Gene Expression - The Basics

PhET Sim Design Document

Alternative Names:
Gene Expression
Gene Expression I
Basic Gene Expression
Protein Synthesis

Design Team:
John Blanco
Mike Klymkowsky
Kathy Perkins

Public URL

https://docs.google.com/a/colorado.edu/document/d/1iZ6MJm7rABKpIz3bssL-UtI1CqV2SdbXq7rhUkVI0hA/edit?hl=en#

Abstract

This sim is intended to help students learn the basic processes involved in gene expression, emphasizing the role of randomness, competition, adaptation and homeostasis.

Contents

Audience

This simulation is intended to be used primarily by high school biology students and early college undergraduates who are taking introductory biology courses, and by the teachers of these students. These students, if they continue to study biology, may end up as the next generation of doctors, nurses, pharmacists, nutritionists, geneticists, drug researchers, and researchers in the areas of basic biology.

There are certain key terms and concepts that students will need to know prior to using this simulation in order to get anything out of it. The following is a list of some of these terms and concepts:

- Cells - Students will need to understand that cells are the basic unit of life
    - cells are distinguished by their limiting membrane.
    - organisms can range from one to trillions of cells.
    - cells can be specialized (even if they carry the same DNA)
- DNA - Students will need to recognize a drawing of a strand of DNA (note that it may be labeled as such) and should know that DNA exists inside cells and encodes information that is important to a cell.
    - That information is the product of mutation and selection.

    - DNA strands have an directionality - they have a 5’ and a 3’ end.
    - DNA exists in cells as a double helix, with the two strands oriented in opposite directions (anti-parallel)
- Protein - Students should understand that proteins are important to the operation of cells.
- mRNA - Students should understand that the synthesis of proteins (translation) depends upon the translation machinery (ribosomes and associated factors) and mRNA, which supplies the information to specify the order of amino acids with a polypeptide
    - mRNA are synthesized beginning at their 5’ end
- Transcription
    - Mediated by DNA-dependent, RNA polymerase (polymerase for short)
    - The start of transcription is controlled by DNA binding transcription factors (proteins)
    - transcription factors bind to DNA with a direction
- Translation
    - Ribosomes engage mRNAs, producing a polypeptide
    - 1 ribosome + 1 mRNA = 1 polypeptide
- There is a lot of terminology in this simulation that is specific to the domain of cell biology, such as ‘polymerase’ and ‘ribosome’. Some familiarity with such terms would be needed to avoid cognitive overload.

Learning Goals

After using this simulation, students should be able to:

- Explain that a gene is a portion of the DNA.
- Explain the basic anatomy of a gene, i.e. that they have a regulatory region, a coding region, and transcription and translation start and stop markers.
- Explain the main sequence of events that occur within a cell that leads to polypeptide (i.e. protein) synthesis.
- Explain the role that probability (stochastic processes) plays in the internal workings of a cell with respect to the expression of a gene.
    - How does its stochastic nature influence gene expression?
    - Why are some stochastic events more “limiting” than others?
        - Show the number of polypeptides in a cell over time
- What are the major molecules and macromolecular assemblies involved in gene expression? Define, in terms of their role within the cell:
    - DNA
    - DNA-dependent, RNA Polymerase
    - transcription factors
    - mRNA
    - ribosomes
- What is the “Central Dogma of Modern Biology” and why is it important?
    - How does the Central Dogma effect evolutionary mechanisms?

- Explain how the levels of the different molecules and macromolecular assemblies influence one another, and therefore also affect the overall operation of the cell.
- Explain why it is that the levels of a given protein in a cell cannot change instantaneously.
    - How might one instantaneously change protein activity?
- Provide a scenario in which the variation in protein level/activity (noise) leads to changes in cell behavior; why doesn't noise always lead to changes in cell behavior?
    - How can noise be suppressed.

Here are some of the relevant benchmarks from The American Association for the Advancement of Science (AAAS). The benchmarks were found by going to the site http://www.project2061.org/publications/bsl/online/index.php and searching on terms like 'DNA' and 'protein'. This simulation is not attempting to address all of these benchmarks, since the scope is simply too large, but the use of the simulation supports many of the goals listed in the benchmarks. The on line benchmarks state:

By the end of 5th grade, students should know that:

- Microscopes make it possible to see that living things are made mostly of cells. 5C/E2a

By the end of the 8h grade, students should know that:

- The way in which cells function is similar in all living organisms.

By the end of the 12th grade, students should know that:

- The genetic information encoded in DNA molecules provides instructions for assembling protein molecules.
- Complex interactions among the different kinds of molecules in the cell cause distinct cycles of activities, such as growth and division. Cell behavior can also be affected by molecules from other parts of the organism or even other organisms.
- Complex interactions among the different kinds of molecules in the cell cause distinct cycles of activities, such as growth and division. Cell behavior can also be affected by molecules from other parts of the organism or even other organisms.
- Different parts of the genetic instructions are used in different types of cells, influenced by the cell's environment and past history. 5B/H6b
- Genes are segments of DNA molecules. Inserting, deleting, or substituting segments of DNA molecules can alter genes. An altered gene may be passed on to every cell that develops from it. The resulting features may help, harm, or have little or no effect on the offspring's success in its environment. 5B/H4*
- The work of the cell is carried out by the many different types of molecules it assembles, mostly proteins. Protein molecules are long, usually folded chains made from 20 different kinds of amino acid molecules. The function of each protein molecule depends on its specific sequence of amino acids and its shape. The shape of the chain is a consequence of attractions between its parts. 5C/H3 [this last sentence is incorrect, it ignores entropy, which arises from interactions with water].

- Within the cells are specialized parts for the transport of materials, energy capture and release, protein building, waste disposal, passing information, and even movement. 5C/H2a

Since Mike K is interested in incorporating the importance of probability and randomness in cell processes, the AAAS database was searched for anything that connected randomness and biology. Nothing was found directly, but there are some benchmarks connected with randomness in general that might be relevant. They are:

- It is not always easy to recognize meaningful patterns of change in a set of data. Data that appear to be completely irregular may be shown by statistical analysis to have underlying trends or cycles. On the other hand, trends or cycles that appear in data may sometimes be shown by statistical analysis to be easily explainable as being attributable only to randomness or coincidence. 11C/H9** (SFAA)

Questions about Goals

- What do we want to teach about transcription factors? If so, what do we want them to learn?
    - They connect to DNA? MK: They are proteins, they bind to DNA with a certain affinity, there is lots of DNA, typically combinations of transcription factors are required to recognize a gene, they act in combination with non-DNA binding proteins to recruit and activate DNA-dependent, RNA polyerase.
    - They influence how much of a protein is produced? indirectly
    - They have different affinities? for each other, other proteins, and different DNA sequences
    - They fall off after a while? they are “knocked off” by thermal motion.
    - They have a finite lifetime? they have a finite life-time, because of degradative systems in the cell.

- Do we want to teach about regulatory regions on the Gene? If so, what are the key points? MK: sure, major points. Bound by transcription factors (often multiple factors bindnin regulatory region. Can lead to gene activation or repression through competition between distinct factors. Transcription factors themselves are gene products, and so their synthesis is regulated in the same manner. Networks of factors active within a cell; which active will determine which genes active in a particular cell.
- MK: cell is continuous; with 3.5 billion year history,
- Gene anatomy - Do we agree that it is important (some of the biologist interviewed said that it is) and, if so, what exactly do we want students to come away with? MK: This is silly, since regulatory and coding regions can overlap. (perhaps i do not understand what "gene anatomy" means).

Non-Goals

There is a lot to know about how cells work, and this simulation can only address a small portion of that knowledge. Here are some areas that are related to gene expression and protein synthesis in cells that will not be covered in this simulation.

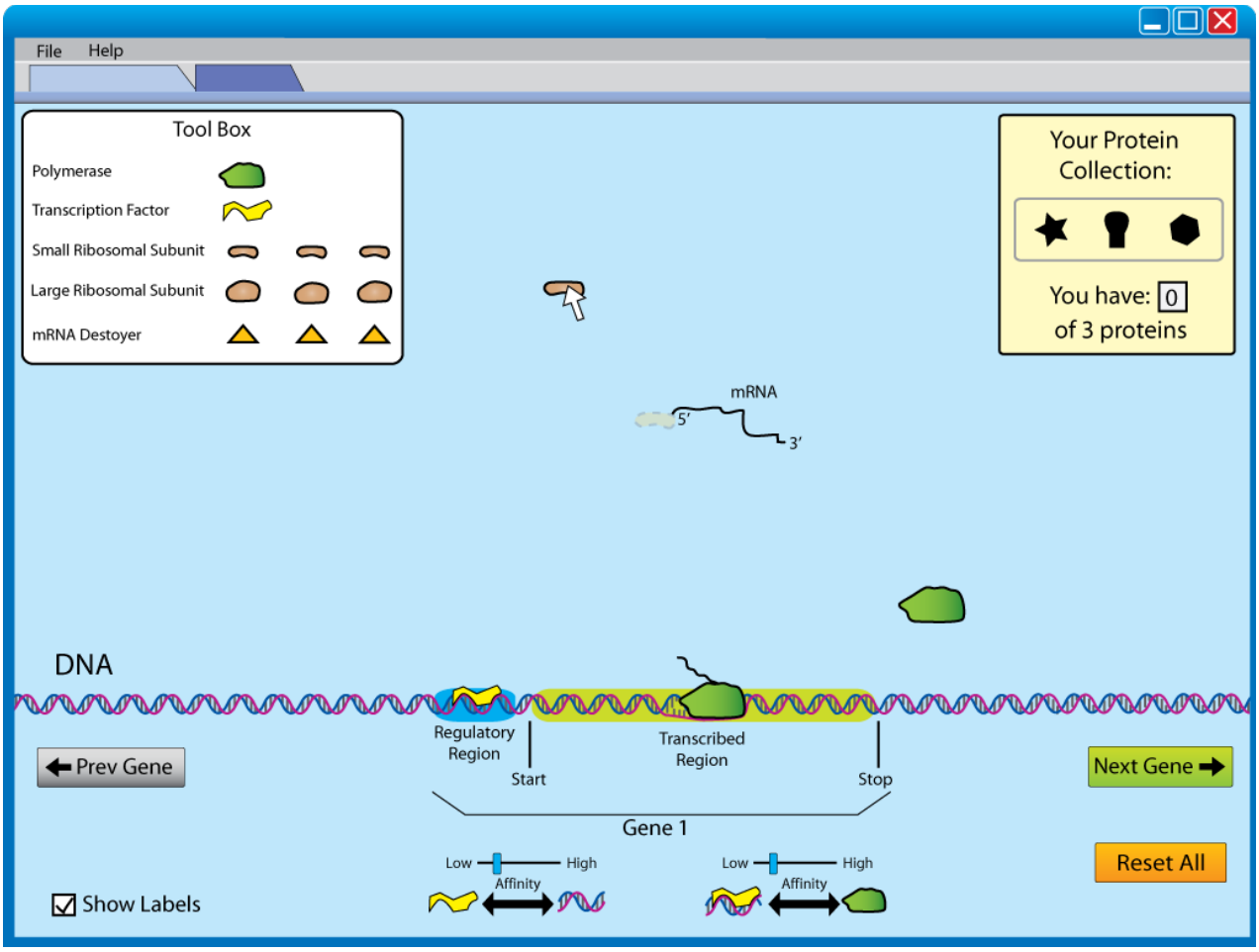
- The roles of the various proteins in the cells, i.e. why proteins are important.
- Non-coding sequences, aka “Junk DNA”. This in not junk, but simply non-coding, non-regulatory (as far as we know. Important to understand that many genes on a chromosome (i.e. Single dna molecule)
- Protein folding, other than just showing some folding into different shapes by the synthesized proteins.
- Processing of mRNA that occurs after transcription but before translation. This is important in eukaryotes, important to make simplification explicit (note it in the teaching tips).

Mockup

First Tab

The intent of the first tab is to convey some of the basic concepts and terms related to gene expression within a cell. The user sees a strand of DNA with areas that are marked and highlighted to indicate the location and basic structure (a.k.a. anatomy) of a gene. They can navigate to other genes by pressing the “Next Gene” or “Previous Gene” buttons. They can control the affinity of the transcription factor for the DNA and of the RNA polymerase for the combined DNA and transcription factor.

The user will need to manually move the various components, such as the polymerase and ribosome, around within the cell. They will get hints about where things should be placed by being shown “ghost” outlines at the location where the item may be placed.

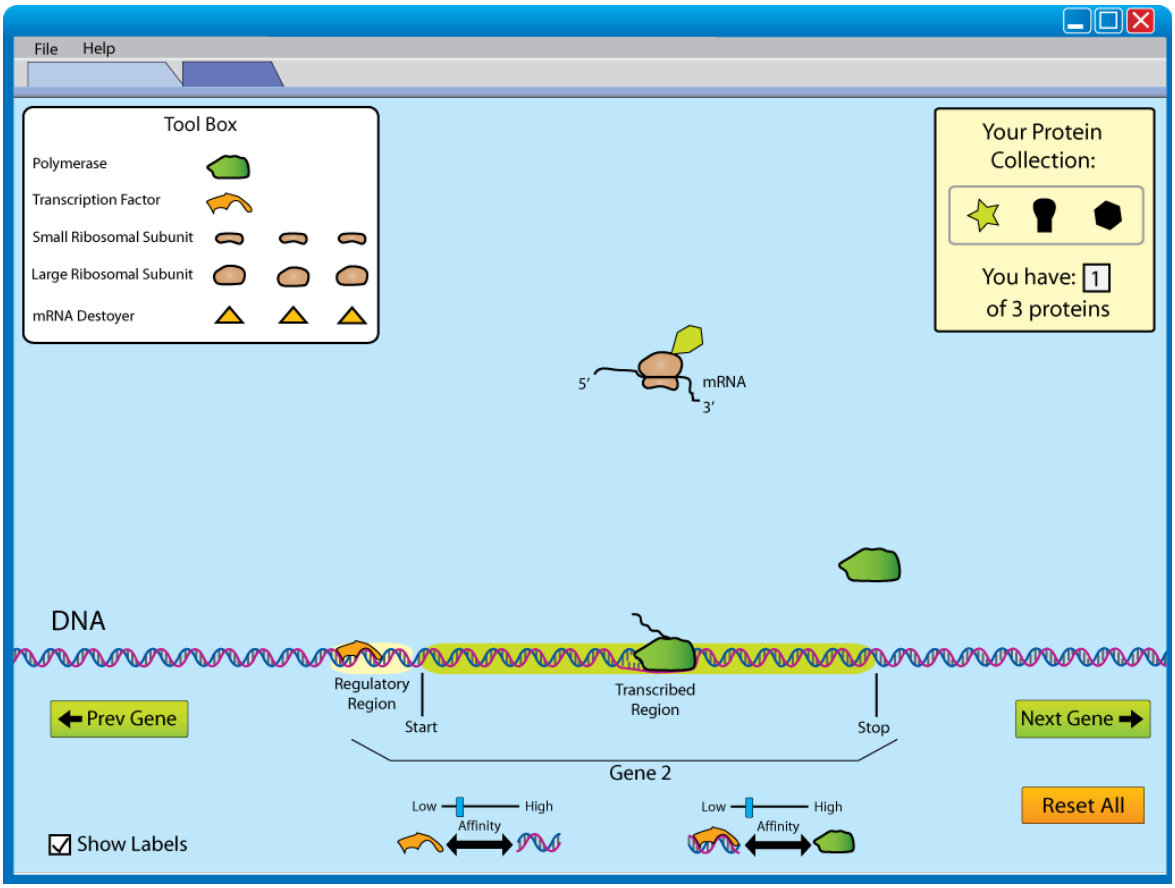


The drawing above depicts the first tab after the user has interacted with it for a while. They have figured out how to attach the transcription factor to the DNA and have placed a polymerase on the DNA, which subsequently transcribed the DNA into the mRNA shown in the middle. They also added a second polymerase (the one that is shown in the process of transcription). The user is shown holding a ribosome, and the sim is giving them a hint about where to put it.

Once the mRNA is transcribed, it will produce a protein shaped like one of the slots in the collection box. The user can drag the protein into the box once it has been fully produced.

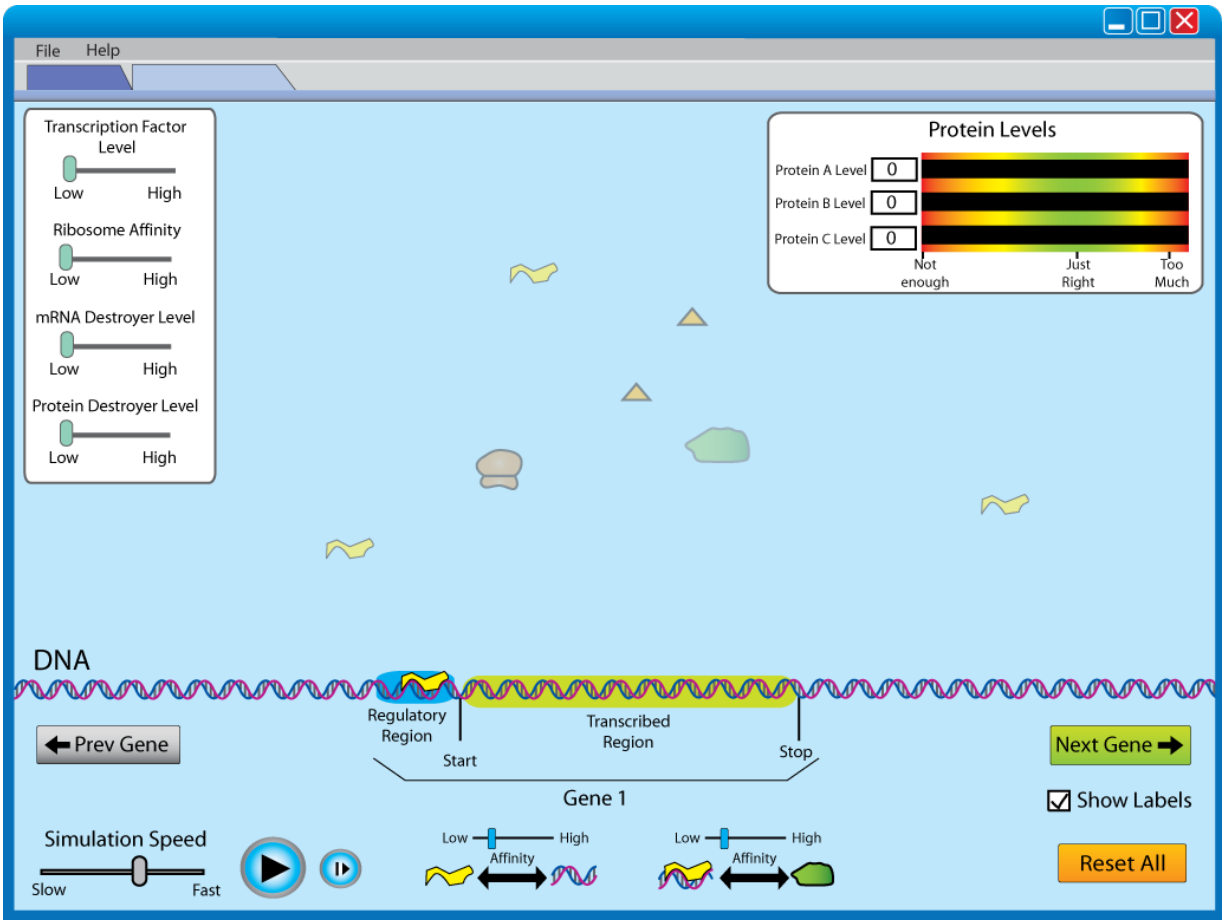
In order to create the other proteins, the user will have to navigate to each gene by using the “Next Gene” and “Previous Gene” buttons.

The drawing below depicts the situation where the user is part way through the process of populating the protein collection.



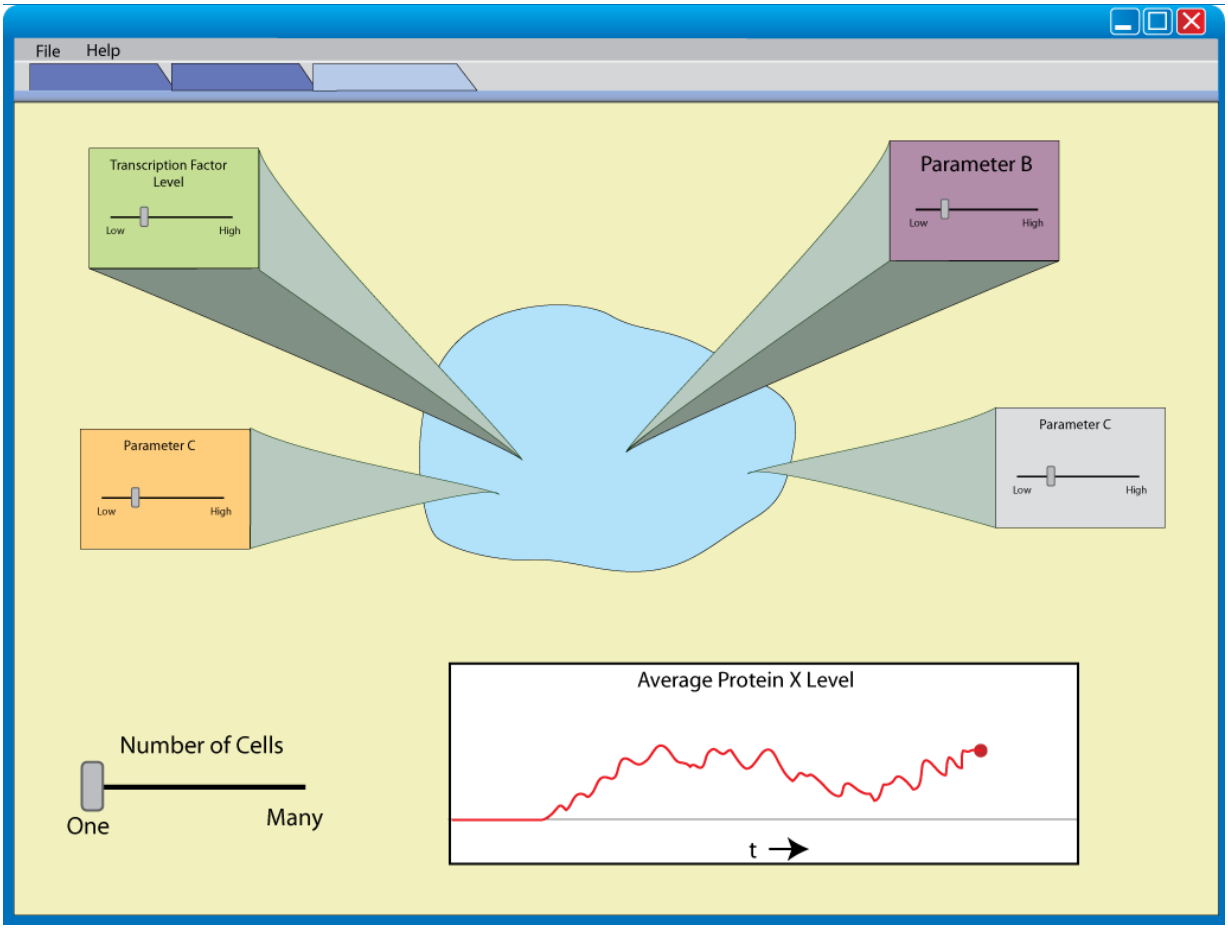
Second Tab

In the second tab, the user will control various parameters of a cell and see how those parameters affect the way that the various biomolecules interact with the DNA and with each other, and how this ultimately affects the levels of various proteins. The drawing below illustrates.

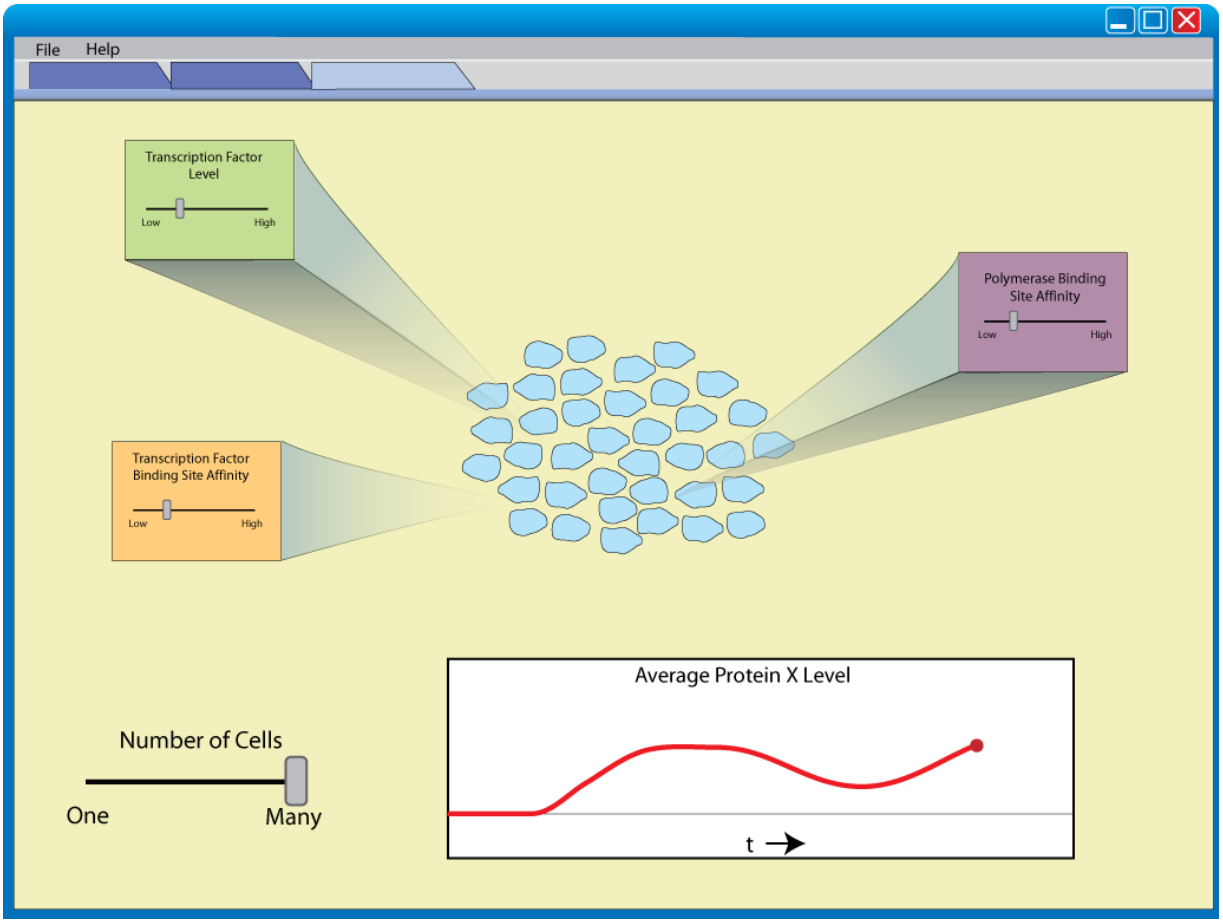

File
Help
Transcription Factor Level
Low
High
Ribosome Affinity
Low
High
mRNA Destroyer Level
Low
High
Protein Destroyer Level
Low
High
Protein Levels
Protein A Level
0
Protein B Level
0
Protein C Level
0
Not enough
Just Right
Too Much
DNA
Regulatory Region
Start
Transcribed Region
Stop
Gene 1
Prev Gene
Next Gene
Show Labels
Simulation Speed
Slow
Fast
Low
High
Affinity
Low
High
Affinity
Reset All

Third Tab

On the third tab, the user controls the levels of various parameters for a given cell from a perspective outside the cell. While they will be able to get close to a level, the value will always be changing, i.e. it will be 'noisy'. The drawing below illustrates this.



The user will be able to increase the number of cells and, as they do, they will see that the average protein level will be much more steady. The drawing below shows this.



Questions and Issues

- Will it be clear to users that what they are seeing in this simulation is occurring inside of a cell?
- We had talked about allowing the introduction of mutations at one point, but this point was was never fully resolved. Do we want to have this capability?

Things to Note in the Teacher Tips

This section describes certain assumptions and simplifications that were made that should be noted somewhere that is easy for users and potential users to find. The Teacher Tips may be the most appropriate place, but there may be other places as well.

- The genes depicted in the simulation are substantially shorter in terms of the number of base pairs they contain than real genes. On average, a real gene in a prokaryotic organism is 1000 base pairs, whereas the genes in this simulation contain less than 100 base pairs. (The numbers in this statement should be verified with Mike Klymkowski or some other knowledgeable person).

- The ribosome is added to the active area from the tool box as one unit. In reality, the two portions that make up the ribosome exist separately in the cell's cytoplasm and assemble at the start of mRNA transcription, and then disassemble again.
- The sizes and the shapes that are depicted for the various biomolecules were inspired by drawings in "The Machinery of Life" by David S. Goodsell.

Suggestions for Activities

- Suggested probing question: What would happen if there were a lot more ribosomes?
- Suggested probing question: What if a polymerase were transcribing DNA and it encountered another polymerase moving in the opposite direction?
- Suggested probing question: How many different ways can you regulate gene expression?

Prior Art

This section contains links to and descriptions of resources that are already available on line (as of 2/28/2011).

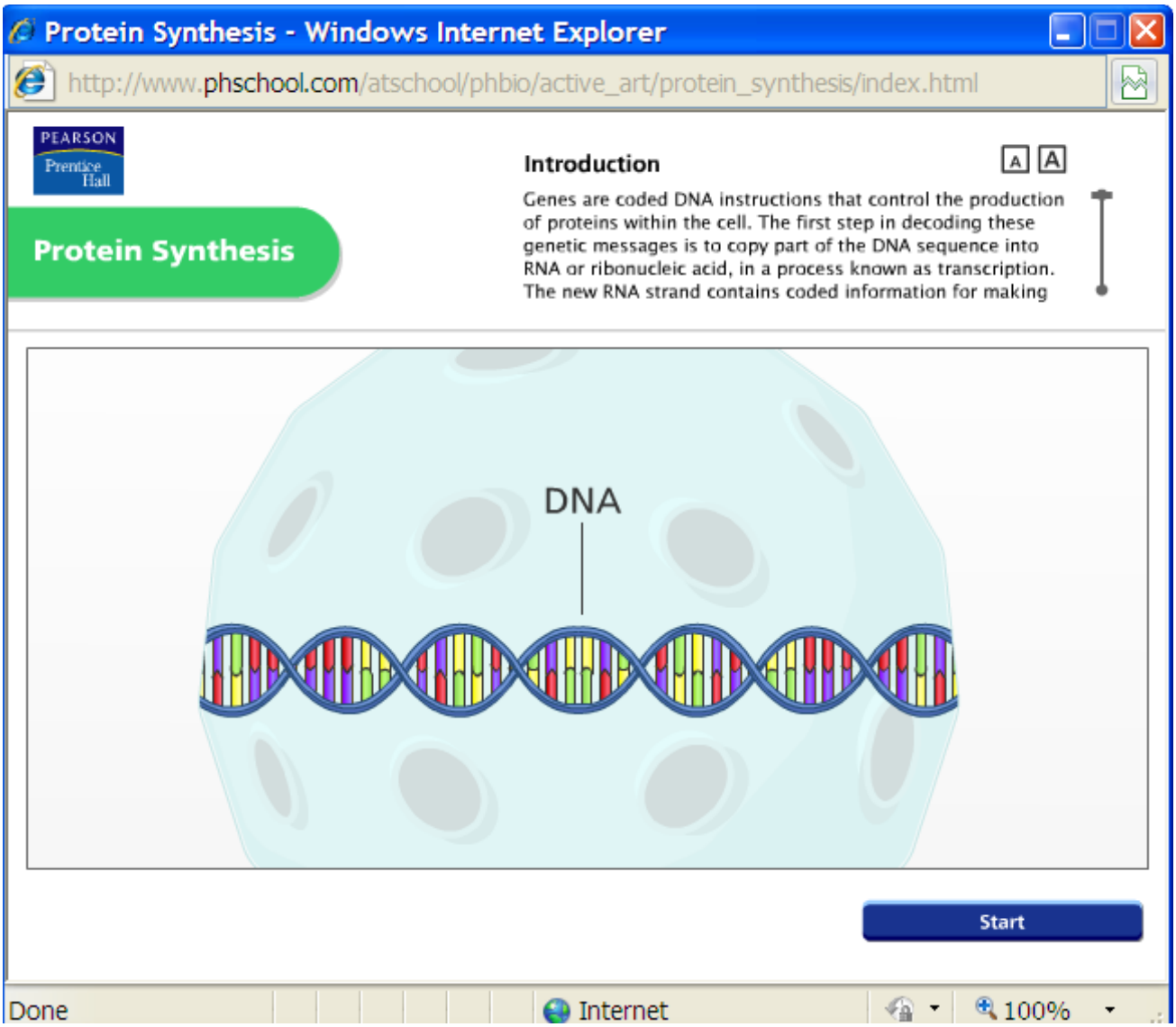
DNA Interactive

This web site contains a bunch of Flash resources - mostly animations but some interactive things as well. Jenny Knight brought this to our attention. Seems quite good.

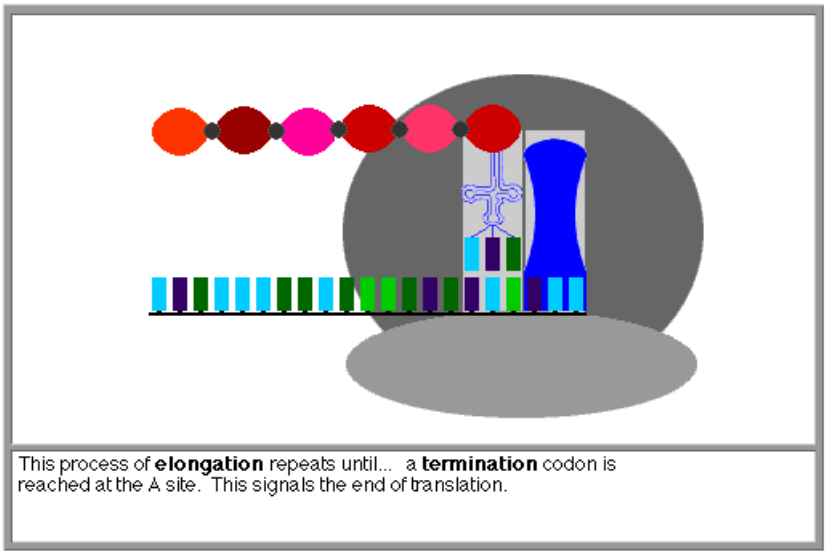
http://www.dnai.org/

Protein Synthesis on phschool.com

Trish Loeblein asked the AP Biology teacher at her school for some information on "hot topics" in biology, and he loaned her a text and pointed her to an on-line video that he thinks really helps his students. One can get to it by going to www.phschool.com, entering the code cbp-4123, and searching on "protein synthesis". This seems to be a Shockwave resource. Here is what the start screen looks like:



The following is a snapshot of a simple Shockwave animation of translation found at http://www.fed.cuhk.edu.hk/~johnson/teaching/genetics/animations/translation.htm



The following is a link to a site where the user interactively creates an mRNA sequence, then manually drags amino acids from a codon table on the to the RNA to create the protein. You create a protein with 7 amino acids.

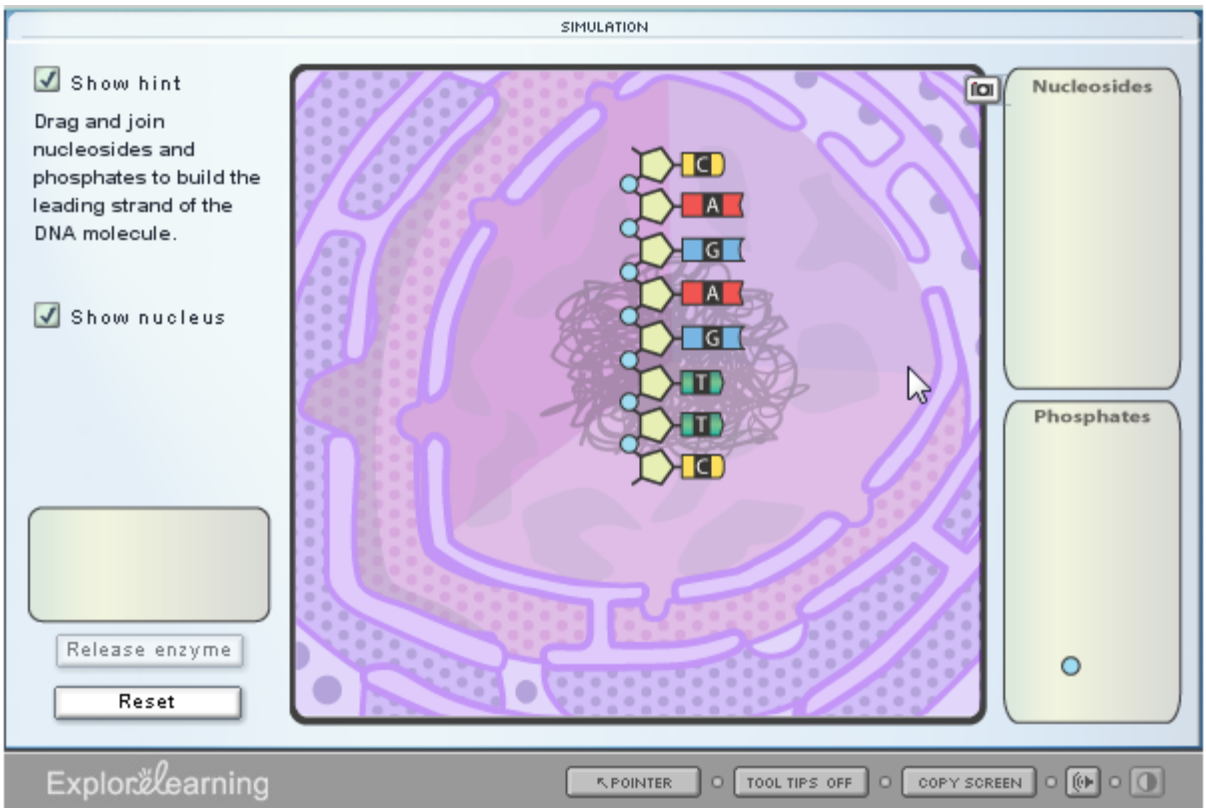
http://learn.genetics.utah.edu/content/begin/dna/transcribe/

Another resources with some interactive aspects. Mostly, though, it is more of a slide show type of thing.
http://www.wiley.com/college/boyer/0470003790/animations/translation/translation.htm

This is a page with a list of Flash resources for teaching about protein synthesis and DNA Genetics. Some of these are decent.

Gizmos

Gizmos has a very simple simulation where the users drag bases out, attach a phosphate, and make a chain. That's it. No pairing of bases or folding or anything. It has a background drawing of a nucleus, but there is no interaction with this.



Interviews

Gene Expression Basics (0.00.19)
April 13, 2012

Student Background
College freshman, male, no college level biology, some PhET experience (CCK)

Priority Fixes:

The student had appeared to have some difficultly mapping the different biomolecules to their respective shapes once they were all removed from the toolbox. I suggested to John B. they he leave a faded version in the tool box as a legend, he has already implemented this change in the latest version and it looks great.

For Discussion:

It might be useful to make the “delta” for the placement of transcription factors and such a little larger…a couple of times the student had difficulty placing the transcription factors and getting them to stick on the 1st tab.

Reflections:

From a user-interface point of view, the sim appears to be completely intuitive and easy to use. The student found all controls and manipulated them with ease. Although it took the student a bit of time to understand the learning goals of the sim, he eventually discovered all the important features of the sim, including understanding that different genes make different proteins (the collection boxes worked well), as well as seeing how the various transcription factors allowed or blocked the action of the RNA polymerase. On the second tab, he again seemed to understand all the main points and how the affinity and such affected mRNA production. On the final tab, it took him a bit of time, but again he eventually understood how each control related to the average protein level. He correctly interpreted affinity as “sticking” to the gene, and specifically mentioned “fluctuating levels”, but he did not clearly articulate the idea of a stochastic process.

Overall, I think the sim is in publishable form after a few minor tweaks (one of which John has already implemented).

Gene Expression Basics (0.00.22)
June 20, 2012

Student Background
College freshman, female, bio for engineers, no PhET experience

Priority Fixes:
The “snap” area for the ribosome should be increased in tolerance

For Discussion:

The zoom feature should be reconsidered (currently the thought was to take it out of the main sim and make it a teacher control)

The “real cells” picture, we decided should have a slightly different caption more focused on the learning goals.

Reflections:

The sim was very easy for the student to use and interpret. The student generated her first protein in less than one minute from starting the sim and collected the protein almost immediately after creating it. She did explore for a bit before collecting all the proteins (which seems fine), and used the reset all before all proteins had been collected (she had only collected one at that point). However in less than 10 minutes of using the sim, she had collected all proteins, after having come back to the first tab. In addition, she found and used all controls (including second tab “negative transcription factor”) and transitioned several times between tabs 2 and 3 to better understand the learning goals.

Overall, the sim seems clear, intuitive, and focused on the learning goals. This student even seemed to grasp the stochastic nature of the many cells vs. the one cell behavior in the average protein grasp.

With the changes mentioned above implemented, the sim appears ready for publication.

Meeting Notes

Date: 1/3/2012

Attendees: Mike Klymkowsky, Emily Moore, John Blanco

General Notes

- Version reviewed: 0.00.17
- Emily showed interviews of two students to Mike, and we spent a fair amount to time discussing the students’ interaction with the simulation.
- Several suggestions for activities came up. These have been entered into the “activities” section as well as here.

Activity Suggestions

- Suggested probing question: What would happen if there were a lot more ribosomes?
- Suggested probing question: What if a polymerase were transcribing DNA and it encountered another polymerase moving in the opposite direction?
- Suggested probing question: How many different ways can you regulate gene expression?

First Tab Feedback

- Would like to try starting zoomed out, see how students respond in interviews.
- The protein that is now colored to by fluorescent green should be barrel shaped, since that is the shape of the real GFP.
- Should allow biomolecules to go below the DNA as well as above. This is less likely to create some sort of misunderstanding about the spatial relationship between the molecules and the DNA strand.
- We should consider taking out the negative transcription factors. May want to just do some interviews with and some without.
- When both transcription factors are present in the regulatory region of the gene, they should change shape.
- Revised 1D walk looks good (MK later came back and said that biomolecules should be attached a little longer than in the reviewed version).
- Relative size of the ribosome is acceptable. Still somewhat troubling that it comes "pre-assembled" from the tool box.
- Label on mRNA is good.
- Different colors for non-fluorescent proteins is good.

Second Tab Feedback

- Molecules should be allowed to move all over the place, not just above the DNA strand.
- Max concentration of transcription factors is too high, should be maybe half of the shown amount.

Third Tab Feedback

- Play/pause button looks good.

Date: 2/13/2012

Attendees: Mike Klymkowsky, John Blanco

First Tab

- MK quite likes the zoom out capability on the 1st tab.
- 1D walk on 1st tab should be more obvious and should not last as long.
- We should have multiple transcription factors for the 2nd and 3rd genes.
- Transcription factors should not be in exactly the same place, but should overlap.
- With multiple transcription factors, there doesn't need to be any order dependence.
- OK to kick off biomolecules that get in the way of polymerase when transcribing.
- Ribosomes should be larger. In reality, they are roughly 5x as massive as polymerase.
- Protein growth is acceptable, but not ideal. If we could make it do some kind of folding, that would be better.

Second Tab

- Overall design looks good.
- MK suggests hiding the negative transcription factor initially to keep things simpler, and having a button or something to show it for a sort of “advanced mode”.
- Meter should not have “Too Little”, “Too Much”, and “Just Right” markings. It implies that there are levels that are correct, and there aren’t. It should just have “none” and “lots”.
- Affinity should only affect how long a biomolecule stays connected to the DNA strand, not the ease with which it connects.

Third Tab

- MK is okay with the idea of a visibility meter.
- Okay to use a log scale on the 3rd tab to make the noisiness more apparent but still have max protein level stay on the chart.

Date: 1/13/2012

Attendees: Mike Klymkowsky, John Blanco

3rd tab (Multiple Cells)

- Suggest darker background so that fluorescence is clearer.
- No spots - entire cell should glow - there are lots of individual protein molecules, they don't aggregate, and they move.
- Cells should be oblate spheroid, all same diameter, different lengths and orientation. We're basically saying that they are Coli. Irregular roundish shape currently used isn't very real.

1st Tab (Cell Gene Expression)

- Suggestion: Zoom out would be clearer if there is just a single cell.
- No organelles necessary.
- Maybe make the cells look more 3D.

Date: 1/6/2012

Attendees: Mike Klymkowsky, George Emanuel, John Blanco

3rd tab (Multiple Cells):

- Make the cell backgrounds be gray so that the color change releated to fluorescence is more apparent.
- Do whatever else is possible to make the fluorescence more apparent.
- Add a parameter: Protein Stability. We discussed the name for this, considering "Protein Half Life" and some other alternatives, and settled on "Protein Stability".
- When protein stability is high, there should be no flickering, when low, flicker, it would be great.

1st tab (Manual Gene Expression):

- MK felt that overall it looks good.
- MK is OK with the idea of not having affinity sliders on the first tab - said "it's already a bit busy".
- Background for tab one should be gray too.
- No need to try to depict any protein folding. It's kind a can of worms. Having them just "inflate" as they do now seem reasonable.

General

- Overall looking good.
- MK is teaching a class this semester, would be interested in incorporating the simulation. Would use in mid to late January, so it would be good to have a stable prototype version available then.

Date: 11/3/2011

Attendees: Tom Perkins, Kathy Perkins, John Blanco

- Suggestion: make the mRNA destroyer look a bit Pacman-ish.
- Make motion more random walk, a bit faster, longer free path, doesn't gravitate towards the DNA.
- Polymerase can be pulled to DNA from within one or two nanometers (about a diameter of the polymerase).
- 3rd tab: Have the cells change color as the level changes so that they can see the differences in individual cells. This would somewhat emulate the production of a flourescent protein.
- Idea on 3rd tab: Make the line on the graph change color as the concentration of the represented protein goes up and down so that it matches the colors portrayed by the cells. This would help connect the idea that the color and the concentration are correlated.
- Tom will have some time to put into this in a few days, and suspects that he will have something to discuss in about two weeks from when he starts on it.

- Tom may want to meet at some point with Mike Klymkowsky to make sure that we are all on the same page.

Date: 9/9/2011

Attendees: MK & JB

- Version reviewed: 0.08, link: http://www.colorado.edu/physics/phet/dev/gene-expression-basics/0.00.08/
- Q: How is the 1D random walk on the DNA?
    - A: Good, but it would be better if the molecules stayed on the DNA a little bit longer so that they could possibly traverse farther.
- Recommendation: Prevent overlap of biomolecules on the DNA. Can't have two in the same place at the same time on the DNA. Need to prevent with transcription factors, polymerase.
- Recommendation: Have all transcript factors available in every tool box, user has to find the right one for the particular gene.
- Q: How is conformational change of polymerase? There are many other options we could review.
    - A: Looks good as is.
- Bug: Doing a reset after transcription factor was present, then putting polymerase on, it will transcribe.
- Recommendation: mRNA should fold up, leaving the 5' end to where the ribosome can get to it.
- Q: What should happen when users pull the polymerase off the DNA when it is transcribing? Should we even allow them to do this?
    - A: Don't allow users to do this. It is a "non-physical event".
- Q: How is the transcription speed?
    - A: Good.
- Recommendation: If transcription starts and there is another non-transcribing polymerase on the DNA, the non-transcribing one should be bounced off.
- What should the first tab be named?
- Q: Should transcription factor fall off eventually?
    - A: Yes. About 20x to 50x longer than the random times.
- Recommendation: Make the mRNA a little thicker, and maybe "fuzzy" so that it looks sort of like if you blow it up you could see the bases.
- Q: How does the relationship between the transcription factors and the polymerase look? Because it can be tweaked a lot.
    - A: Looks fine as is.

Date: 8/19/2011

Attendees: MK & JB

- Dev version reviewed: http://www.colorado.edu/physics/phet/dev/gene-expression-basics/0.00.05/
- Q: How is the wandering motion?
    - A: Looks good.
- Q: How is the attachment to the DNA (in non-specifc points only at current time)?
    - A: RNA Polymerase should to a "1D diffusion", which means that it should do a random walk back and forth a little bit and them come off. If it does this and hits its real binding site, it should take off.
- Recommendation: Transcription factor should be bound much longer when put on its target site.
- Q: How do transcription factors look?
    - A: Fine. As long as they are smaller than the polymerase, their shape doesn't matter much.
- Recommendation: Polymerase should change shape and/or color when it gets together with the transcription factor. Then it kicks off the transcription.
- Recommendation: The hint for polymerase should INCLUDE the positive transcription factor.
- Recommendation: Transcription factors should be at top of tool box.
- Q: Can we use a smaller title for positive and negative transcription factors?
    - A: Can use + and - for transcription factor names instead of words, should preceed name.
- Q: Where should mRNA come out of polyerase?
    - A: mRNA can come out the back side of the polymerase. Can leave extended or can start to fold up.
- Q: How big should the regulatory regions be?
    - A: Fine as they are.
- Q: Any ideas how we can convey the idea that the transcription factor must be in place for the polymerase to attach?
    - A: The hint for polymerase should show the transcription factor too. When the polymerase encounters the transcription factor on the DNA, it should change shape and then take off transcribing.
- Q: How should the hints for RNA destroyer work? Do we need them?
    - A: Hints can be at either end and it can be destroyed from either end.
- Q: What if they pick up a ribosome before there is any mRNA? We would give them no hints, right?
    - A: Yep.
- Q: Where should transcription factors and poloymerase go on the regulatory region?
    - A: Polymerase should be right at the boundary, because once it starts move it is transcribing. Probably makes sense to have the transcription factors to the left of that.

- Q: Should the biomolecules bump into each other? Right now they don't - they just overlap.
    - A: Overlap is fine. Give sense of 3D.
- Q: Where should protein come out of the ribosome?
    - A: Top is fine.

Date: 7/28/2011

Attendees: MK & JB

Q: Agreed that it is okay to have the users move things manually on the first tab in order to get things started?
A: Yes, but there will be some automatic movement, such as falling off of the gene, after initial adding.

Q: Should we have a different tool box for each gene? Implemented this - OK?
A: Yes.

Q: How many twists should we use per gene?
A: As many as we can reasonably do. 3000 base pairs, or 30 twists, is a realistic gene size. We will need to be smaller than this, so we should note it in the Teacher Tips document.

Q: Is "Tool Box" an appropriate title?
A: MK proposed "Molecular Tool Box"

Q: How is the size of the polymerase? Kathy thought maybe too small.
A: Yes, too small. See pg 46 of Goodsell for an illustration, let's do it to that scale.

Q: How is the size of the transcriptioni factor?
A: Should be bigger, but not as big as polymerase.

Q: Suggestion by PhET team: Polymerase like Lac Operon. OK? Kind of fit together?
A: MK doesn't like, due to the very geometric shape and the very different look from the one in Goodsell's book. JB will follow up on this with Noah P and Kathy P.

Q: Team wants ribosome to be one thing instead of the current two. How strongly do you feel?
A: OK to have it pre-assembled, but we should note that in the Teacher Tips document (or wherever it makes sense).

- Discussed zoom concept, MK really likes it because it provides context. Cell at a distance, then move in.
- Idea: Have things showing at the macro level that indicate protein levels, zoom in, change things, what happens?

- We should communicate clearly (in whatever place it makes sense) that we've made a lot of simplifications and assumptions.
- We should depict conformation change when polymerase binds to the right site to make it clear that something changes when this binding occurs.
- After further consideration, we should clearly say that this is bacteria, because it is too hard to be ambiguous. Early in the design we said that we should try to be ambiguous between prokaryotic vs eukaryotic.
- Add trascriptional activator and transcriptional repressor.
- MK is now questioning whether or not we want to be able change affinities on the first tab.
- Things should bind loosely to DNA if placed in the wrong place, then float off. This is because all of these molecules have a basic affinity for DNA.
- It would be good if the transcription factors fall off after a little while, even if they are in the right place.
- Discussion reinforced the idea that we should have both activator and repressor from the beginning.
- On 2nd tab we should allow them to control the affinities and levels of repressors and activators.
- MK: We should have fixed affinities on the first tab, changable on the 2nd tab. JB needs to think about how this will look on 2nd tab.
- Make the longest gene as wide as the screen.
- Panel 2: Thoughts: Zoom out a bit, specify gene length, affinities and concentrations for transcription factors, see how that affects concentration of proteins. Possible idea to introduce: Half life of the polypeptide.

Date: 6/28/2011

MK & JB attending

- Generic form of gene expression - this is what we want to do. It is not specific to pro- or eukaryotic cells.
- Mike is concerned that we are "going to get nit-picked to death" because we have to make so many choices about what to include, exclude, and simplify.
- We discussed having both positive and negative transcription factors, and Mike felt like it is okay to leave off negative transcription factors from the 1st tab, but we should perhaps include them in subsequent tabs.
- When asked what he thought the name of the sim should be, MK said "Generic Gene Expression Basics"
- MK had the suggestion that we should consider a sim on cooperation as a biochem concept, one that students could use to learn about the concept and how it applies to gene expression. In this context, multiple transcription factors can cooperate to create a situation in which the RNA polymerase can bind to the DNA.

- We should consider having allosteric factors in the later tabs. These factors change the transcription factors to make them more or less likely to bind to the DNA, so they affect the overall rate of transcription and thus protein production.
- MK is not particularly fond of the idea of having users manually move the various components around manually in the first tab, since this may potentially be misleading. It may create the idea that something moves things around, rather than the idea that they bounce around randomly and bond/align based on this motion. But it is hard to envision an alternative that is interactive, so is okay with it for now.
- The ribosome is currently drawn in the first tab as having both parts pre-assembled. This is not correct. The mRNA connects first with the Small Ribosomal Subunit (SRS) and then this complex connects with the Large Ribosomal Subunit. That is when translation occurs. We should have these two units separated and have the user assemble them. JB will update the drawings accordingly.
- There were some hand-written notes from this meeting that were scanned and put into SVN.

Date: 6/21/2011

Attendees: MK, JB

Meeting was short due to a meeting that MK needed to join.

- We reviewed the current version of the sim, and the DNA depiction and navigation looked good to MK.
- Parameters that can be controlled on the 3rd tab should be:
    - Transcription Factor Concentration
    - Transcription Factor Binding Site Affinity
    - Polymerase Binding Site Affinity

Date: 6/3/2011

Attendees: MK, JB, KP

- Tab 1: We should have different transcription factors for the different genes. It is an open question as to how to allow the users to manipulate the affinity and figure out which transcription factor goes with which gene.
- There are 10 base pairs per twist of DNA in real life. We should try to model this if possible.
- Different regulatory regions should have different colors so that it is clear to users that they are not the same, i.e. the base pairs are different.
- A real transcription factor would bind to about 10 base pairs, so it is about 1 twist. Current one is too small, we should make it larger.

- The minimum size for a real gene is roughly 100 base pairs (10 twists). We should do ones with 100, 150, & 200.
- MK: There a LOT of ribosomes in a cell.
- Ribosomes are bigger than polymerase. John or Kathy should look up the relative scale in order to come up with something reasonable, since in the current drawings the polymerase is bigger.
- MK: It is OK to not depict tRNA at all, and have protein just comes out of the ribosome as the mRNA is processed. Amino acids will also not be shown prior to being used as a piece of a peptide.
- The graph in the 3rd tab should depict the average level when there are lots of cells, not the total.
- MK: "An important concept is that stochastic processes can produce meaningful behavior." We discussed whether this should be an explicit learning goal, and decided that it was perhaps too advanced. We should probably have a related, but simpler, goal, such as: Probability is important in the inner workings of a cell.
- One thing that we want people to get is that there is a time delay between change of the various factors, such as the amount of a transcription factor, and the response in terms of the quantity of a given protein.
- We discussed whether we want to include mutations in this sim, and decided against.
- Kathy will talk to her brother for the model used in the 3rd tab. Mike feels that he could do it, but Tom has more directly relevant experience.
- Since there is no separation depicted between the DNA and the ribosomes can start before mRNA fully translated. This is how it works in prokaryotic cells.

Date: 5/31/2011

Attendees: Kathy P and John B

The role of competition and randomness are key learning goals for this sim.

Codons, individual bases, and base pairs will not be depicted.

The affinity sliders currently shown on tab 1 should also exist on tab 2.

The post-processing of mRNA (which I believe has a more specific term, but I don't know what it is) is NOT a learning goal, and will not be depicted.

The secondary structure of mRNA, which is the folding and joining that it does, is not a learning goal and will not be depicted.

Learning Goal: Understand that the levels of a protein in a given cell will have some randomness, but with large numbers of those cells, the randomness evens out (or washes out) and becomes steady once the levels become somewhat stable.

Tab 2 - Things are faded looking and come forward as they become relevant to whatever is being shown, such as transcription.

Tab 3 - Users can control the number of cells and several parameters, and can see the levels of certain proteins (or whatever) change as parameters are changed, and can see that more cells evens out the levels.

Tab 3 - We will probably need some help on the model for this. Mike K may know what is needed. If not, perhaps Tom Perkins.

Kathy also made notes in a pencast notebook for this meeting.

Date: 3/23/2011

Attendees: Mike Klymkowski, Kathy Perkins, John Blanco

The full meeting was recorded on a pencast, which can be obtained from Kathy P or John B. There were several items worth noting:

Mike: “The goal here is to understand how regulating the transcription factor and regulating the stability of the RNA can get me a particular level (of mRNA or a partcular protein)”.

Mike: “Where do you start, and where do you simplify, and what is really important”?

In the latter portion of the meeting, Mike was advocating having a the first tab just show mRNA production and having the users try to get to the appropriate level. John and Kathy thought that this may not be interesting enough for the student, i.e. that they wouldn’t understand the reason to try to get to a given level and wouldn’t be incented to do it.

Decision, not sure when it was made: Do not distinguish between prokaryotic and eukaryotic cells.

Notes from 2/17/2011 meeting to discuss gene simulation

Attendees: Kathy Perkins, Jenny Knight, Jia, Nancy, John Blanco

Learning Goals from Jenny:

- Know the basic DNA structure

- Base pairing, which bases pair with which
- Take a gene sequence and turn it in to RNA and protein (using the codon table)
- If there is a mutation in the DNA, what is the impact on the RNA and the protein and why?
- Get students to understand a shift due to an insertion.
- Alternative splicing is important for every student to understand, since it is the reason why we can get away with fewer genes. Allows making various proteins from one gene.
- Enzymes do this. It reacts with the RNA. Splicing occurs at the RNA level.
- It is important for students to understand that there are portions of the DNA that are start sequences that transcription factors bind to.
- Directionality is huge.
- Initiation, elongation, termination are key.
- Sigma or transcription factors.
- Maybe learning goal: Understand that energy (in the form of ATP) is required for transcription.
- "anatomy of a gene" – promoter region is on the DNA, and the start codon is there also but is there mostly because the RNA needs it (I think).
- The "coding region" of the gene.
- The "processing" step – "intron removal"
- Nancy "base pairing is really important"

Consensu: Would be best to be generic and NOT worry about the difference between prokaryotic or eukaryotic.

Term: Anti-condon
Term: Codon is 3 RNA bases.
Term: tRNA
Question (for us): What cuts out the pieces (introns)?

Jia pointed out that students didn't understand that promoter was not part of the coding region.

Jenny teaches a human genetics class for non-majors.

Term: Codon matching

Note: Students often have confusion about U's and T's in DNA.

They mentioned 5 prime end and and 3 prime end several times. This has to do with directionality.

DNAi.org – Has an animation (non-interactive) that Jenny likes that shows the process.

DNA should "melt open" as the polymerase as it is transcribed.

Polymerase should be bigger than binding region.

Add a transcription factor to the tool box, and have them put it on DNA

Just color the binding region

Shouldn't be too lock and key-ish.

Rather than adjusting the polymerase level, do the transcription factor instead.

Term: AUG

Term: Methylation event or a histone modification.

They don't teach the affinity stuff at intro. They don't think that it's unimportant, just haven't done it. Jenny thinks the epigenetic stuff would be more important.

First tab could be much more detailed and allow them to learn more.