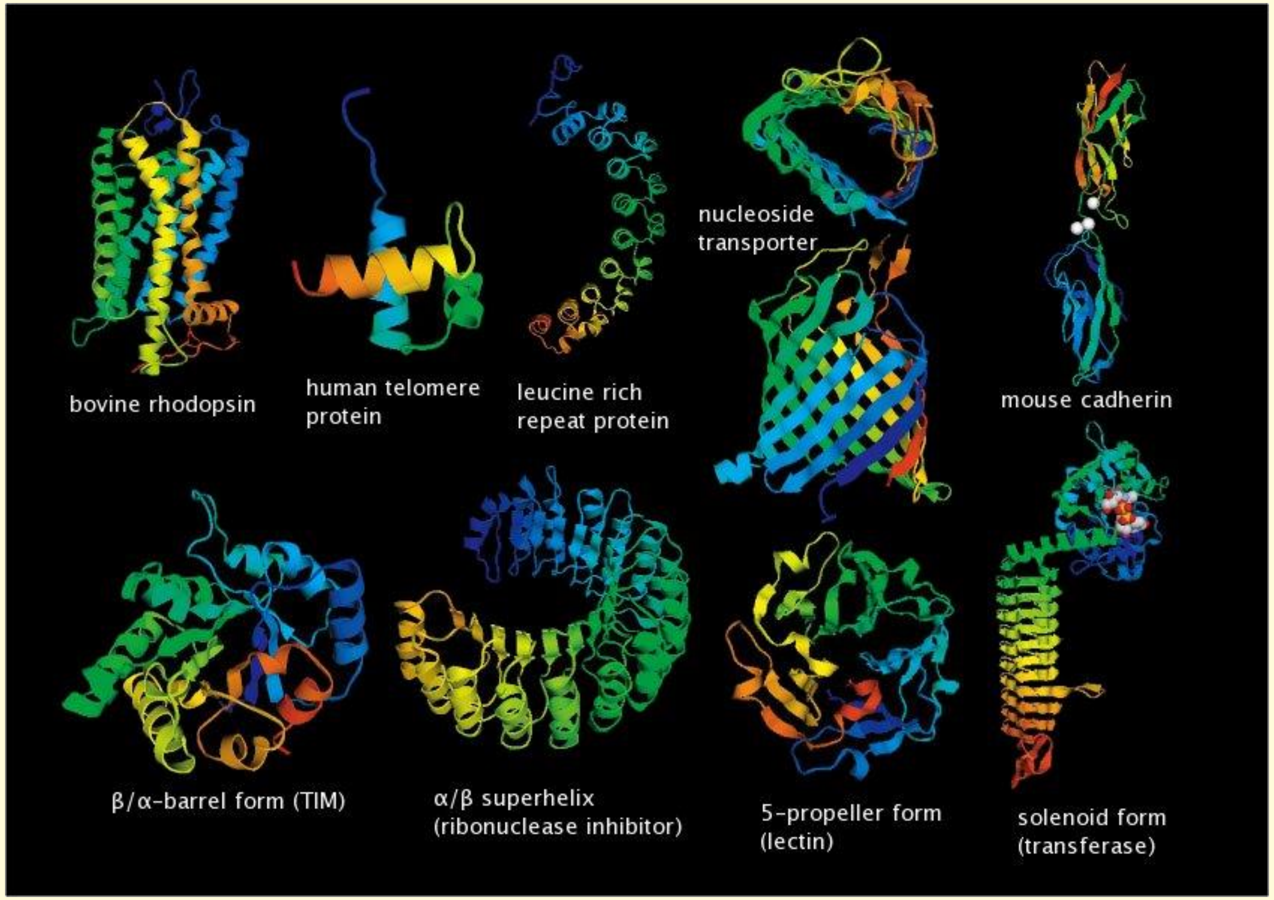


- The MAP Kinase pathway causes a cell to change behavior in response to the presence of a molecule outside the cell.
- Physical interactions between a cell-surface receptor and the signaling molecule lead to changes in the physical interactions that the receptor makes inside the cell.
- These changes set off a cascade of chemical reactions involving phosphorylation of proteins, which changes their behavior.

Image from: https://en.wikipedia.org/wiki/MAPK/ERK_pathway

How are proteins made?

- Functional proteins are large, intricately folded molecules whose physical structure is critical to function.
- The study of these shapes is known as **structural biology.**

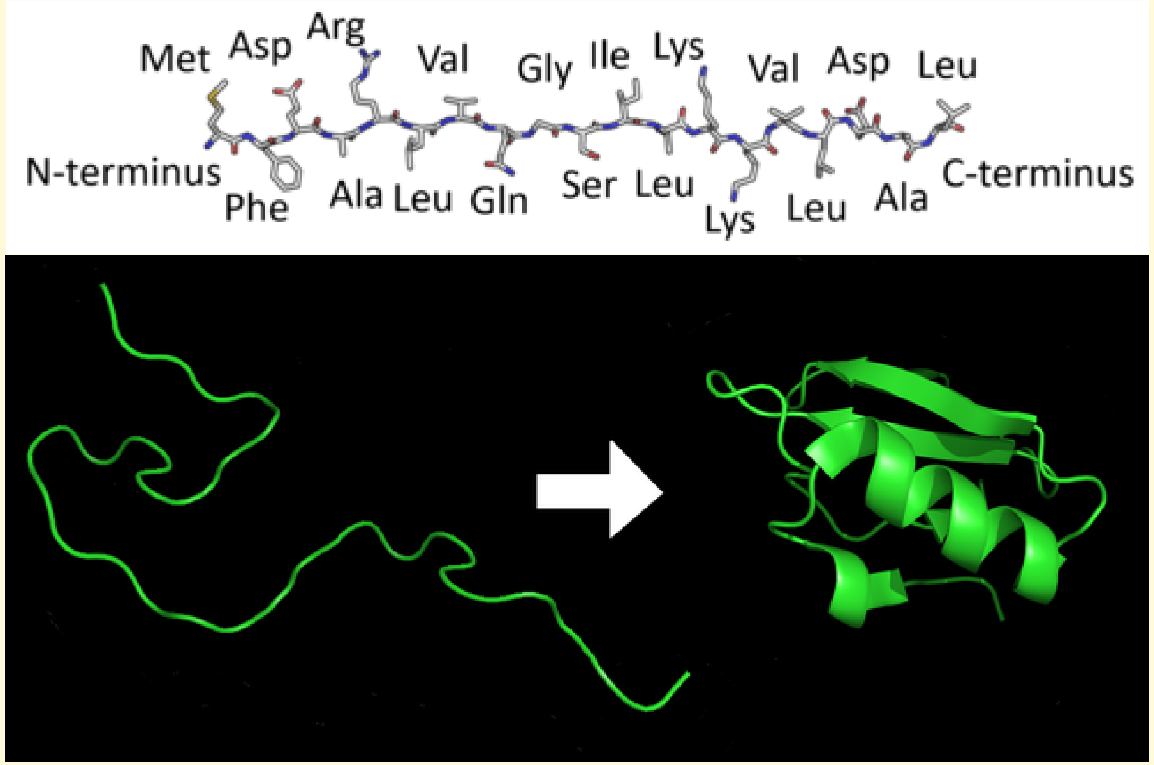


bovine rhodopsin
human telomere protein
leucine rich repeat protein
nucleoside transporter
mouse cadherin
β/α-barrel form (TIM)
α/β superhelix (ribonuclease inhibitor)
5-propeller form (lectin)
solenoid form (transferase)

Image from: http://pdbj.org/eprots/index_en.cgi?structure%20and%20function%3AStructure%20of%20protein

How are proteins made?

- The underlying structure of proteins is a linear polymer of amino acids, or polypeptide.
- The process of adopting the final 3-D structure necessary for function is known as protein folding.



Images from: https://en.wikipedia.org/wiki/Protein_primary_structure

https://en.wikipedia.org/wiki/Protein_folding

How are proteins made?

- The polypeptide is made based on information stored in DNA!
- The segment of DNA that encodes one protein is called a gene.
- There are two key steps in this process: transcription and translation.

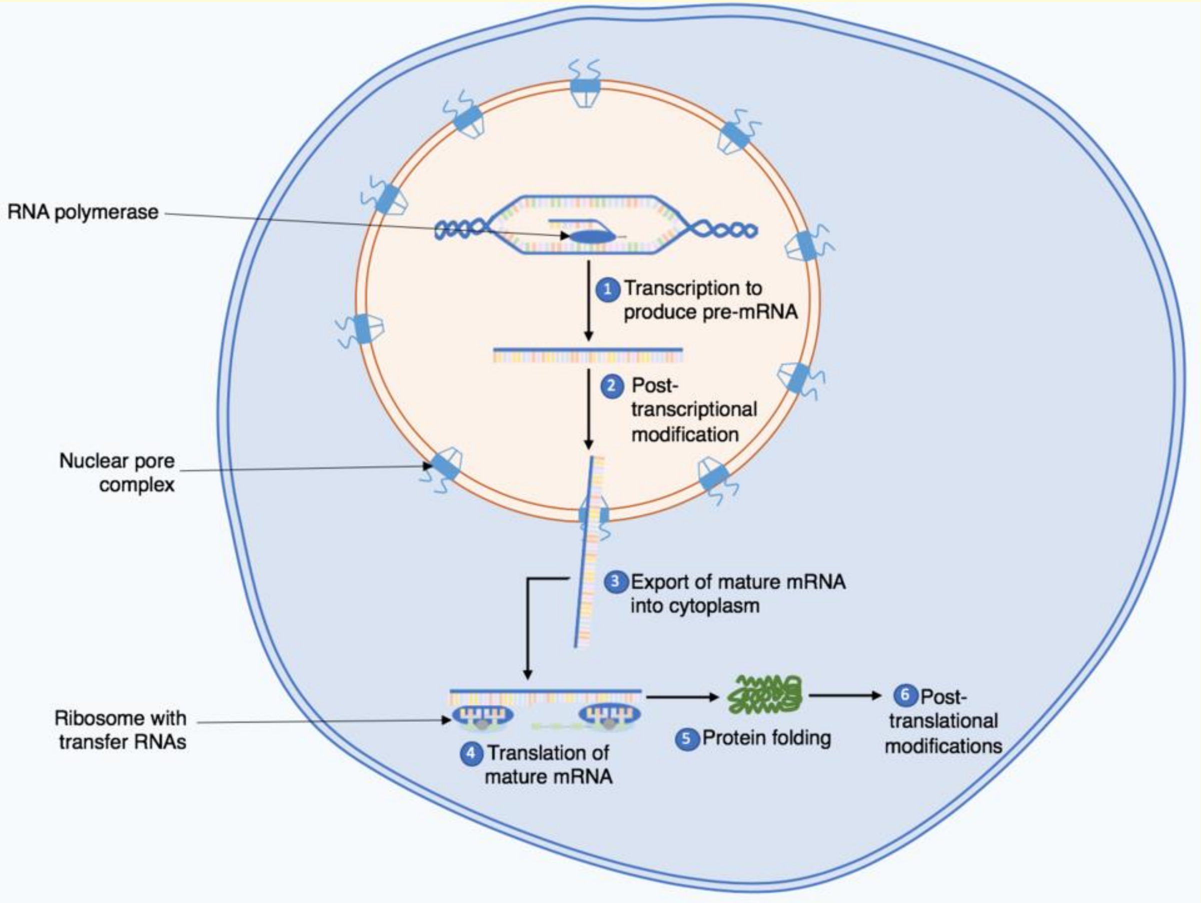

RNA polymerase
1 Transcription to produce pre-mRNA
2 Post-transcriptional modification
Nuclear pore complex
3 Export of mature mRNA into cytoplasm
Ribosome with transfer RNAs
4 Translation of mature mRNA
5 Protein folding
6 Post-translational modifications

Image from: https://en.wikipedia.org/wiki/Protein_biosynthesis

How are proteins made?

- In the process of **transcription**, a section of DNA that encodes a protein is used as a template to make messenger RNA (mRNA) that also encodes the protein.
- The DNA is unchanged in this process.

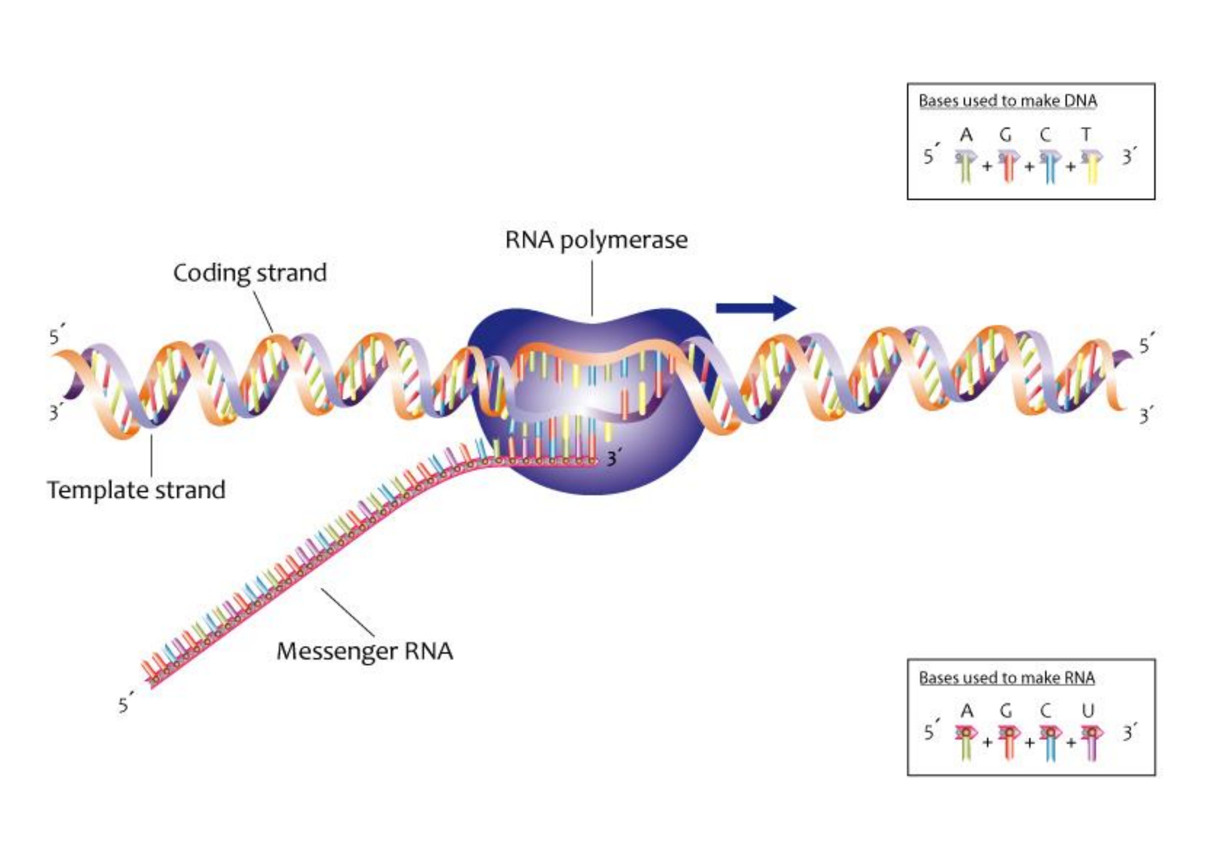


Image from: https://en.wikipedia.org/wiki/Bacterial_transcription

How are proteins made?

- In the process of **translation**, a messenger RNA (mRNA) is used as a template to synthesize the polypeptide (which then folds into the final protein structure).
- The mRNA is unchanged in this process.

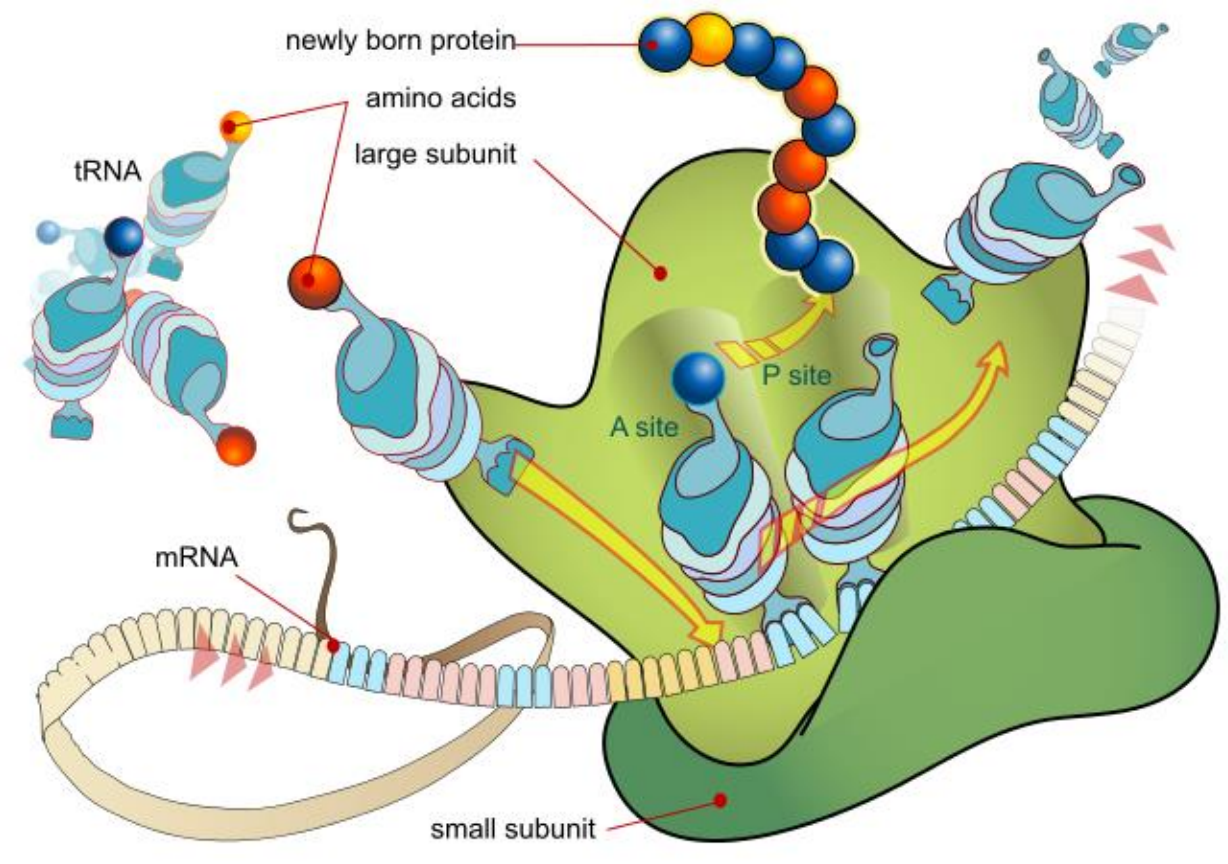


Image from: https://en.wikipedia.org/wiki/Translation_(biology)

The Central Dogma

- The concept that proteins are encoded by DNA and produced through an mRNA intermediate is combined with the concept of DNA replication (to form an identical copy) in what we call The Central Dogma of Molecular Biology.

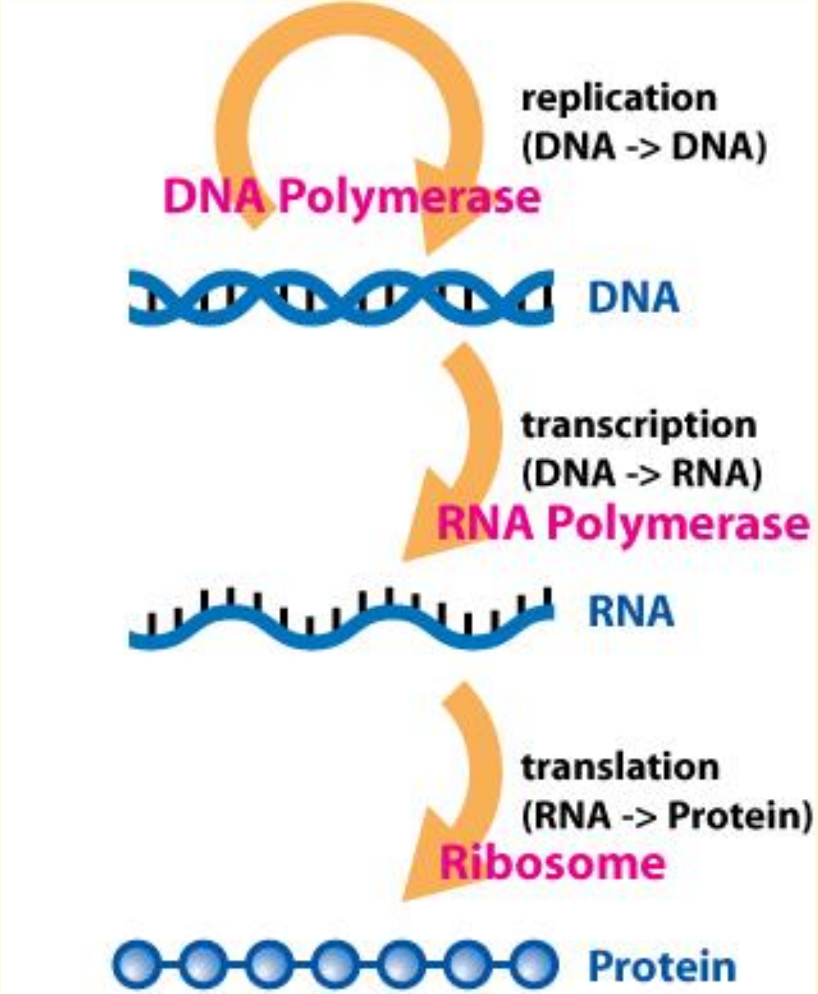


Image from: https://en.wikipedia.org/wiki/Central_dogma_of_molecular_biology

Gene Regulation

- A cell’s DNA typically contains the code for many more proteins than are present in the cell at any given time; and various proteins can present in very different amounts.
- The human genome encodes for perhaps 20,000 different proteins, but only about 10,000 of these appear to be expressed in a “typical” human cell.
- Studies in yeast (S. cerevisiae) have found that some proteins are present in numbers greater than 500,000 per cell, while other are present in numbers less then 10 per cell!
- How is this controlled?

Transcriptional Regulation

- The transcription of a gene requires the assembly of a complex of proteins that help DNA polymerase get started.
- This process can be controlled!
- Translation is much less regulated.

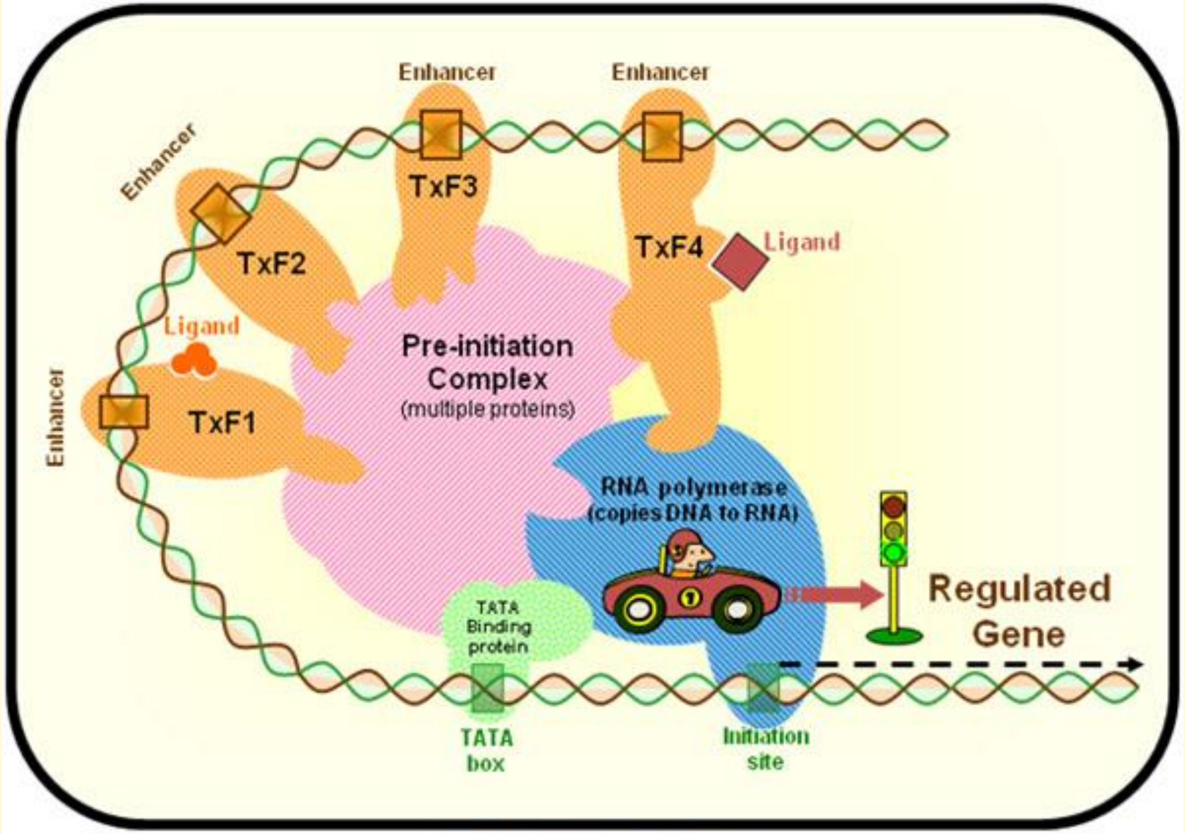


Figure 2 - Transcription Factors

Image from: http://www.monsanto.com/products/pages/transcription-factors-in-plants.aspx

Transcription Factors

- Proteins known as transcription factors bind to DNA and/or proteins involved in transcriptional initiation.
- They often act cooperatively, with more than one molecule needed to have a full effect.

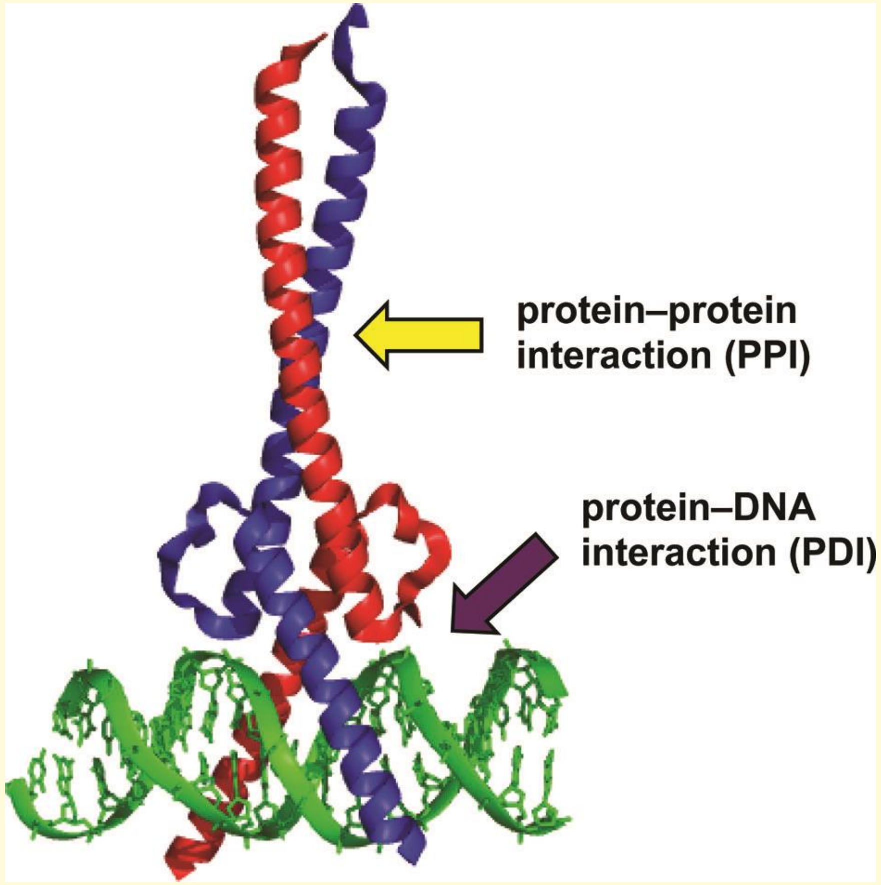
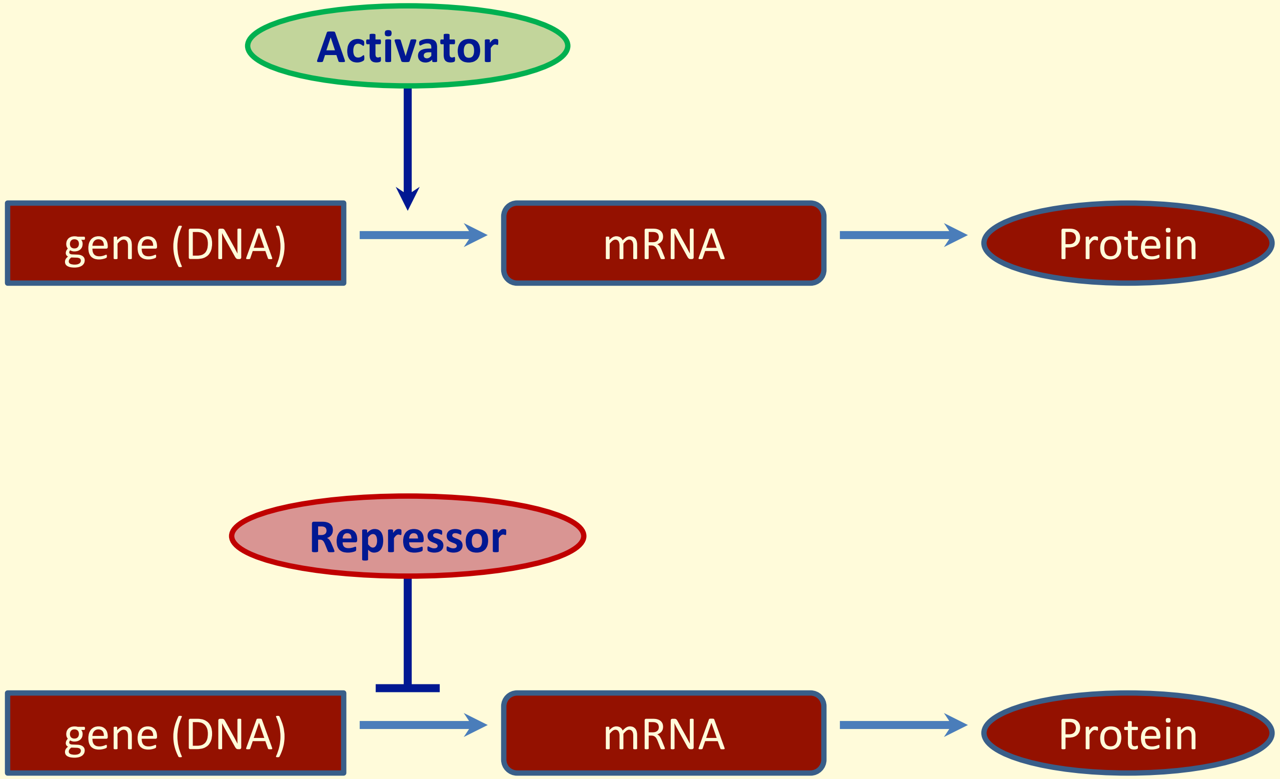

protein–protein interaction (PPI)
protein–DNA interaction (PDI)

Image from: http://blogs.rsc.org/md/2012/03/13/review-article-on-small-molecule-inhibitors-of-dimeric-transcription-factors-a-challenging-area/

Transcription Factors

- Activators, or enhancers, make transcription occur FASTER.
- Repressors make transcription occur SLOWER.
- It is not uncommon for the same transcription factor to act as an activator in some contexts and as a repressor in others.



Summary

- Proteins carry out the bulk of cellular activities, but their sequences (and thus structures and functions) are encoded in DNA (genes).
- Protein synthesis involves two non-destructive steps, transcription and translation, which are part of the Central Dogma.
- The primary source of the regulation of gene expression occurs at the transcriptional step, with protein transcription factors playing an essential role.

Mini-Lecture 4

Mass-Action Kinetic Models of Metabolism and Signal Transduction

Building a mathematical model of biochemical processes.

- What are the fundamental processes involved?
  - Physical interactions of proteins (“binding”).
  - Enzyme-catalyzed reactions.
  - RNA synthesis (transcription).
  - Protein synthesis (translation).
  - Movement of proteins and other molecules.
- What type of mathematical model might we use?
- What approximations can we make?

Mass-Action Kinetics: A 170-year-old modeling framework.

- Many of the fundamental biochemical processes are chemical reactions.
- The rate of chemical reactions is dependent on the concentrations of the molecular reactants.
- The rate of a chemical reaction can be expressed as a rate of change in concentration over over time.

Reaction	Rate Law
$X \rightarrow Y$	$\frac{d[Y]}{dt} = \frac{-d[X]}{dt} = k_1[X]$
$A + B \rightarrow AB$	$\frac{d[AB]}{dt} = \frac{-d[A]}{dt} = \frac{-d[B]}{dt} = +k_2[A][B]$

Mass-Action Kinetics: A 170-year-old modeling framework.

- When we have multiple reactions, we can write a rate law for each reaction.
- The rates of change of concentration for each molecule are given by adding up the contributions for each reaction.
- The resulting model is a **system of ordinary differential equations (ODEs).**

Reactions	Rate Laws	Rates of Change
$A + B \rightarrow AB$	$v_f = k_f[A][B]$	$\frac{d[A]}{dt} = -v_f + v_r$
$AB \rightarrow A + B$	$v_r = k_r[AB]$	$\frac{d[B]}{dt} = -v_f + v_r$
		$\frac{d[AB]}{dt} = +v_f - v_r$

Mass-Action Enzyme Kinetics.

- We can apply these rules to biochemical systems simply by breaking own each process into its fundamental steps.

$$E + S \overset{k_1}{\rightarrow} ES$$
$$ES \xrightarrow{k_{-1}} E + S$$
$$ES \overset{k_2}{\rightarrow} EP$$
$$EP \overset{k_3}{\rightarrow} E + P$$

$$\frac{d[E]}{dt} = -k_1[E][S] + k_{-1}[ES] + k_3[EP]$$
$$\frac{d[ES]}{dt} = +k_1[E][S] - k_{-1}[ES] - k_2[ES]$$
$$\frac{d[EP]}{dt} = +k_2[ES] - k_3[EP]$$
$$\frac{d[S]}{dt} = -k_1[E][S] + k_{-1}[ES]$$
$$\frac{d[P]}{dt} = +k_3[EP]$$

Simplified Enzyme Kinetics.

- Biochemistry is complicated; even simple systems lead to systems of many equations with many parameters.
- Simplified rate rules can be derived that work well under certain conditions.
- Michaelis-Menten kinetics (1913):

$$E + S \rightarrow E + P$$

$$\frac{d[P]}{dt} = \frac{-d[S]}{dt} = \frac{k_{cat}[E][S]}{K_m + [S]}$$

Maud Menten, MD, PhD, 1879-1960: https://en.wikipedia.org/wiki/Maud_Menten

Simplified Enzyme Kinetics.

- Rate laws can be derived for enzymes with more complicated behavior as well.
  - Cooperative binding: $$v = \frac{k_{cat}[E][S]^h}{\left(K_{1/2}\right)^h + [S]^h}$$
  - Competitive inhibition: $$v = \frac{k_{cat}[E][S]}{K_m\left(1+\frac{[I]}{K_I}\right)+[S]}$$
  - Non-competitive inhibition: $$v = \frac{k_{cat}[E][S]}{\left(1+\frac{[I]}{K_I}\right)(K_m+[S])}$$
  - Obligate activation: $$v = \frac{k_{cat}[E][S]}{K_m\left(1+\frac{K_A}{[A]}\right)+[S]}$$

Modeling Metabolic Networks.

We build models for metabolic networks just by:

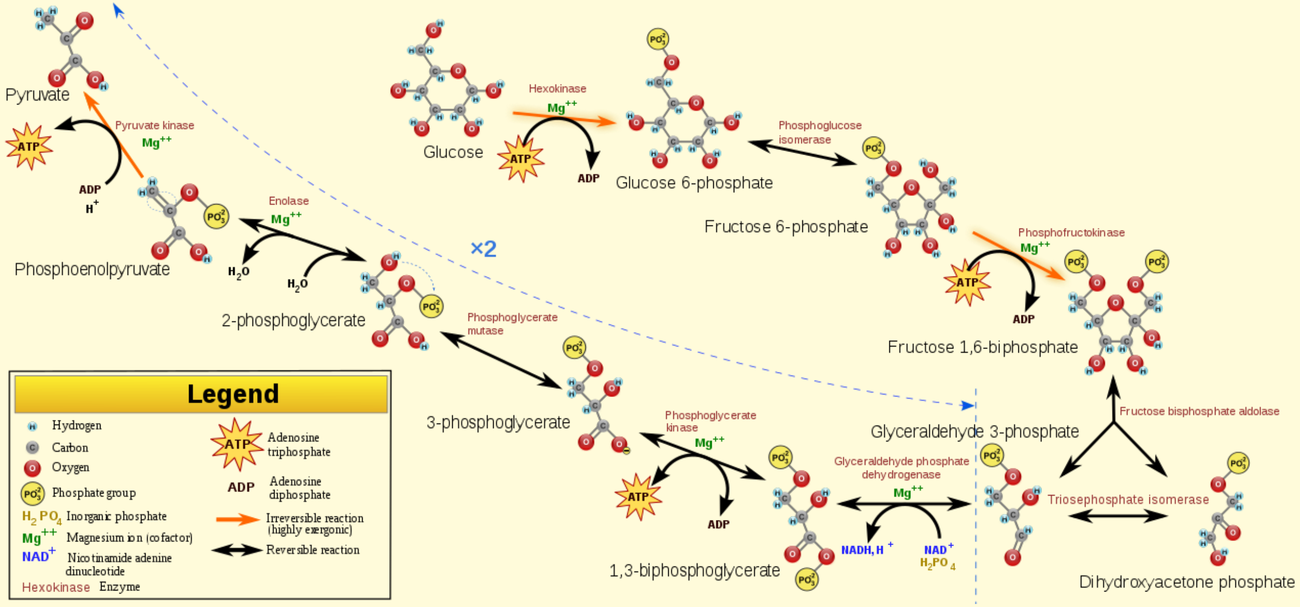
- Defining each reaction, identifying substrates (reactants), products, enzymes, etc.
- Writing a rate law for each reaction.
- Writing an expression for the rate of change of each molecular species as a sum of rate laws:
  - Positive if the species is a product.
  - Negative if the species is a reactant.

Modeling Metabolic Networks.

Glycolysis:

1. $Glu + ATP \xrightarrow{HK} G6P + ADP$
2. $G6P \xrightarrow{PGI} F6P$
3. $F6P + ATP \xrightarrow{PFK} F16P + ADP$
4. $F16P \xrightarrow{FBA} G3P + DHAP$
5. $DHAP \xrightarrow{TPI} G3P$
6. $G3P + NAD^+ \xrightarrow{FBA} 13BPG + NADH_2$

$$\vdots$$



Modeling Metabolic Networks.

Glycolysis:

1. $Glu + ATP \xrightarrow{HK} G6P + ADP$
2. $G6P \xrightarrow{PGI} F6P$
3. $F6P + ATP \xrightarrow{PFK} F16P + ADP$
4. $F16P \xrightarrow{FBA} G3P + DHAP$
5. $DHAP \xrightarrow{TPI} G3P$
6. $G3P + NAD^+ \xrightarrow{FBA} 13BPG + NADH_2$

$$\vdots$$

$$v_1 = \frac{k_{cat,1}K_{a,1}[HK][ATP][Glu]}{K_{M,1}+[Glu]}$$

$$v_2 = \frac{k_{cat,2}[PGI][G6P]}{K_{M,2}+[G6P]}$$

$$v_3 = \frac{k_{cat,3}K_{a,3}[PFK][ATP][F6P]}{K_{M,3}+[F6P]}$$

$$v_4 = \frac{k_{cat,4}[FBA][F16P]}{K_{M,4}+[F16P]}$$

$$v_5 = \frac{k_{cat,5}[TPI][DHAP]}{K_{M,5}+[DHAP]}$$

$$v_6 = \frac{k_{cat,6}[FBA][G3P][NAD^+]}{K_{A,6}[NAD^+]+K_{B,6}[G3P]+[G3P][NAD^+]}$$

$$\vdots$$

Modeling Metabolic Networks.

Glycolysis:

1. $Glu + ATP \xrightarrow{HK} G6P + ADP$
2. $G6P \xrightarrow{PGI} F6P$
3. $F6P + ATP \xrightarrow{PFK} F16P + ADP$
4. $F16P \xrightarrow{FBA} G3P + DHAP$
5. $DHAP \xrightarrow{TPI} G3P$
6. $G3P + NAD^{+} \xrightarrow{FBA} 13BPG + NADH_2$

$\vdots$

$$\frac{d[Glu]}{dt} = -v_1$$

$$\frac{d[G6P]}{dt} = +v_1 - v_2$$

$$\frac{d[ATP]}{dt} = -v_1 - v_3 - \cdots$$

$$\frac{d[ADP]}{dt} = v_1 + v_3 + \cdots$$

$$\frac{d[F6P]}{dt} = +v_2 - v_3$$

$$\frac{d[F16P]}{dt} = v_3 - v_4$$

$$\vdots$$

Summary

- Mass-action kinetics provides a framework for biochemical kinetics.
- Approximate forms can be derived for various types of enzyme activity.
- This approach gives a model in the form of a system of ordinary differential equations: one for each molecular species in the system.

Mini-Lecture 5

Modeling Gene Expression and Regulation

Gene expression is mechanistically complicated

- Both transcription and translation involve many individual chemical reactions.
  - Assembly of an initiation complex
    - Particularly complicated in transcription.
  - Repeated extension of a polymer.
    - Dependent on charged transfer RNA in translation.
  - Termination and release.

- Need to be inspired by Mass-Action kinetics, rather than building a mechanistically exact model.

Encapsulate the complexity of transcription/translation in a rate law.

- For each gene, we will track the concentration of two molecular species, mRNA and of protein:
  - We only track the complete, functional mRNA and protein; no intermediate steps.
- For each molecular species, we will define a rate equation with two terms, synthesis and degradation:

$$\frac{d[X]}{dt} = +SynthesisRate - DegradationRate$$

- This basic framework can be extended to include a lot of detail simply by creating more complicated rate rules:
  - Our models will be as simple as possible.

Degradation (both protein and mRNA)

- Degradation is an essential part of the model; without it all concentration could only increase!
- The simplest model for degradation is a unimolecular process, much like radioactive decay:

$$X \rightarrow \emptyset$$

$$\upsilon_{degradation} = \chi_X[X]$$

- Note: assuming a simple model does not imply that we think the process follows a simple mechanism.
  - We begin with a simple model and see what we can learn; adding complexity is usually easier than removing it!

Protein Synthesis (Translation)

- Translational regulation is generally viewed as being less regulated than transcription.
- The simplest model for translation assumes that the rate of protein synthesis is directly proportional to the amount of mRNA that is present:

$$X_{mRNA} \rightarrow X_{mRNA} + X_{prot}$$

$$\upsilon_{synthesis} = \omega_X [X_{mRNA}]$$

- Again: There are lots of “incorrect” assumptions in this model, but that does not mean it is a bad starting point!

mRNA Synthesis (Transcription)

- We know that transcriptional regulation is essential for many processes affecting cell-fate, and so it can not reasonably be ignored.
- As there are different types of regulation, and each will need to be considered individually:
  - Constituitively expressed.
  - Obligate activator.
  - Repressor.
  - Many more complicated mechanisms.

Constituitively active transcription

- Some genes are expressed at approximately the same level under almost all conditions.
- This can be modeled by a constant rate of mRNA synthesis.

$$X_{gene} \rightarrow X_{gene} + X_{mRNA}$$

$$v_{synthesis} = \mu_X$$

- Constituitive activity may well actually result from elaborate control mechanisms, but if those mechanisms are not the object of study, it may be reasonable to ignore them.

Transcription with an obligate activator

- Some genes absolutely require the presence of a particular transcription factor in order to be expressed:

$$X_{gene} \cdot TF \rightarrow X_{gene} \cdot TF + X_{mRNA}$$
$$v_{synthesis}([TF] = 0) = 0$$

- When the transcription factor is present in excess, some maximal synthesis rate is reached:

$$v_{synthesis}(\lim([TF] \rightarrow \infty)) = \mu_X$$

- In between these extremes we model the rate of mRNA synthesis using a function that varies smoothly from zero to the maximal rate:

$$v_{synthesis} = \xi_X([TF])\mu_X$$

Fraction bound of a transcription factor

- We may think of ξ_X as representing the fraction of time that the transcription factor is bound to the gene (specifically to the promoter).
  - The overall synthesis rate is an average over the time that the TF is not bound ($\nu = 0$) and the time that it is bound ($\nu = \mu_X$).
- A commonly used function to describe a fraction-bound is the Hill Equation, which has all the desired properties:

$$\xi_X([TF]) = \frac{[TF]^h}{(K_{1/2})^h + [TF]^h}$$

- For the obligate activator, the overall rate of mRNA synthesis is then:

$$v_{synthesis} = \left(\frac{[TF]^h}{(K_{1/2})^h + [TF]^h}\right)\mu_X$$

A note on the Hill Equation

$$f([X]) = \frac{[X]^h}{(K_{1/2})^h + [X]^h}$$

- The Hill Equation can be derived from first principles (although doing so makes many assumptions.)
- The derivation suggests a meaning for the Hill coefficient, h; it is a measure of the degree of "cooperativity" of binding.
  - It can loosely be related to the number of molecules that must act together to have an effect.
- $K_{1/2}$ is the concentration of half occupancy.

Repressed transcription

- In the simplest context, a repressor works in exactly the opposite way as an obligate activator.

$$X_{gene} \rightarrow X_{gene} + X_{mRNA}$$

$$X_{gene} \cdot TF \rightarrow No\ reaction.$$

- When the transcription factor is absent, there is some maximal synthesis rate, and when it is present in excess, the synthesis rate is zero:

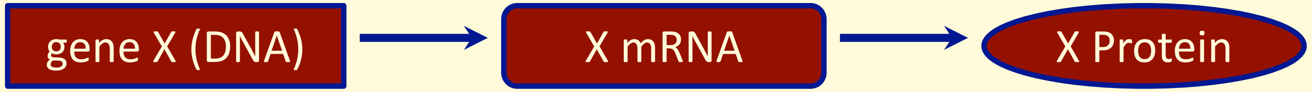
$$v_{synthesis}([TF] = 0) = \mu_X \qquad v_{synthesis}(\lim([TF] \rightarrow \infty)) = 0$$

- The fraction-bound can be used to give an overall synthesis rate:

$$v_{synthesis} = \big(1 - \xi_X([TF])\big)\mu_X$$

Putting it all together: A constituitively expressed gene.

- We now have all the tools we need to write a model for the expression of a gene.



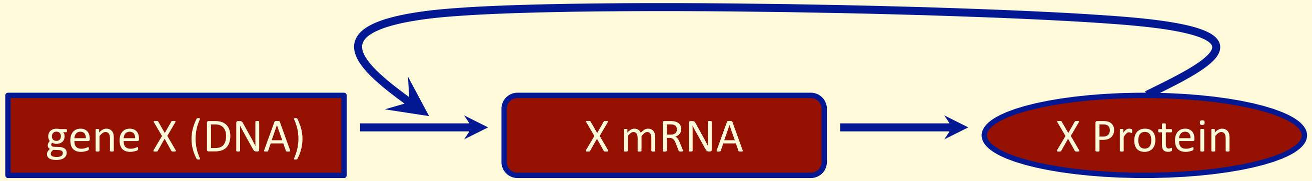
Protein: $$\frac{d[X_{prot}]}{dt} = \omega_X [X_{mRNA}] - \chi_{X_{prot}} [X_{prot}]$$

mRNA: $$\frac{d[X_{mRNA}]}{dt} = \mu_X - \chi_{X_{mRNA}} [X_{mRNA}]$$

Synthesis **Degradation**

Putting it all together: An auto-regulated gene.

- Some transcription factors can affect their own expression! Consider an auto-activator:



$$
\begin{aligned}
\frac{d[X_{prot}]}{dt} &= \omega_X [X_{mRNA}] - \chi_{X_{prot}} [X_{prot}] \\
\frac{d[X_{mRNA}]}{dt} &= \left(\frac{[X_{prot}]}{K_{1/2} + [X_{prot}]} \right) \mu_X - \chi_{X_{mRNA}} [X_{mRNA}]
\end{aligned}
$$

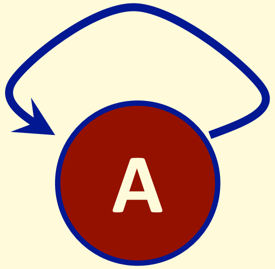
Synthesis

Degradation

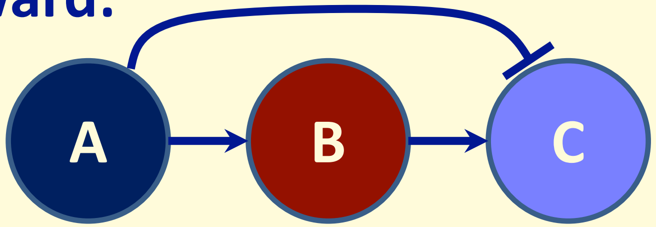
Building models of gene-regulatory networks.

- Gene regulatory network arise when transcription factors affect each other's expression.

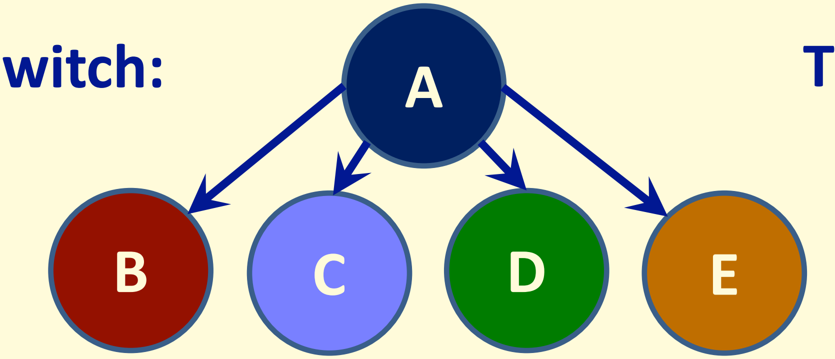
Auto-enhancer:



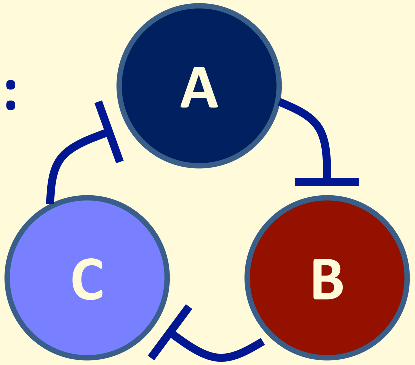
Negative feedforward:



Master switch:



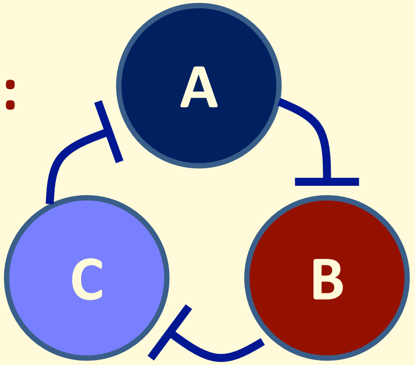
Triple-negative feedback:



Building models of gene-regulatory networks.

- We write a pair of equations for each gene, using a common framework for each one:

Triple-negative feedback:



$$\frac{d[A_{prot}]}{dt} = \omega_A[A_{mRNA}] - \chi_{A_{prot}}[A_{prot}]$$

$$\frac{d[B_{prot}]}{dt} = \omega_B[B_{mRNA}] - \chi_{B_{prot}}[B_{prot}]$$

$$\frac{d[C_{prot}]}{dt} = \omega_C[C_{mRNA}] - \chi_{C_{prot}}[C_{prot}]$$

$$\frac{d[A_{mRNA}]}{dt} = \left(1 - \frac{[C_{prot}]}{K_{C\to A}+[C_{prot}]}\right)\mu_A - \chi_{A_{mRNA}}[A_{mRNA}]$$

$$\frac{d[B_{mRNA}]}{dt} = \left(1 - \frac{[A_{prot}]}{K_{A\to B}+[A_{prot}]}\right)\mu_B - \chi_{B_{mRNA}}[B_{mRNA}]$$

$$\frac{d[C_{mRNA}]}{dt} = \left(1 - \frac{[B_{prot}]}{K_{B\to C}+[B_{prot}]}\right)\mu_C - \chi_{C_{mRNA}}[C_{mRNA}]$$

Summary

- Transcription, translation and macromolecular degradation can all be modeled using equations inspired by mass-action kinetics.
- In simple models, regulation is limited to transcription.
- Models of complex gene-regulatory networks can be constructed simply by reproducing a common equation structure for each gene.

Mini-Lecture 6

Simulating systems of ordinary differential equations

How to we work with systems of ODEs?

- We’ve developed a modeling framework for cellular physiological processes from enzyme catalysis through transcription and translation.
- Systems of differential equations, with a rate of change for each molecular species.
- Is the rate of change of concentrations what we really care about?

Concentrations matter more!

- If individual proteins carry out function at the molecular level, the total number (amount) of each protein must affect its function at at a cellular level.
- How do we get total concentrations from rates of change?
- Compare to knowing velocity (speed) and wanting to know position (distance): a problem in integration!

$$v = \frac{dx(t)}{dt} \rightarrow x(t) = x(0) + \int_0^t v(t')dt'$$

- Note that we need an initial condition to solve the problem.

Integrating a system of ODEs.

- We have a general expression for the rate of change of each variable, in terms of the same variables.

$$\frac{dx(t)}{dt} = f\big(x(t), y(t)\big), \qquad \frac{dy(t)}{dt} = g\big(x(t), y(t)\big)$$

- We want to solve multiple integrals:

$$x(t) = x(0) + \int_0^t f(t')dt', \qquad y(t) = y(0) + \int_0^t g(t')dt'$$

- The derivatives are not *explicit* functions of time, but rather *implicit* functions.
- The time dependence is imbedded in the time variation of the variables, which is exactly what we are trying to find.

A solution: Small time steps.

- For a small time interval, we can use approximations:

$$x(t_0+\Delta t) = x(t_0) + \int_{t_0}^{t_0+\Delta t} f(t')dt', \qquad y(t_0+\Delta t) = y(t_0) + \int_{t_0}^{t_0+\Delta t} g(t')dt'$$

$$\frac{dx(t)}{dt} = f\big(x(t), y(t)\big), \qquad \frac{dy(t)}{dt} = g\big(x(t), y(t)\big)$$

- Choose $\tilde{f}(t) \approx f(t)$ and $\tilde{g}(t) \approx g(t)$ then:

$$x(t_0+\Delta t) \approx x(t_0) + \int_{t_0}^{t_0+\Delta t} \tilde{f}(t')dt', \qquad y(t_0+\Delta t) \approx y(t_0) + \int_{t_0}^{t_0+\Delta t} \tilde{g}(t')dt'$$

- The approximate functions are *explicit* functions of time, so we can solve the integral directly.

The Forward Euler approximation

$$x(t_0+\Delta t) \approx x(t_0) + \int_{t_0}^{t_0+\Delta t} \tilde{f}(t')dt', \qquad y(t_0+\Delta t) \approx y(t_0) + \int_{t_0}^{t_0+\Delta t} \tilde{g}(t')dt'$$

- Perhaps the simplest approximation is a constant function:

$$\tilde{f}(t) \equiv f(t_0) = \frac{dx(t_0)}{dt}, \qquad \tilde{g}(t) \equiv g(t_0) = \frac{dy(t_0)}{dt}$$

- The integral becomes a simple product:

$$\int_{t_0}^{t_0+\Delta t} \tilde{f}(t')dt' = f(t_0)(t+\Delta t) - f(t_0)(t) = f(t_0)\Delta t$$

- The Forward Euler approximation:

$$x(t_0+\Delta t) \approx x(t_0) + \left(\frac{dx(t_0)}{dt}\right)\Delta t, \qquad y(t_0+\Delta t) \approx y(t_0) + \left(\frac{dy(t_0)}{dt}\right)\Delta t$$

A simple approximation

- For a small time interval, we can use approximations:

$$x(t+\Delta t) = x(t) + \int_{t}^{t+\Delta t} f(t')dt', \qquad y(t+\Delta t) = y(t) + \int_{0}^{t+\Delta t} g(t')dt'$$

$$\frac{dx(t)}{dt} = f\big(x(t), y(t)\big), \qquad \frac{dy(t)}{dt} = g\big(x(t), y(t)\big)$$

- Choose $\tilde{f}(t) \approx f(t)$ and $\tilde{g}(t) \approx g(t)$ then:

$$x(t+\Delta t) \approx x(t) + \int_{t}^{t+\Delta t} \tilde{f}(t')dt', \qquad y(t+\Delta t) \approx y(t) + \int_{0}^{t+\Delta t} \tilde{g}(t')dt'$$

- The approximate functions are *explicit* functions of time, so we can solve the integral directly.

An iterative approach.

- To solve the full problem over a longer period of time, we just apply the same update rule iteratively:

$$
\begin{array}{rcl}
x(\Delta t) & = & x(0) + f\big(x(0), y(0)\big)\Delta t \\
x(2\Delta t) & = & x(\Delta t) + f\big(x(\Delta t), y(\Delta t)\big)\Delta t \\
x(3\Delta t) & = & x(2\Delta t) + f\big(x(2\Delta t), y(2\Delta t)\big)\Delta t \\
x(4\Delta t) & = & x(3\Delta t) + f\big(x(3\Delta t), y(3\Delta t)\big)\Delta t \\
 & \vdots &
\end{array}
\qquad
\begin{array}{rcl}
y(\Delta t) & = & y(0) + g\big(x(0), y(0)\big)\Delta t \\
y(2\Delta t) & = & x(\Delta t) + g\big(x(\Delta t), y(\Delta t)\big)\Delta t \\
y(3\Delta t) & = & x(2\Delta t) + g\big(2x(\Delta t), y(2\Delta t)\big)\Delta t \\
y(4\Delta t) & = & x(3\Delta t) + g\big(3x(\Delta t), y(3\Delta t)\big)\Delta t \\
 & \vdots &
\end{array}
$$

$$
\frac{dx(t)}{dt} = f\big(x(t), y(t)\big), \qquad \frac{dy(t)}{dt} = g\big(x(t), y(t)\big)
$$

- We need to provide initial conditions; the algorithm gives us every other value.
- ***Important: We must recompute the derivatives every step!***

Forward Euler in Matlab.

- Iterative algorithms are ideally suited for computational solution.
- Consider the system of ODEs defined by:

$$\frac{dx}{dt} = x - 5y \qquad \frac{dy}{dt} = y + 5x$$

- Can we integrate this system for 5 seconds, beginning at $\big(x(0), y(0)\big) = (5,5)$?

Forward Euler in Matlab.

$$\frac{dx(t)}{dt} = x(t) - 5y(t)$$

$$\frac{dy(t)}{dt} = y(t) + 5x(t)$$

$$(x(0), y(0)) = (5, 5)$$

Integrate for 5 seconds.

```
% 1. Initialization of variables
% ******************************
% Set number of steps and step size
% Initialize arrays for all variables (including time)
% Set boundary conditions

% 2. Integration loop
% *******************
for (i=1:numsteps-1)
  % Compute derivatives for each variable
  % Update all variables
end

% 3. Visualization of results
% ***************************
% Plot each variable vs time.
% Plot the path in the plane (or space) of the variables
```

Forward Euler in Matlab.

$$\frac{dx(t)}{dt} = x(t) - 5y(t)$$

$$\frac{dy(t)}{dt} = y(t) + 5x(t)$$

$$(x(0), y(0)) = (5, 5)$$

Integrate for 5 seconds.

```
% 1. Initialization of variables
% ******************************

% Set number of steps and step size
totaltime = 5;
deltat = 0.01;
numsteps = totaltime/deltat;

% Initialize arrays for all variables and for t
x = zeros(numsteps,1);
y = zeros(numsteps,1);
t = zeros(numsteps,1);

% Set boundary conditions
x(1) = 5;
y(1) = 5;
t(1) = 0;
```

Forward Euler in Matlab.

$$\frac{dx(t)}{dt} = x(t) - 5y(t)$$

$$\frac{dy(t)}{dt} = y(t) + 5x(t)$$

$$(x(0), y(0)) = (5, 5)$$

Integrate for 5 seconds.

```
% 2. Integration loop
% *******************

for (i=1:numsteps-1)

  % Compute derivatives for each variable
  dx_dt = x(i)-5*y(i);
  dy_dt = y(i)+5*x(i);

  % Update all variables
  x(i+1) = x(i) + dx_dt*deltat;
  y(i+1) = y(i) + dy_dt*deltat;
  t(i+1) = t(i) + deltat;

end
```

Forward Euler in Matlab.

$$\frac{dx(t)}{dt} = x(t) - 5y(t)$$

$$\frac{dy(t)}{dt} = y(t) + 5x(t)$$

$$(x(0), y(0)) = (5, 5)$$

Integrate for 5 seconds.

```
% 3. Visualization of results
% ***************************
% Plot each variable vs time.
figure(1)
subplot(2,1,1)
plot(t,x)
ylabel('Value of X')
subplot(2,1,2)
plot(t,y)
ylabel('Value of Y')
xlabel('Time’)

% Plot the path in the x-y plane.
figure(2)
plot(x,y)
xlabel('Value of X')
ylabel('Value of Y')
```

Notes on Forward Euler

- Forward Euler is conceptually easy and easy to implement.
- For many systems it gives reasonable answers, but only with a small enough time step!
    - With a large time step, it can be quite inaccurate.
- Many other algorithms exist using a similar concept, but a better approximation (e.g. Runge-Kutta).
    - Beyond the scope of this class.
    - Matlab has many of these as built-in functions.

Summary

- Systems of ODEs can be integrated numerically using an iterative approach.
- Numerical integration is intended for solution by computers, not by hand!
- The Forward Euler algorithm is a conceptually easy method: the derivatives are updated each time step and then used to find the values at the next point in time.
- The accuracy of these methods depends on the choice of the time step used.

Mini-Lecture 7

Numerical analysis of mass-action kinetic systems.

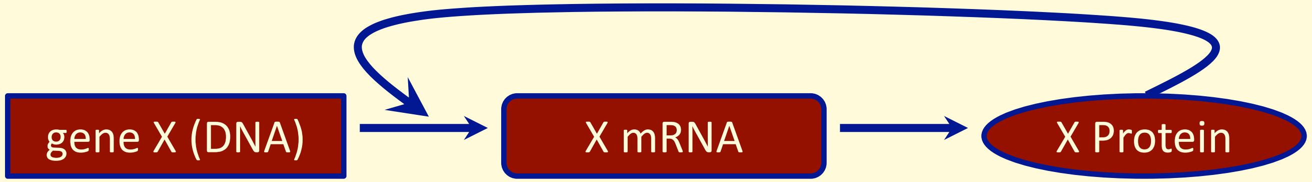
Simulating the behavior of a gene-regulatory circuit.

- We can write down a system of ODEs for a given problem.

Example: An auto-enhancing gene active as a dimer.

$$\frac{d[X_{prot}]}{dt} = \omega_X[X_{rna}] - \chi_{X_{pro}}[X_{pro}]$$

$$\frac{d[X_{rna}]}{dt} = \left(\frac{[X_{pro}]^2}{(K_{1/2})^2 + [X_{pro}]^2}\right)\mu_X - \chi_{X_{rna}}[X_{rna}]$$



Simulating the behavior of a gene-regulatory circuit.

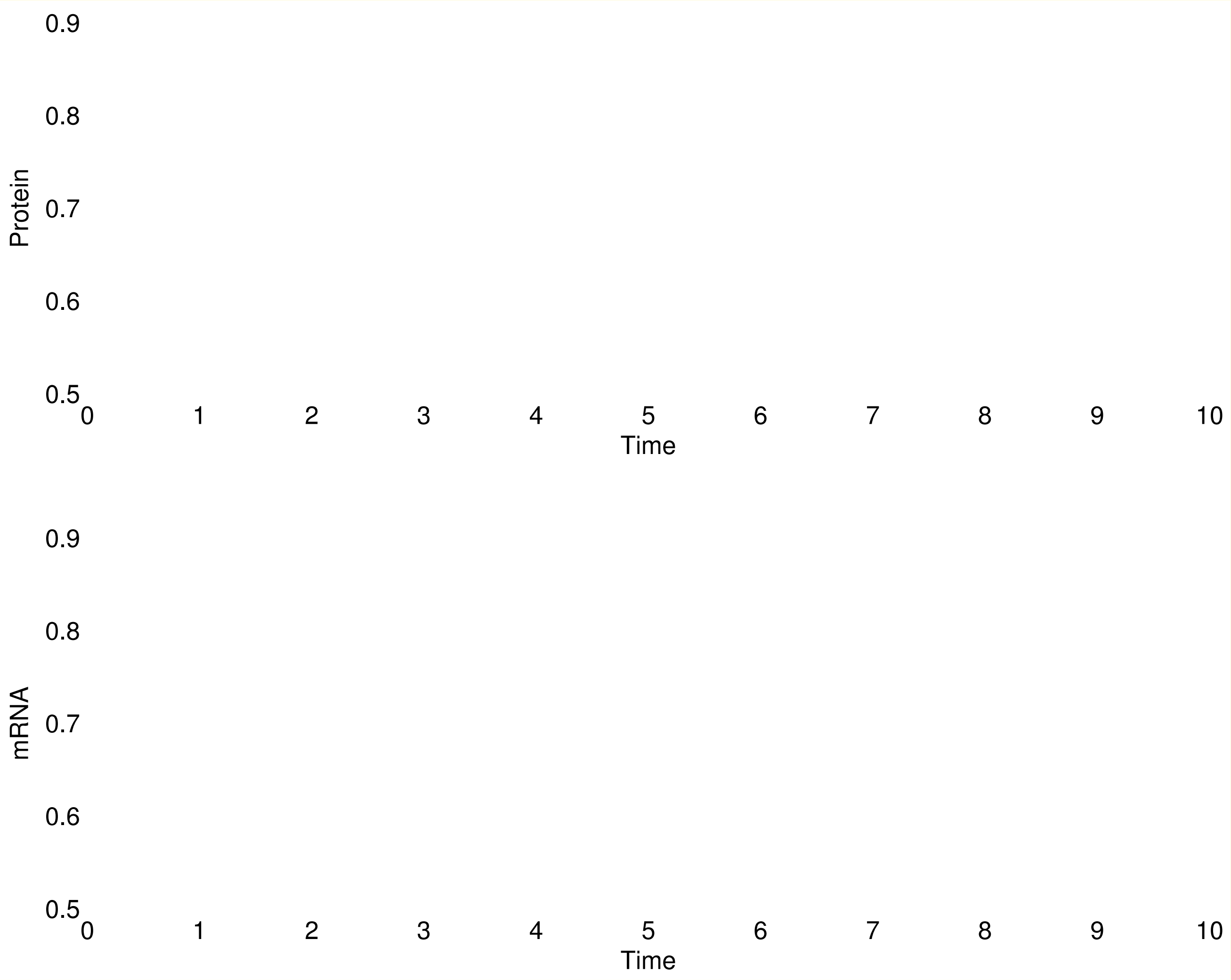
- We can simulate the system using Forward Euler
- Implementation is built around an iterative loop:

```
for (i=1:numsteps-1)
    dxpro_dt = …
    dxrna_dt = …
    xpro(i+1) = xpro(i) + dxpro_dt*deltat;
    xrna(i+1) = xrna(i) + dxrna_dt*deltat;
end
```

- What is the “right” number of steps?
- What is the “right” value of Δt?
- We cannot model a system without addressing these questions!

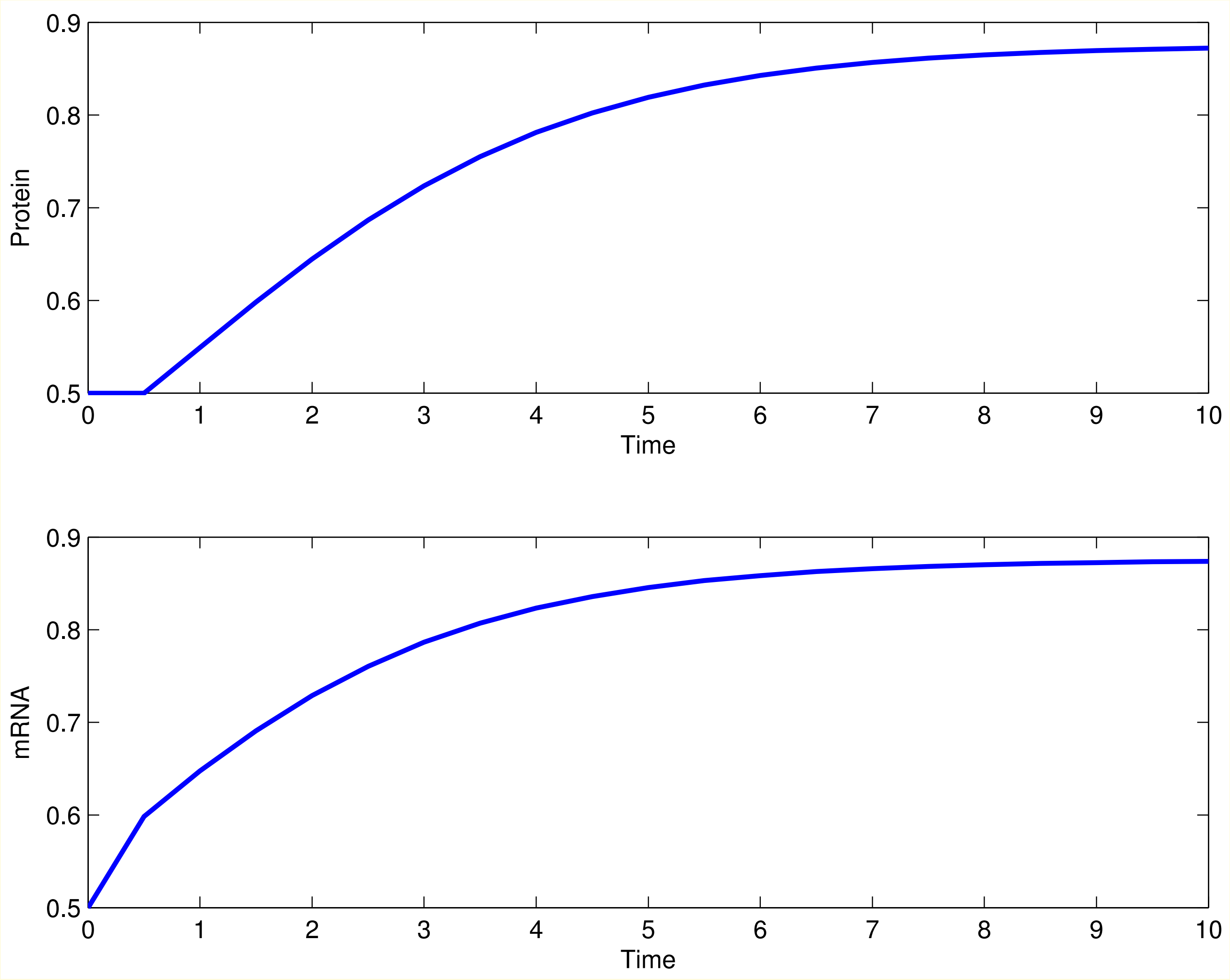
Varying the time step with constant total time.

- $t_{total} = 10\ s$
- $\Delta t = 1\ s$
- $n = \frac{t_{total}}{\Delta t} = 10$



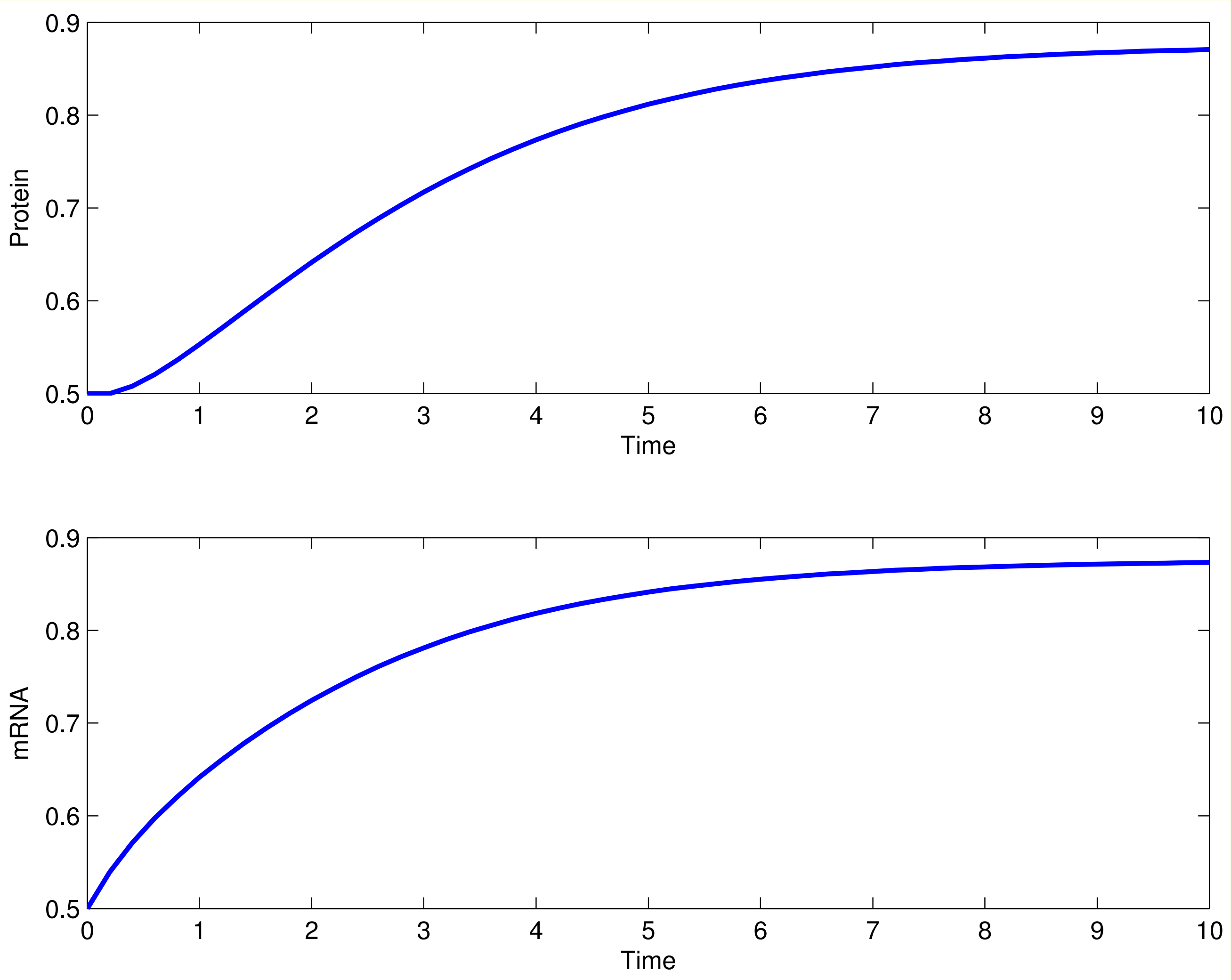
Varying the time step with constant total time.

- $t_{total} = 10\ s$
- $\Delta t = 0.5\ s$
- $n = \frac{t_{total}}{\Delta t} = 20$



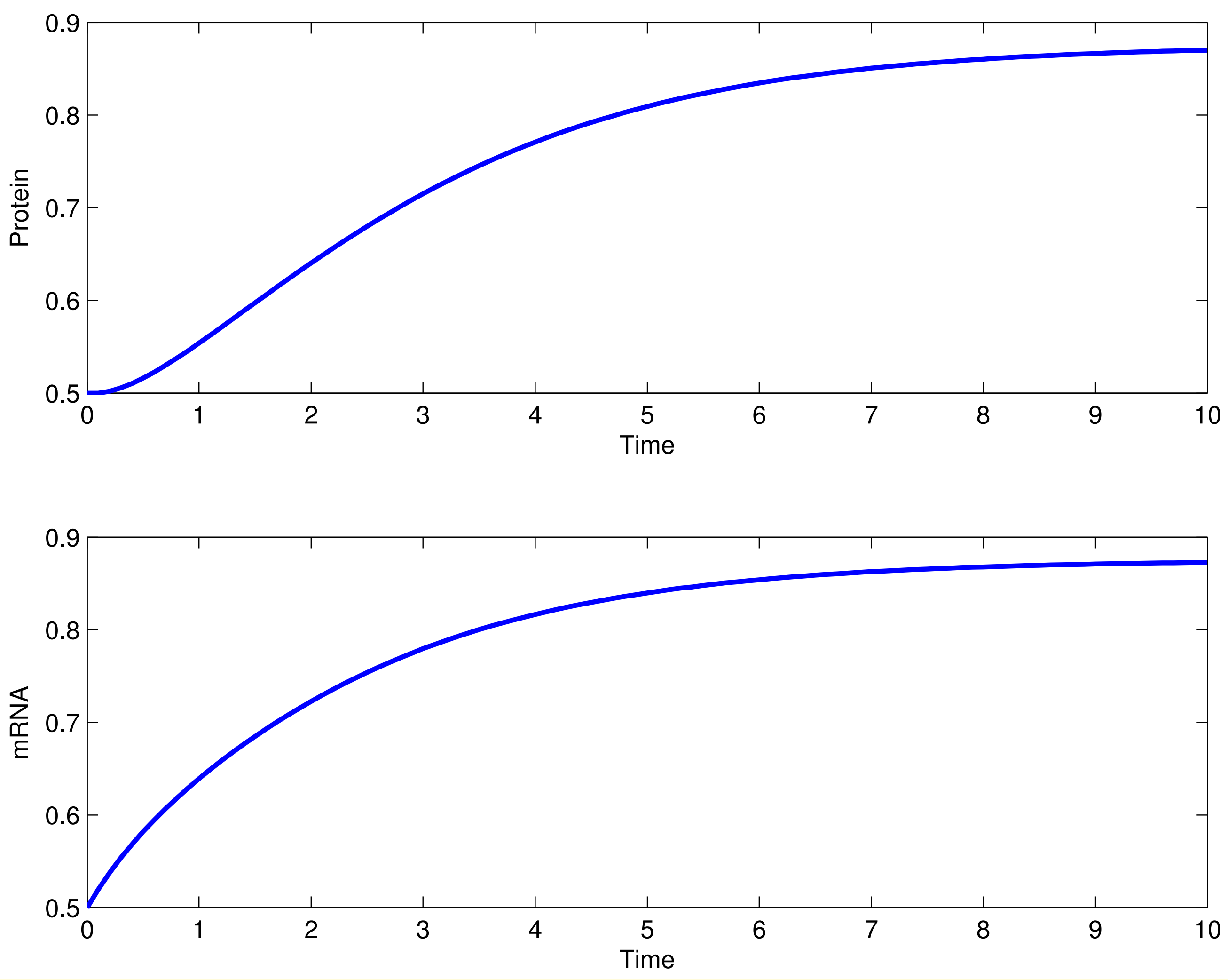
Varying the time step with constant total time.

- $t_{total} = 10\ s$
- $\Delta t = 0.2\ s$
- $n = \frac{t_{total}}{\Delta t} = 50$



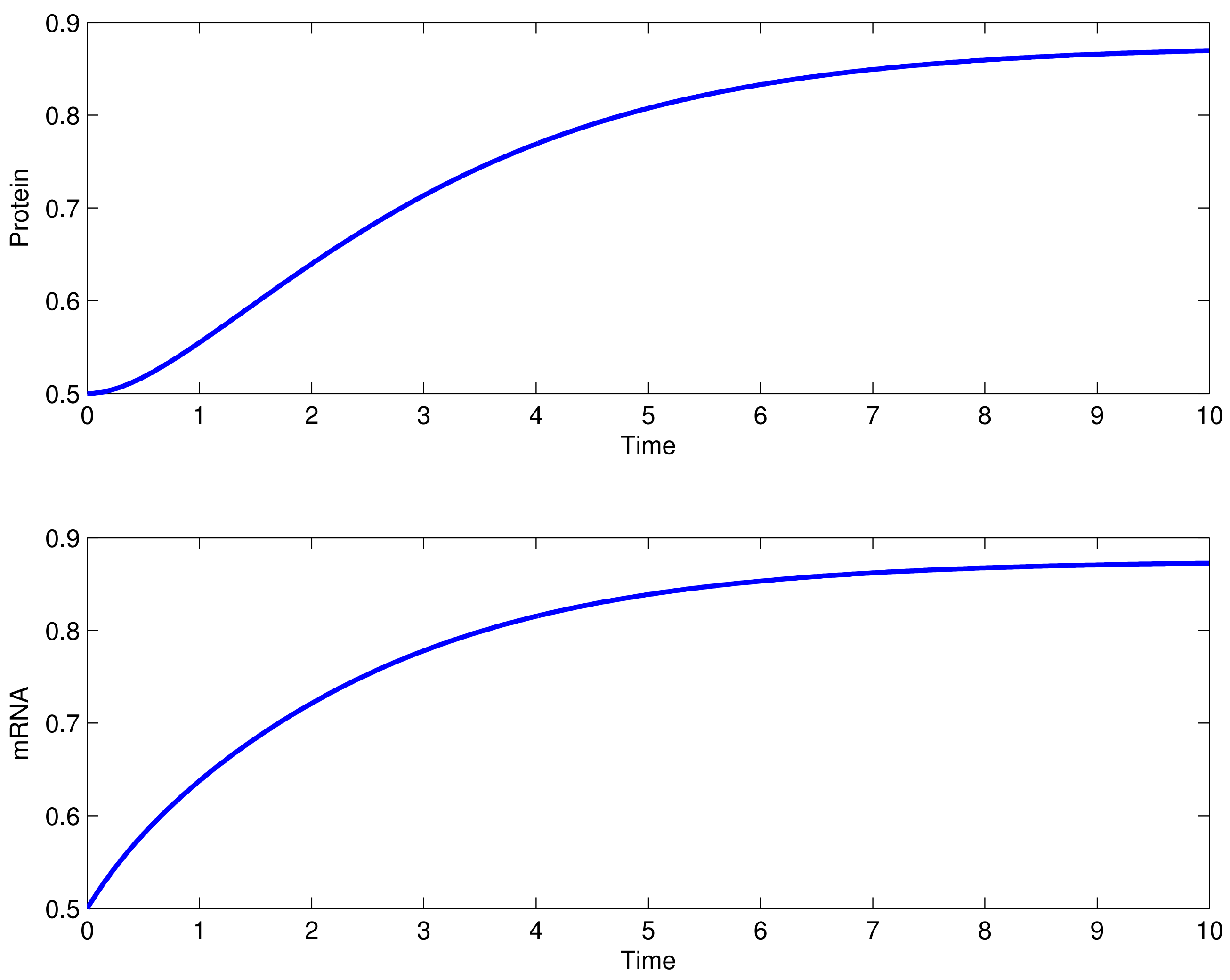
Varying the time step with constant total time.

- $t_{total} = 10\ s$
- $\Delta t = 0.1\ s$
- $n = \frac{t_{total}}{\Delta t} = 100$



Varying the time step with constant total time.

- $t_{total} = 10\ s$
- $\Delta t = 0.02\ s$
- $n = \frac{t_{total}}{\Delta t} = 500$

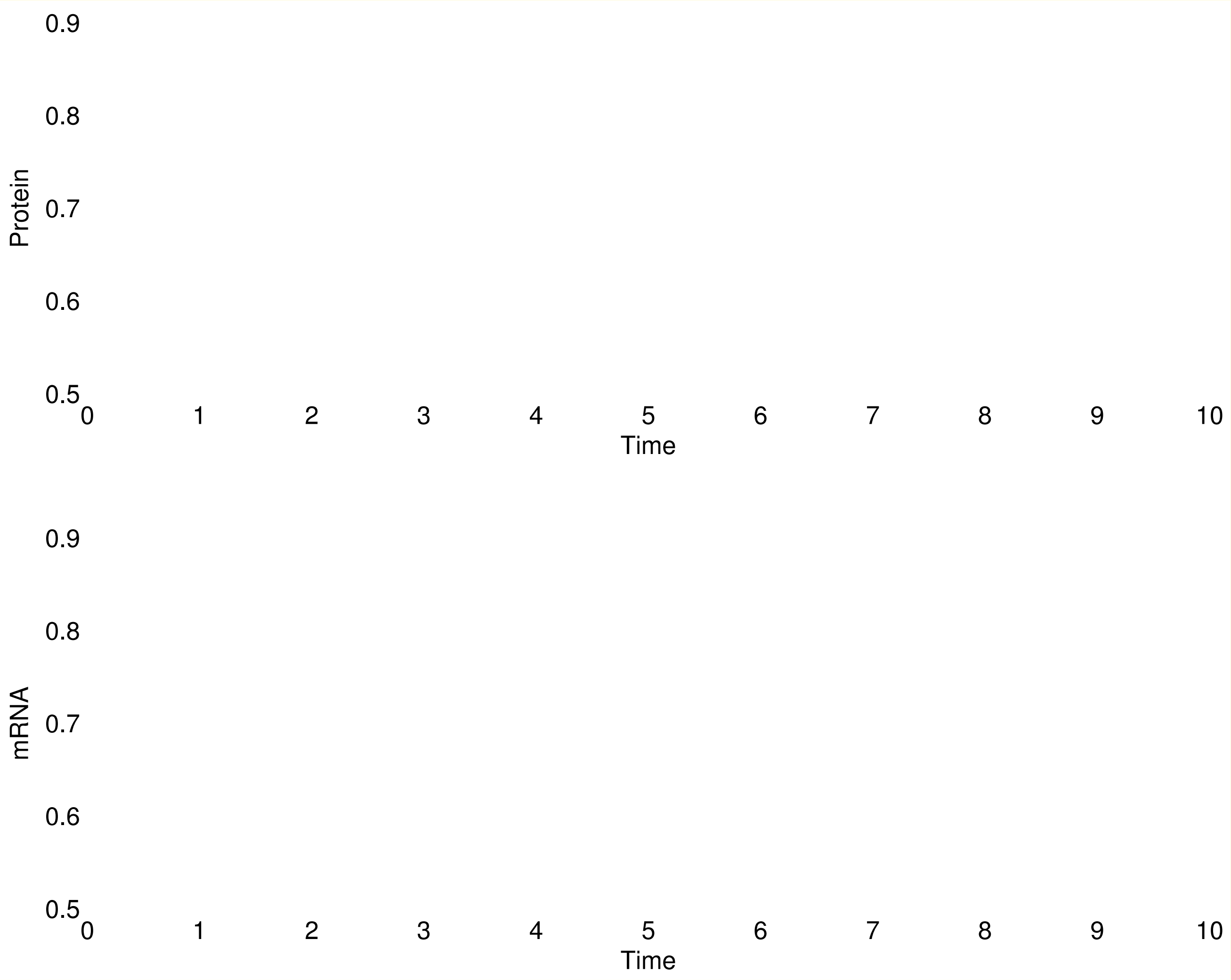


Varying the time step with constant total time.

- When the time step is particularly large, the trajectory can be obviously coarse.
- More subtle differences can occur at moderate step sizes.
- **Check the reasonableness of the time step by repeating the simulation with a smaller value (always divide by at least 2).**
  - Any differences should be very difficult to see!
  - With the total simulation time fixed, the computational cost (number of steps) will increase as the step size is reduced.

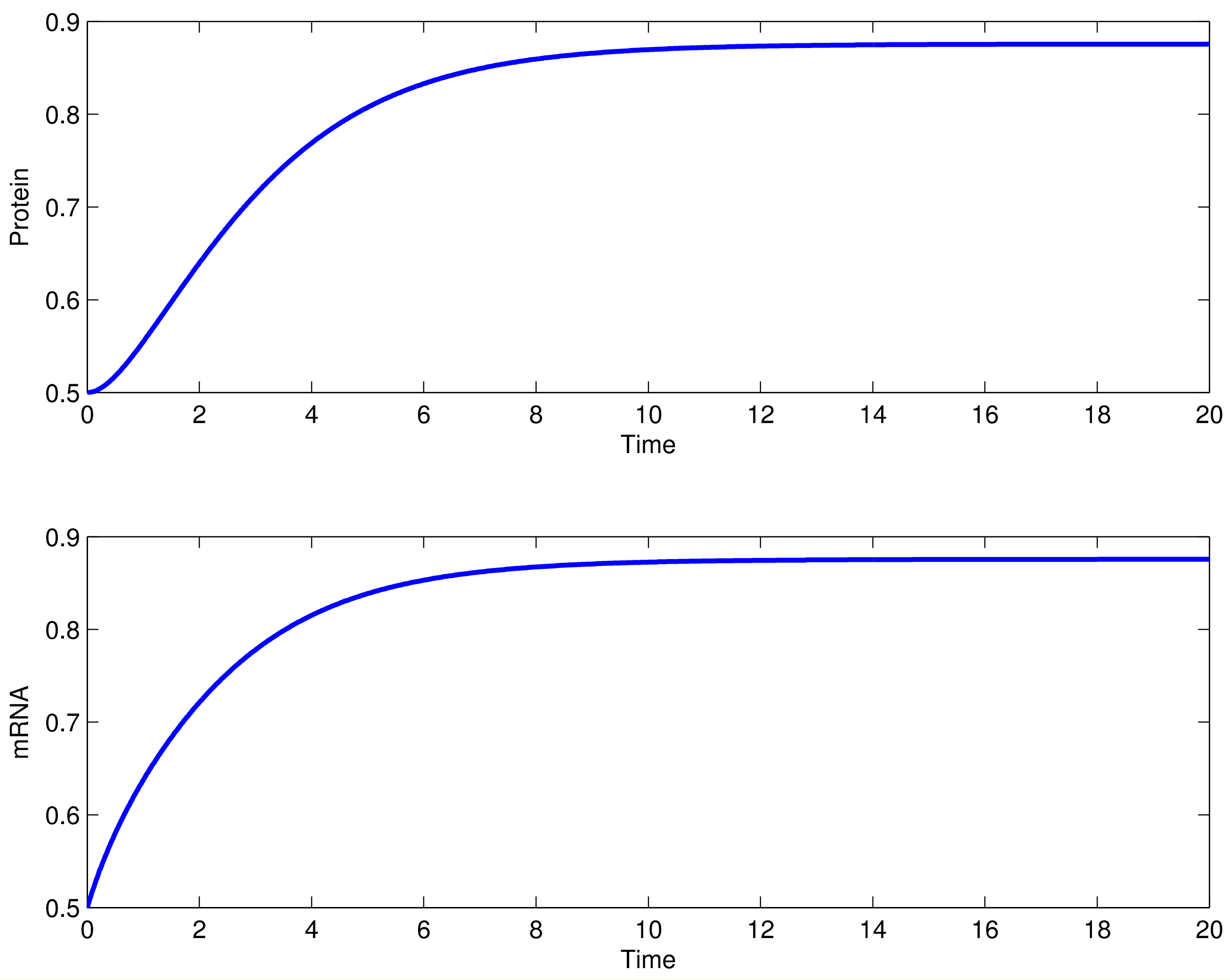
Varying the total time with constant time step.

- $t_{total} = 10\ s$
- $\Delta t = 0.02\ s$
- $n = \frac{t_{total}}{\Delta t} = 500$



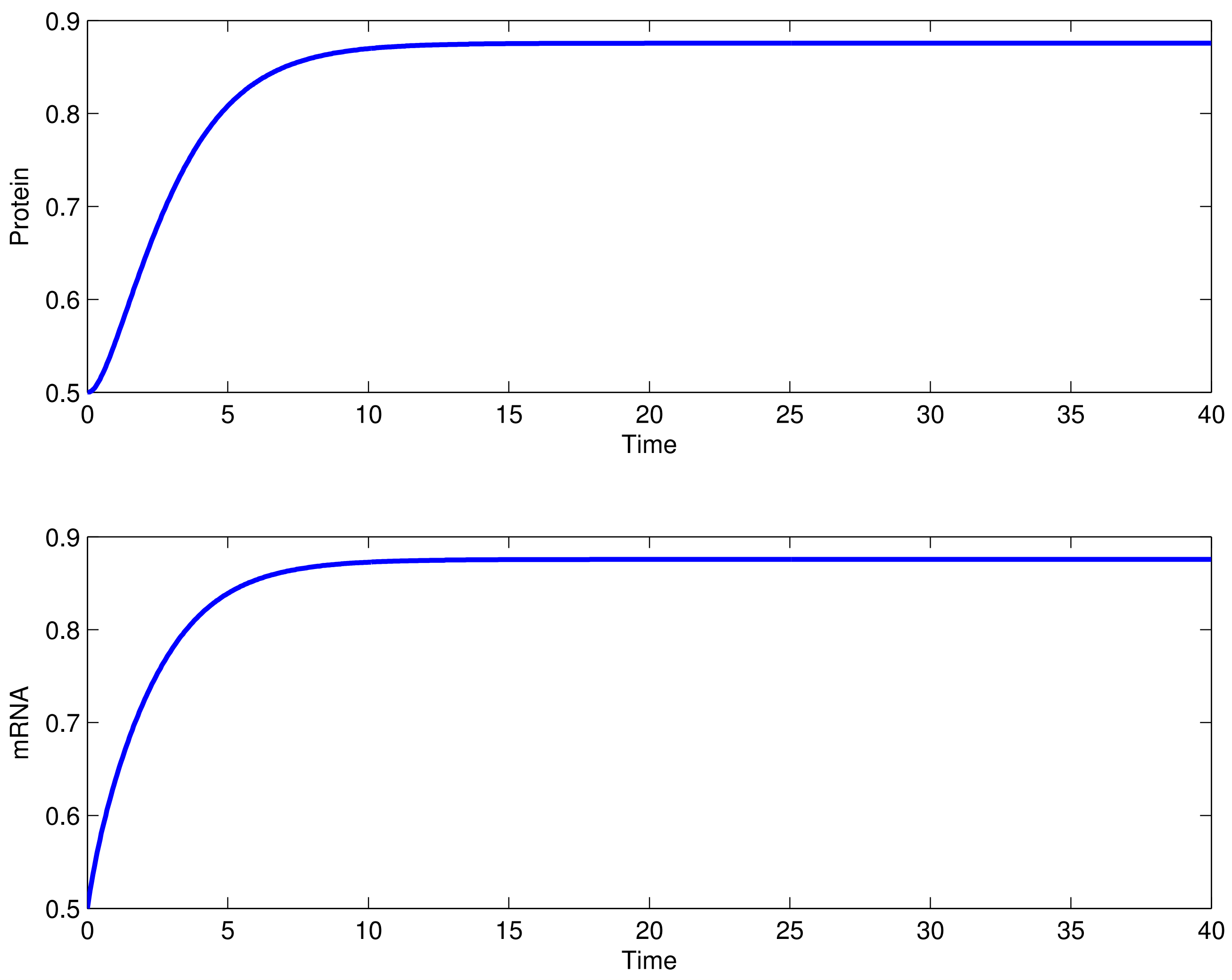
Varying the total time with constant time step.

- $t_{total} = 20\ s$
- $\Delta t = 0.02\ s$
- $n = \frac{t_{total}}{\Delta t} = 1000$



Varying the total time with constant time step.

- $t_{total} = 40\ s$
- $\Delta t = 0.02\ s$
- $n = \frac{t_{total}}{\Delta t} = 2000$



Varying the total time with constant time step.

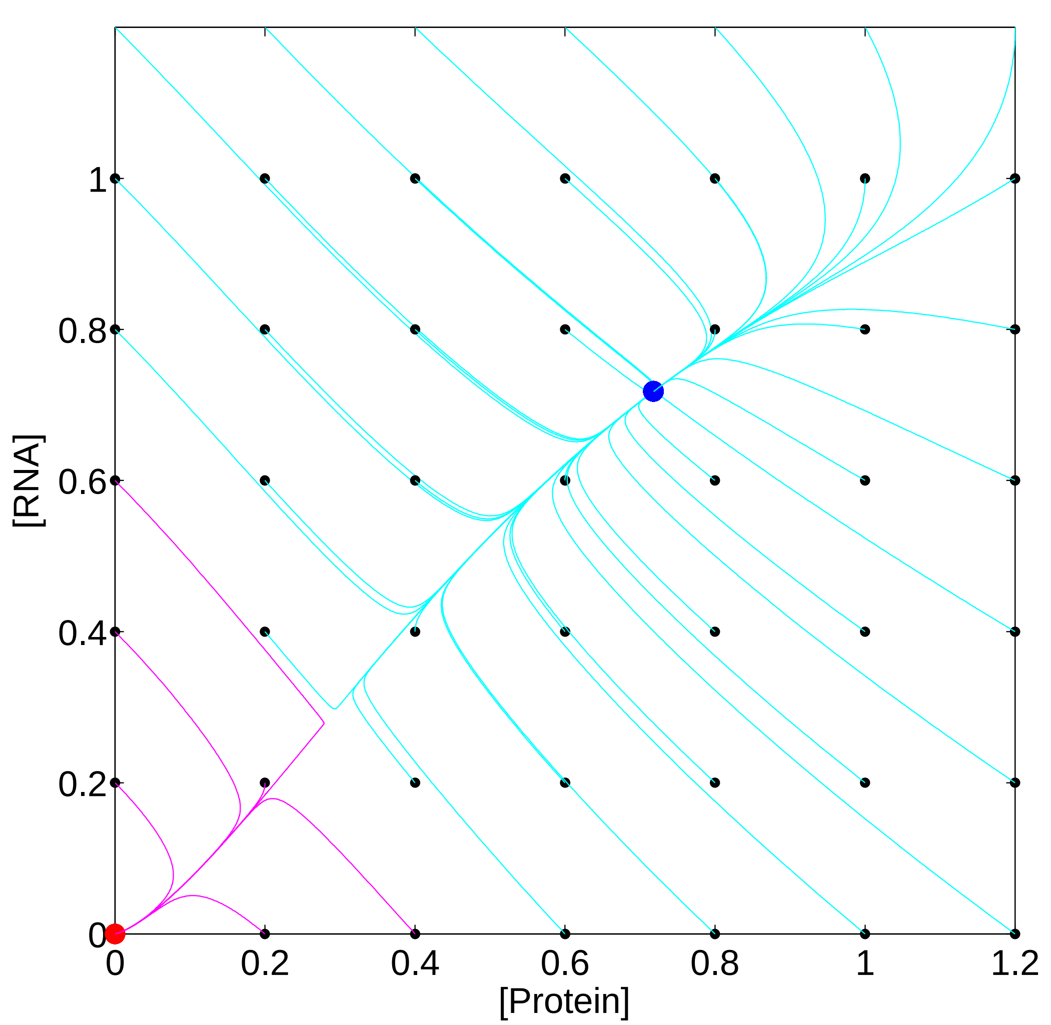
- Does the simulation appear to be approaching an equilibrium or steady state?
- Do you observe any persistent oscillations?
- Do some values appear to continuously grow?
- **Check consistency when doubling the length of the simulation.**
  - If you had reached equilibrium, there should be virtually no change in any variable for the second half of the longer simulation.
  - If there are oscillations, is the period and/or amplitude constant?

Varying the initial conditions.

- Once we have a reasonable idea of what to use for a time step and length of simulation, we may want to explore how initial conditions affect the results.
- It can be hard to compare many initial conditions on plots of concentration versus time; plotting pairs of concentrations against each other can be very informative.
- Each curve represents one simulation, with the end points representing the initial and final values.

Varying the initial conditions.

- $t_{total} = 40\ s$
- $\Delta t = 0.02\ s$
- $n = \frac{t_{total}}{\Delta t} = 2000$
- $[X_{prot}] = 0, 0.2, 0.4, 0.8, 1.0, 1.2$
- $[X_{rna}] = 0, 0.2, 0.4, 0.8, 1.0, 1.2$



Varying the initial conditions.

- When multiple starting conditions lead to a common end point, we have clear evidence for an equilibrium point or steady state.
- Systems can have more than one possible steady state; these may represent different biological states!
- **Important:**
  - If the simulation length isn't long enough, some of the trajectories may stop before reaching an equilibrium point.
  - If it looks like this may be occurring, increase the number of simulation steps!

Summary

- When simulating system of ODEs we need to be thoughtful about our choice of both the time step and the total time.
- Performing computational experiments where these are varied can help us make reasonable choices.
- The effect of initial conditions is often best observed by plotting the trajectories on a “phase plane:” one concentration plotted against another, rather than both plotted against time.
- Equilibrium points are characterized as a common end point for multiple initial conditions.

Mini-Lecture 8

Analytical analysis of mass-action kinetic systems.

Steady states, stationary points and equilibria.

- What does is mean if we observe that a system reaches a steady state?
- It does NOT mean that no reactions are occurring.
- However, the overall concentration of every species is not changing over time.
- Synthesis (production) and degradation (consumption) rate are balanced for every species.
- The system is in equilibrium.

Steady states, stationary points and equilibria.

- What does is mean if we observe that a system reaches a steady state?
- The rate of change in concentration for every species is zero.
- For example:

$$\frac{d[X_{pro}]}{dt} = 0 \qquad \frac{d[Y_{pro}]}{dt} = 0$$

$$\frac{d[X_{rna}]}{dt} = 0 \qquad \frac{d[Y_{rna}]}{dt} = 0$$

- We are at a root of the system of ODEs; a stationary point.

Steady states, stationary points and equilibria.

- In some cases, we can solve directly for the existence and location of these stationary points.

$$\frac{d[X_{prot}]}{dt} = \omega_X[X_{rna}] - \chi_{X_{pro}}[X_{pro}]$$

$$\frac{d[X_{rna}]}{dt} = \left(\frac{[X_{pro}]^2}{(K_{1/2})^2 + [X_{pro}]^2}\right)\mu_X - \chi_{X_{rna}}[X_{rna}]$$

$$\frac{d[X_{pro}]}{dt} = 0 \Rightarrow \omega_X[X_{rna}] = \chi_{X_{pro}}[X_{pro}]$$

$$\frac{d[X_{rna}]}{dt} = 0 \Rightarrow \left(\frac{[X_{pro}]^2}{(K_{1/2})^2 + [X_{pro}]^2}\right)\mu_X = \chi_{X_{rna}}[X_{rna}]$$

Steady states, stationary points and equilibria.

$$\frac{d[X_{prot}]}{dt}=0 \quad \Rightarrow \quad [X_{mRNA}] \quad = \quad \left(\frac{\chi_{X_{pro}}}{\omega_X}\right)[X_{pro}]$$

$$\frac{d[X_{mRNA}]}{dt}=0 \quad \Rightarrow \quad [X_{mRNA}] \quad = \quad \left(\frac{\mu_X}{\chi_{X_{rna}}}\right)\left(\frac{[X_{pro}]^2}{\left(K_{1/2}\right)^2+[X_{pro}]^2}\right)$$

$$\left(\frac{\chi_{X_{pro}}}{\omega_X}\right)[X_{pro}]=\left(\frac{\mu_X}{\chi_{X_{rna}}}\right)\left(\frac{[X_{pro}]^2}{\left(K_{1/2}\right)^2+[X_{pro}]^2}\right)$$

$$\left([X_{pro}]^2-\left(\frac{\omega_X\mu_X}{\chi_{X_{pro}}\chi_{X_{rna}}}\right)[X_{pro}]+\left(K_{1/2}\right)^2\right)[X_{pro}]=0$$

Steady states, stationary points and equilibria.

$$\frac{d[X_{prot}]}{dt} = 0 \quad \Rightarrow \quad [X_{mRNA}] \quad = \quad \left(\frac{\chi_{X_{pro}}}{\omega_X}\right)[X_{pro}]$$

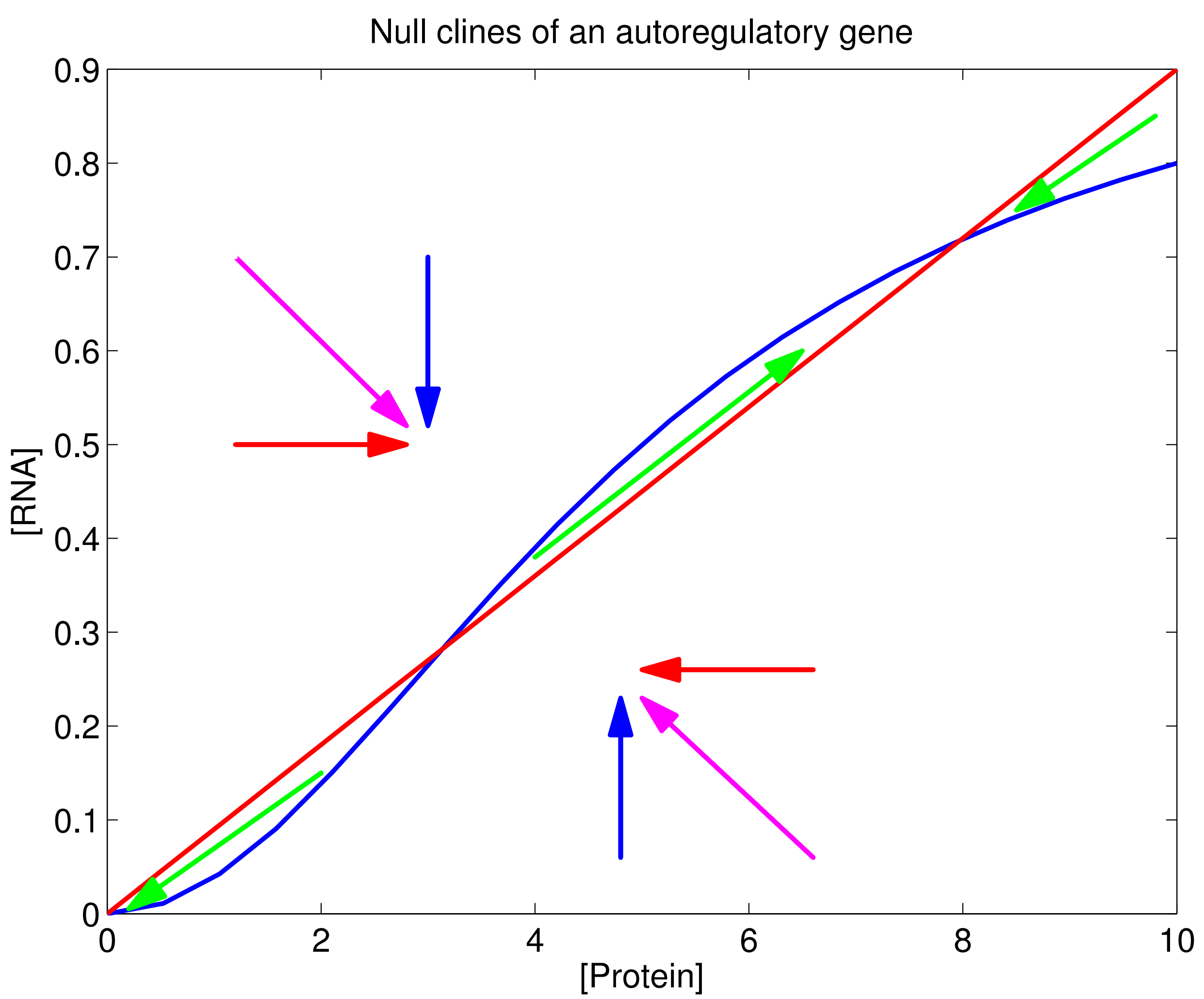
$$\frac{d[X_{mRNA}]}{dt} = 0 \quad \Rightarrow \quad [X_{mRNA}] \quad = \quad \left(\frac{\mu_X}{\chi_{X_{rna}}}\right)\left(\frac{[X_{pro}]^2}{\left(K_{1/2}\right)^2 + [X_{pro}]^2}\right)$$

- The roots of each individual equation are called **null clines**; they are generally curves (surfaces) in the state space.
- Stationary points occur where the null clines for all variables intersect.

Phase plane analysis.

- Null clines often indicate “switching lines” for the sign of the rate of change for one variable.
- The sign of a rate of change cannot change without crossing a null cline if the functions are all continuous.
- If we know the sign of each rate of change inside a region bounded by null clines, then we know those signs for all points in that region.
- We can easily sketch the general direction of motion in each region!

Phase plane analysis.



$$\frac{d[X_{rna}]}{dt} = 0$$

$$[X_{rna}] = \frac{\mu_X}{\chi_{Xrna}}\left(\frac{[X_{pro}]^2}{\left(K_{1/2}\right)^2 + [X_{pro}]^2}\right)$$

$$\frac{d[X_{pro}]}{dt} = 0$$

$$[X_{rna}] = \frac{\chi_{Xpro}}{\omega_X}[X_{pro}]$$

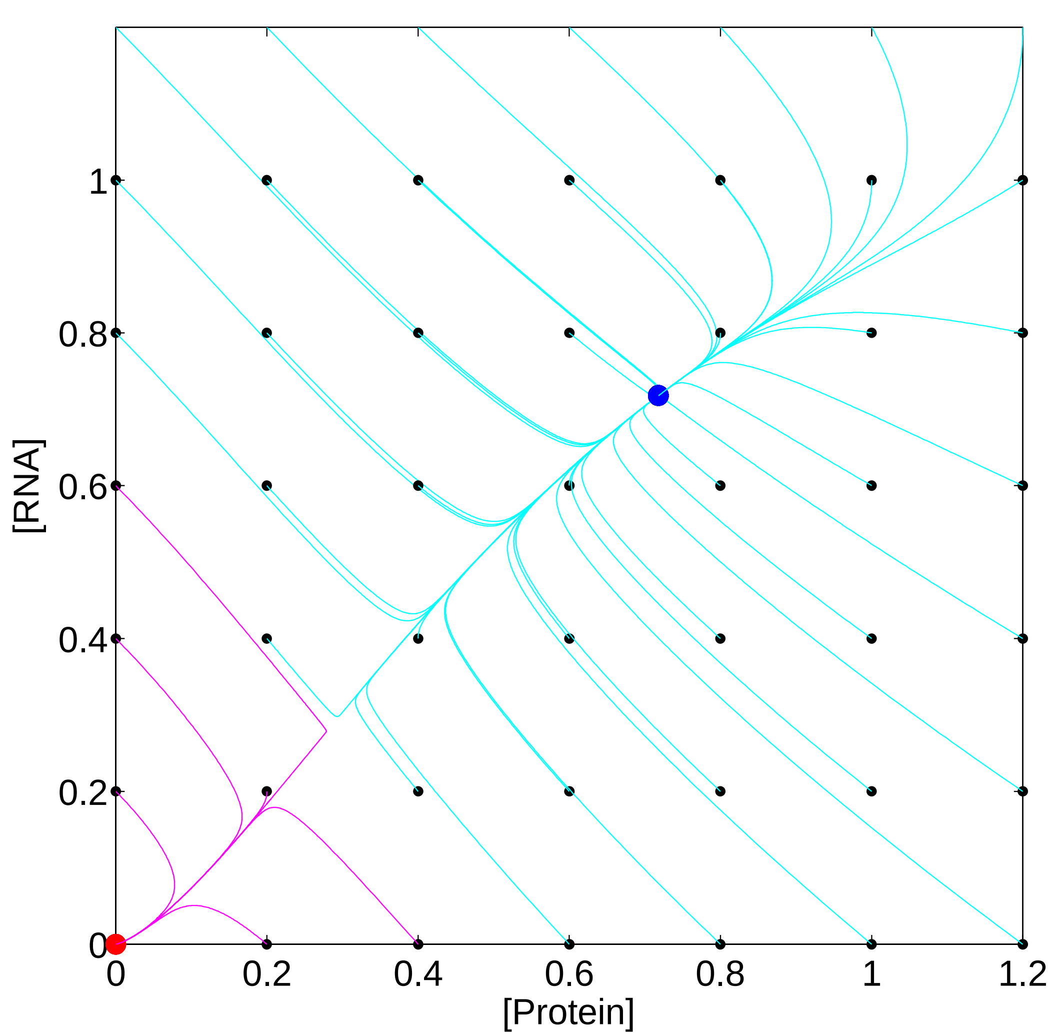
Phase plane analysis.

- Stationary points can be of different types:
  - **Stable points (attractors)** have the general lines of motion pointing towards them.
  - **Unstable points (repellors)** have the general lines of motion pointing away.
  - **Saddle points** have some lines pointing towards them and some pointing away.
  - Points with lines pointing “around” them yield oscillations.
- Note: Any pair of stable points must have a saddle (or unstable) point “in between” them; if a system has more than one stationary point, they can not all be stable.

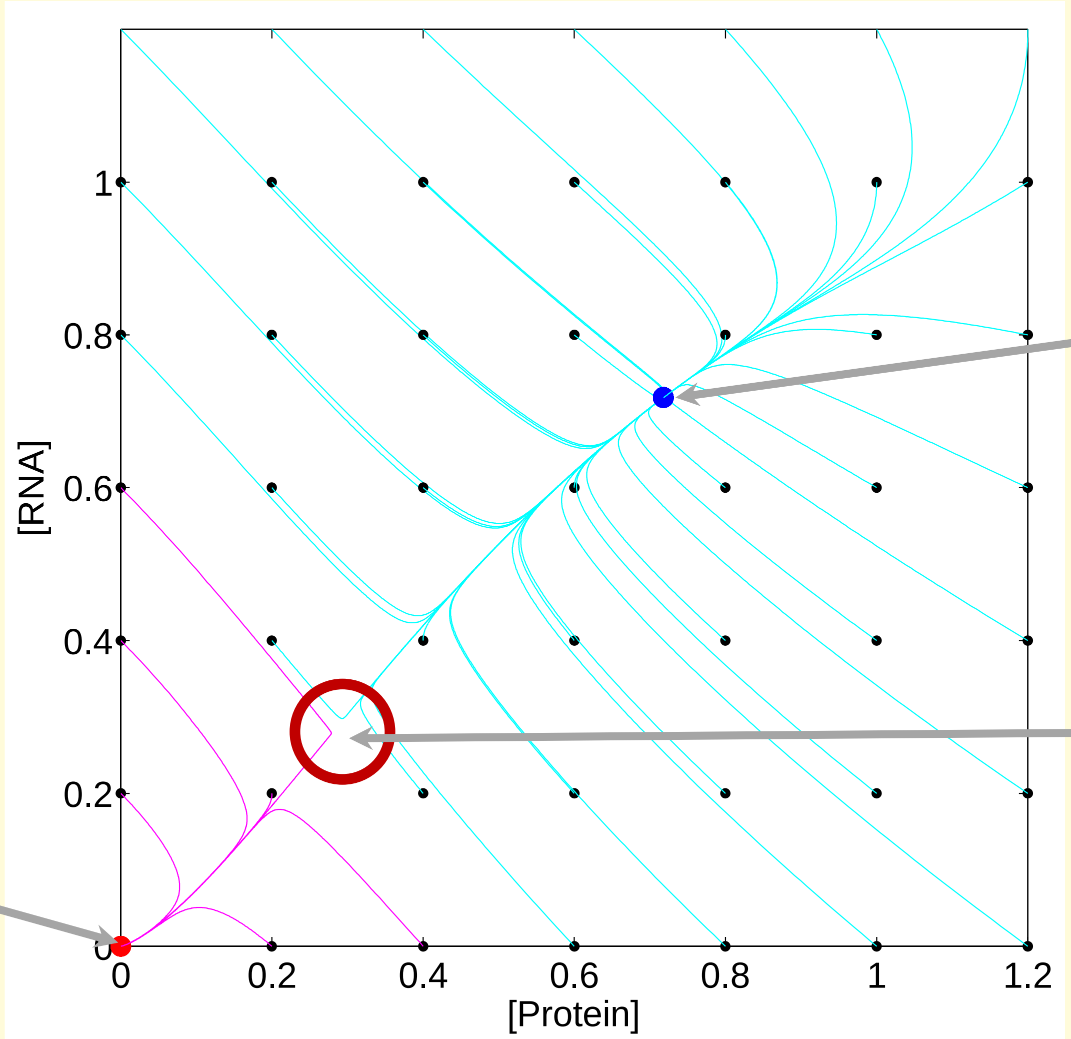
Numerical phase plane analysis.

- When solving for stationary points analytically is difficult, we can use simulations to deduce their location.
  - **Stable points (attractors)** exist at observed equilibria.
  - **Unstable points (repellors) and saddle points** are suggested by the curve of trajectories away from them.
- We can also identify lines of attraction and repulsion, manifolds:
  - **Attracting manifolds** “collect” trajectories as they move towards a stable point.
  - **Repelling manifolds** define boundaries between regions in which all trajectories flow towards a common stable point.

Numerical phase plane analysis.



Numerical phase plane analysis.

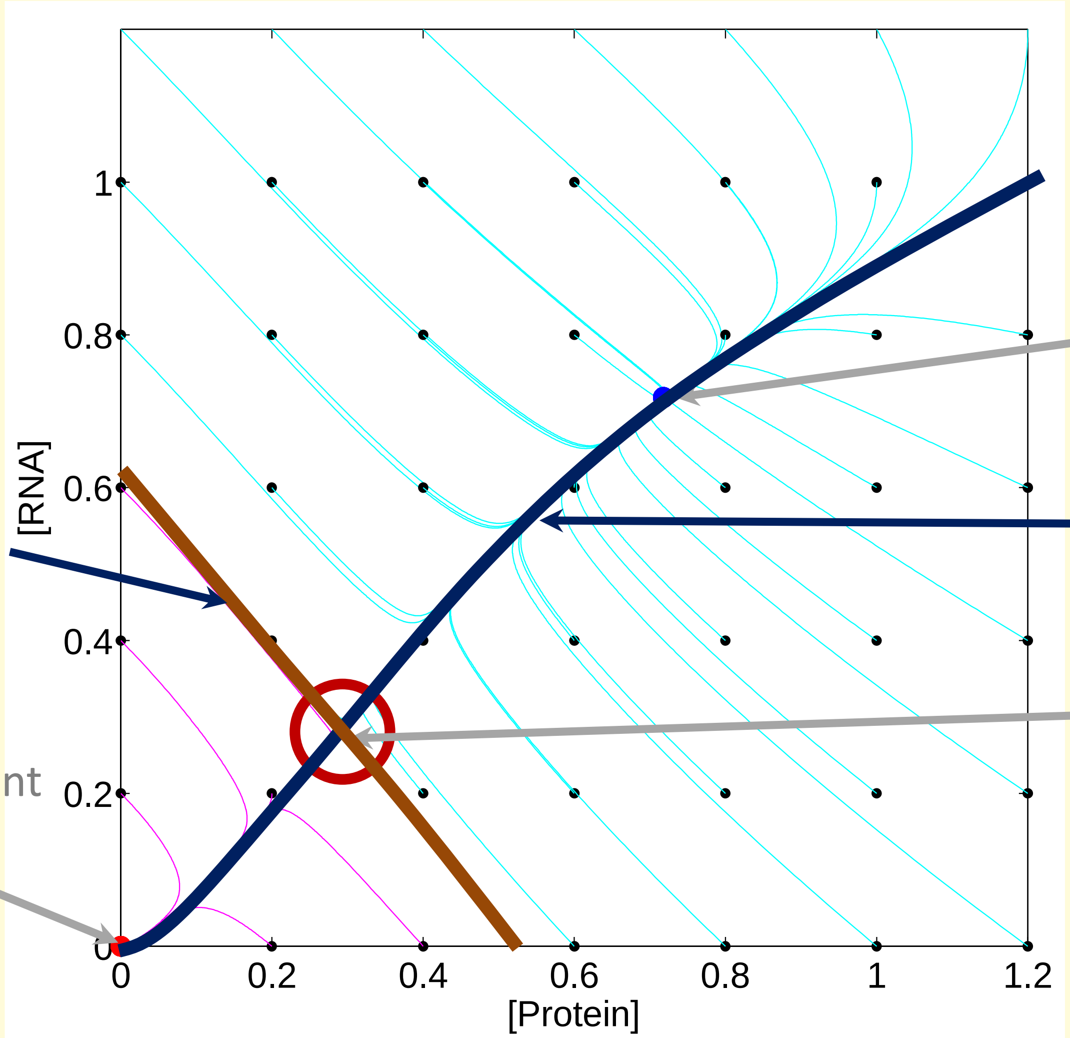


Stable stationary point

Saddle point

Stable stationary point

Numerical phase plane analysis.



Stable stationary point

Attracting manifold

Saddle point

Repelling manifold

Stable stationary point

Summary

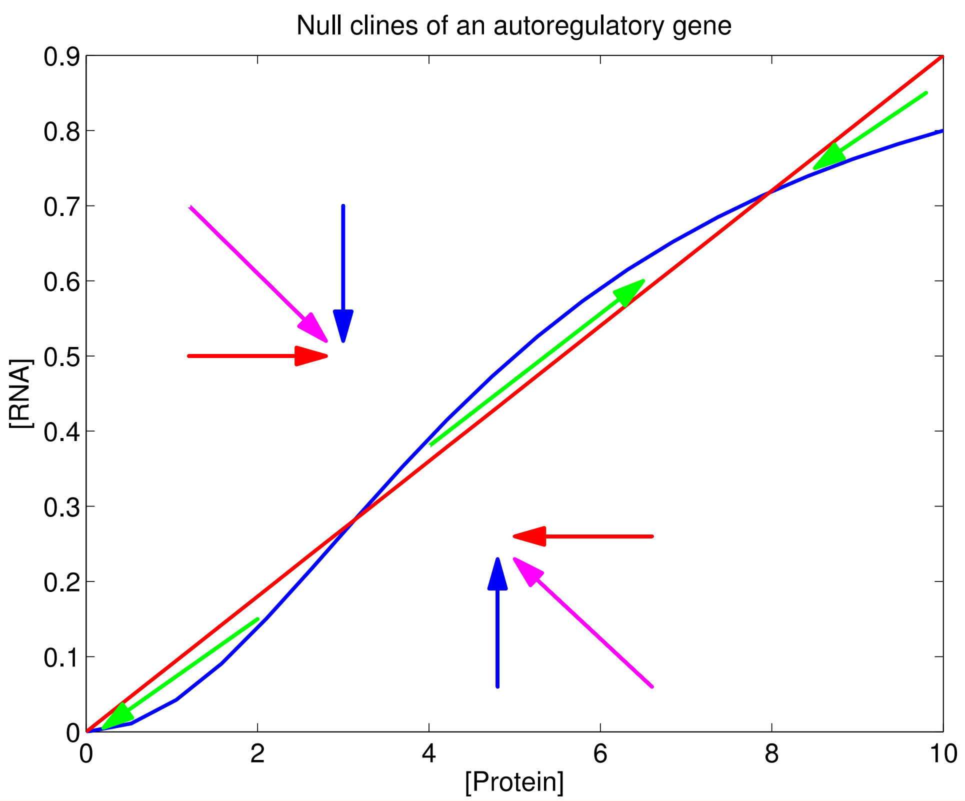
- Equilibrium points are closely related to mathematically-defined stationary points.
- We can find stationary points of a system of ODEs analytically by solving for the intersection of null clines.
- We can also use a graphical analysis based on plotting the null clines to characterize these.
- Numerical analysis can also be used to deduce the position and character of stationary points.

Mini-Lecture 9

From math to biology

Developing insights into function

- What can the structure of the phase plane tell us about the function of the biological system?
- Two stable points:
  - One with no protein or mRNA.
  - One with both mRNA and protein.
- A bistable system: ON and OFF
- Easy to imagine biological utility!

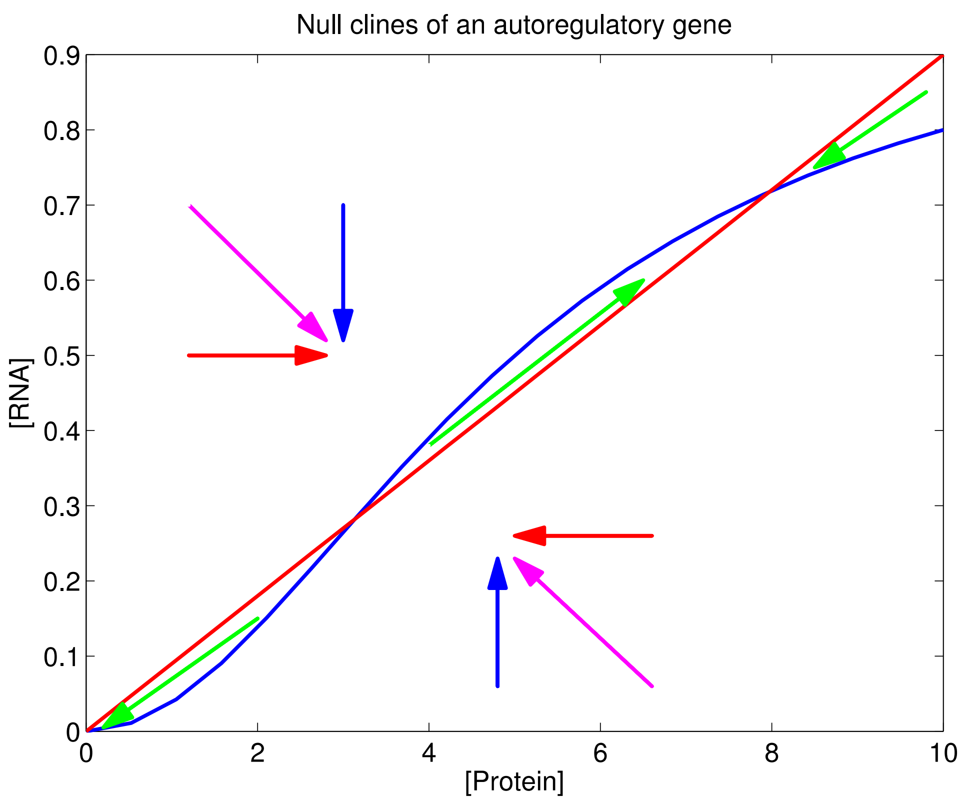


Developing insights into function

- How do we switch between the ON and OFF state?
- In physical/electrical systems, we apply an external force.
  - The force needs to be strong enough and must last long enough to activate a change, but when removed the state is stable.
- Biochemical “forces” include:
  - Addition of a molecular species from another source.
  - A temporary “activation” of a new reaction.
  - A temporary change in the rates of a particular reaction.

Developing insights into function

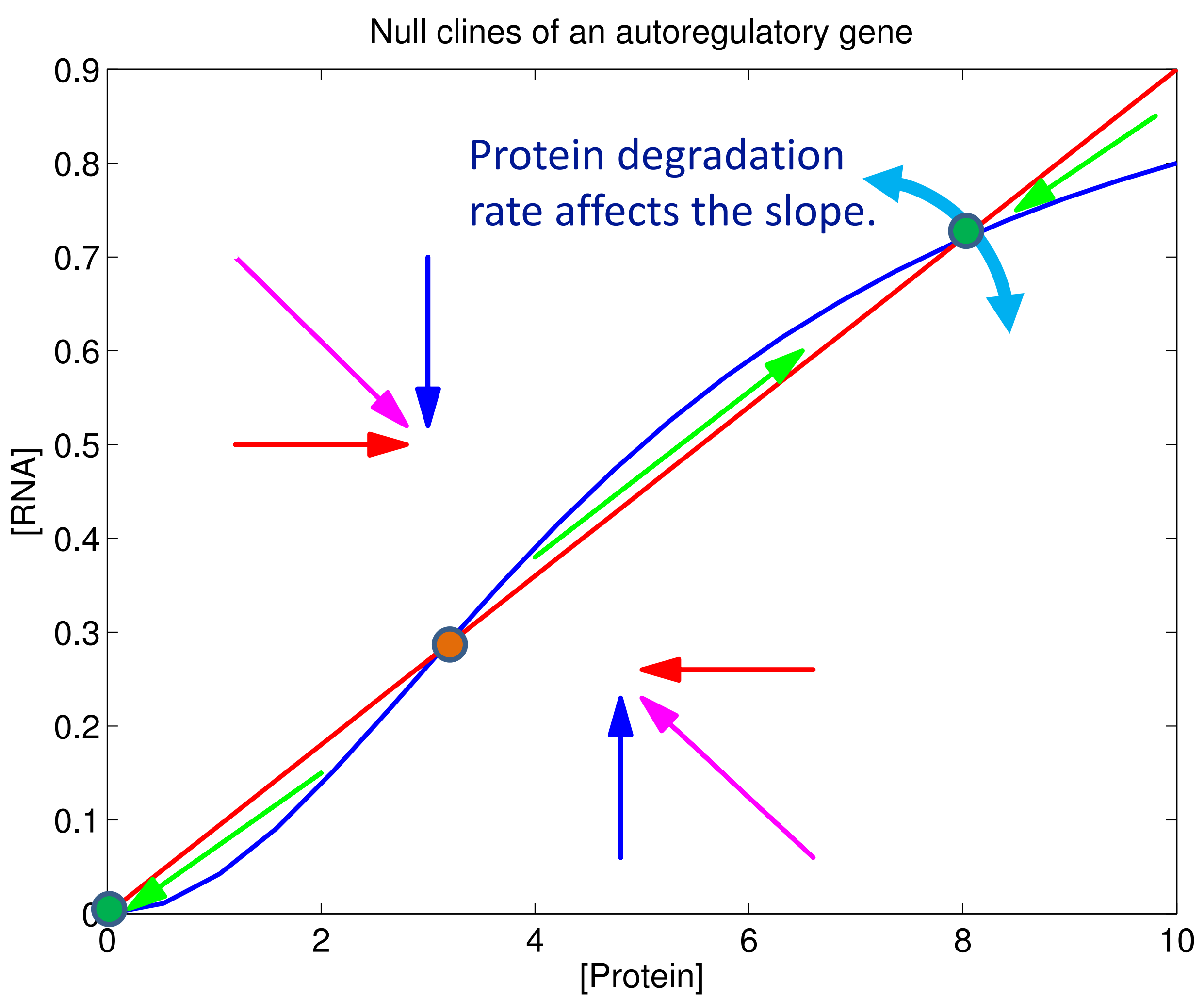
- To turn the system ON:
  - Synthesis ↑ or degradation ↓
- To turn the system OFF:
  - Degradation ↑ or synthesis ↓



$$\frac{d[X_{prot}]}{dt} = \omega_X[X_{rna}] - \chi_{X_{pro}}[X_{pro}]$$

$$\frac{d[X_{rna}]}{dt} = \left(\frac{[X_{pro}]^2}{(K_{1/2})^2 + [X_{pro}]^2}\right)\mu_X - \chi_{X_{rna}}[X_{rna}]$$

Protein degradation as an OFF switch.

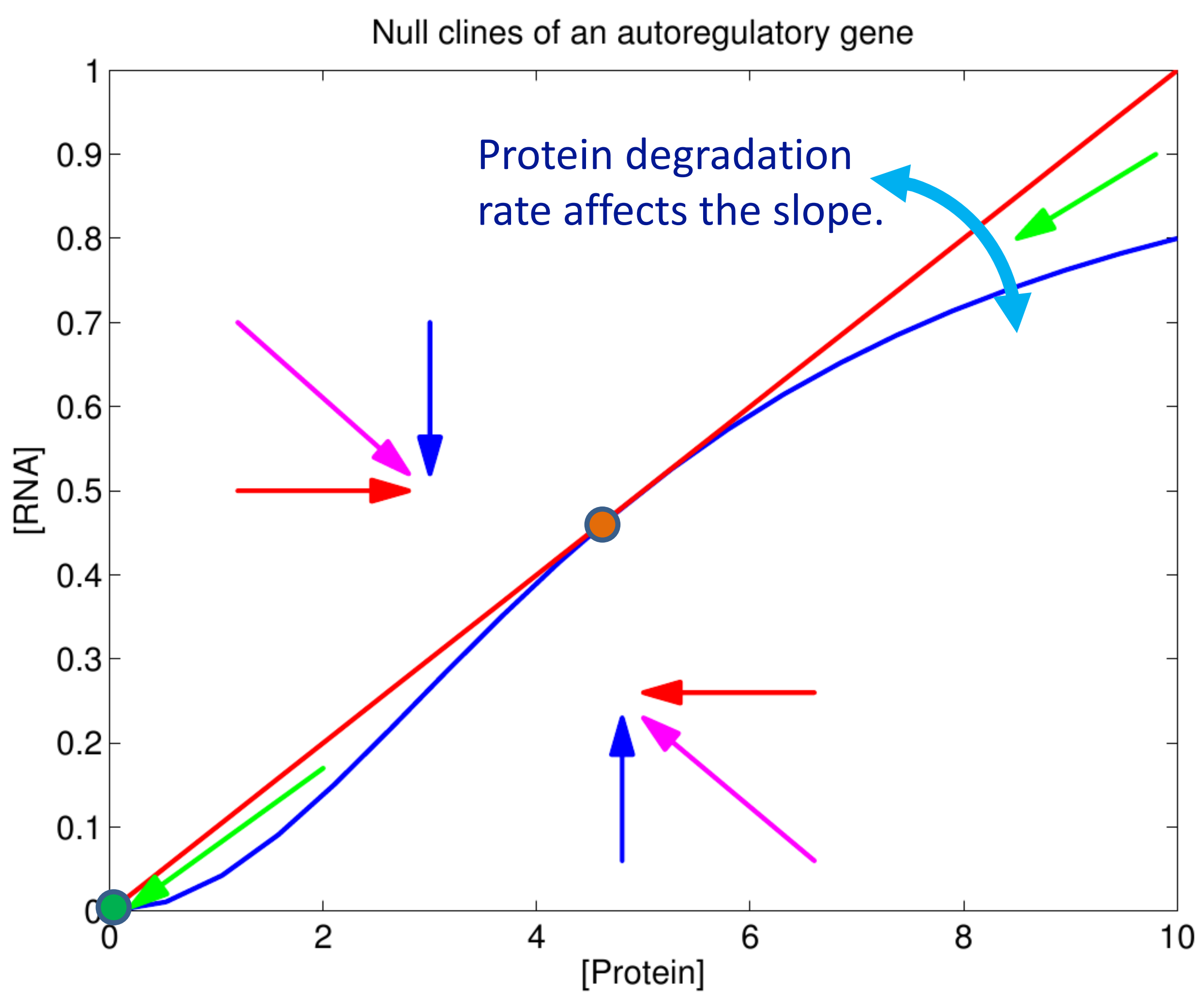


$$[X_{rna}] = \frac{\mu_X}{\chi_{Xrna}} \left(\frac{[X_{pro}]^2}{\left(K_{1/2}\right)^2 + [X_{pro}]^2} \right)$$

$$[X_{rna}] = \frac{\chi_{Xpro}}{\omega_X} [X_{pro}]$$

$$K_{1/2} = 5.0\ M$$
$$\mu_X = 0.50\ M \cdot s^{-1}$$
$$\chi_{Xrna} = 0.50\ s^{-1}$$
$$\omega_X = 2.0\ s^{-1}$$
$$\boldsymbol{\chi_{Xpro} = 0.18\ s^{-1}}$$

Protein degradation as an OFF switch.



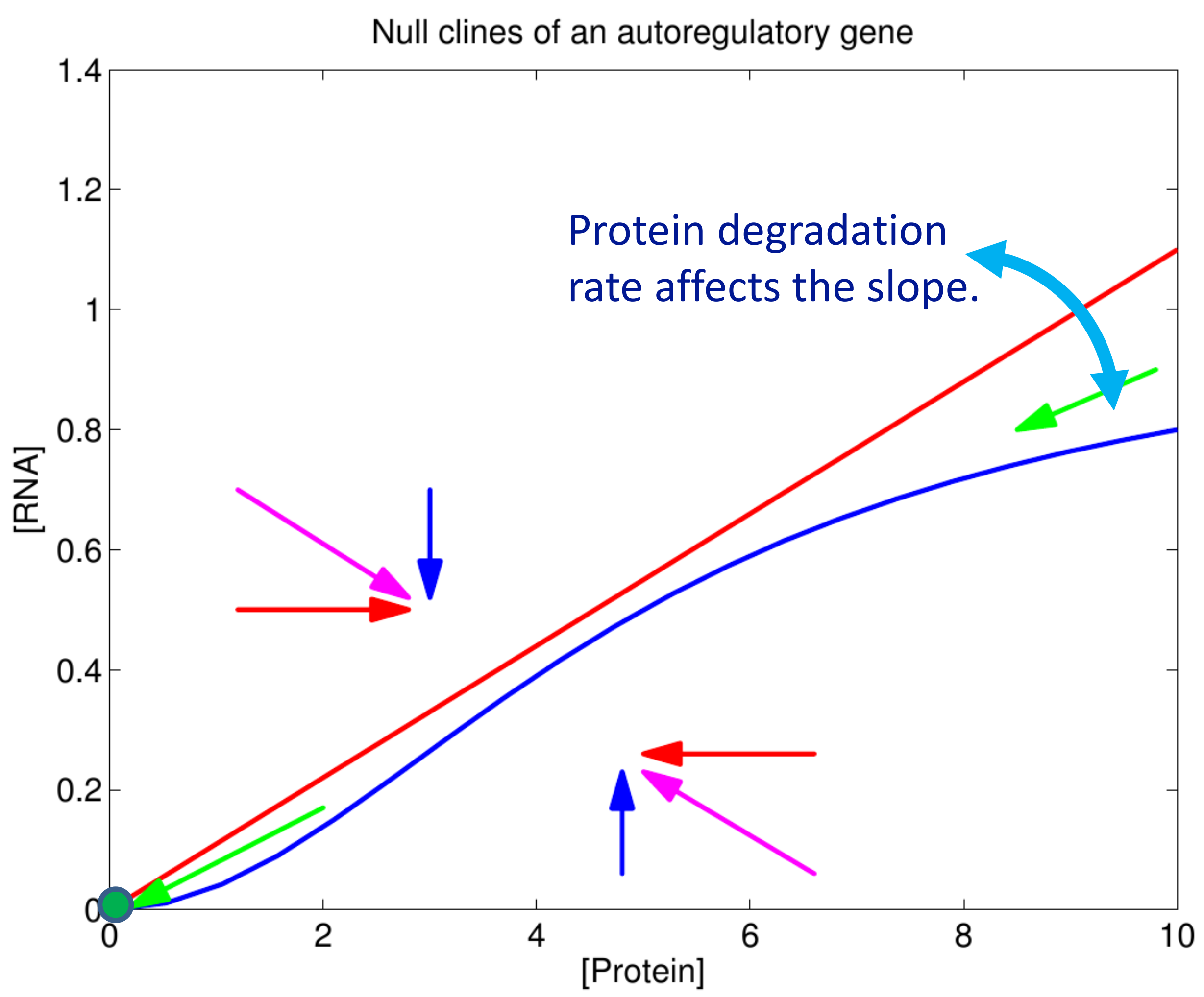
Protein degradation rate affects the slope.

$$[X_{rna}] = \frac{\mu_X}{\chi_{Xrna}}\left(\frac{[X_{pro}]^2}{\left(K_{1/2}\right)^2 + [X_{pro}]^2}\right)$$

$$[X_{rna}] = \frac{\chi_{Xpro}}{\omega_X}[X_{pro}]$$

$$K_{1/2} = 5.0\ M$$
$$\mu_X = 0.50\ M \cdot s^{-1}$$
$$\chi_{Xrna} = 0.50\ s^{-1}$$
$$\omega_X = 2.0\ s^{-1}$$
$$\boldsymbol{\chi_{Xpro} = 0.20\ s^{-1}}$$

Protein degradation as an OFF switch.



Protein degradation rate affects the slope.

$$[X_{rna}] = \frac{\mu_X}{\chi_{Xrna}}\left(\frac{[X_{pro}]^2}{\left(K_{1/2}\right)^2 + [X_{pro}]^2}\right)$$

$$[X_{rna}] = \frac{\chi_{Xpro}}{\omega_X}[X_{pro}]$$

$$K_{1/2} = 5.0\ M$$

$$\mu_X = 0.50\ M \cdot s^{-1}$$

$$\chi_{Xrna} = 0.50\ s^{-1}$$

$$\omega_X = 2.0\ s^{-1}$$

$$\boldsymbol{\chi_{Xpro} = 0.22\ s^{-1}}$$

Protein degradation as an OFF switch.

A saddle-node bifurcation:

- As the protein degradation constant increases, the saddle point and the “ON” stable point get closer together.
- When they meet, they disappear, leaving only the “OFF” stable point.

- When the “ON” state disappears, the system will evolve towards the “OFF” state.
- Once the system is near the origin, a return to the original degradation rate will leave the system “OFF”.
- If the increased degradation “signal” lasts for too short a time, the system may return to the “ON” state.

Summary

- Bistability in an autoregulatory gene can be interpreted biologically as a two state ON/OFF system.
- Biochemical “forces” such as increased degradation can switch a system from one state to another.
- A transient force can lead to a long-term change in state.