
ALTECH

Molecular Switches for Cellular Sensors

by Christina Smolke

Left: Undergrad Jack Lee, Smolke, and grad student Travis Bayer enjoy decaf, regular, and espresso laced with caffeine-sensing cells.

Right: Stem cells make many life choices en route to their eventual careers. After deciding to become a hematopoietic, or blood-forming, cell, various forks in the road lead to immune-system cells including natural killer (NK) cells, red blood cells (erythrocytes), platelet factories (megakaryocytes), and various types of white blood cells. (Graphic adapted from Eckfeldt, et al, *Nature Reviews Molecular Cell Biology*, Vol. 6, pp. 726-736, 2005.)

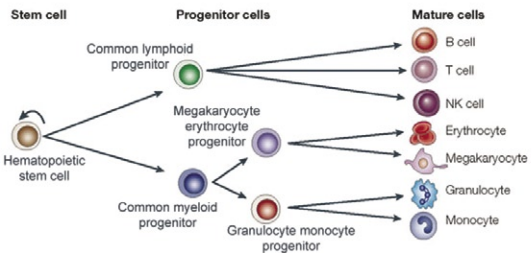
I work in a new field called synthetic biology, which is an amalgam of molecular biology, biochemistry, and control theory. And I'm actually a chemical engineer. Synthetic biologists try to design systems—cells—that will perform some sort of complex task. Now cells do complex things all the time, of course, but what makes synthetic biology different is that it emphasizes robust, predictable design—tiny cellular machines that, like mechanical ones, will reliably do what we want them to do. A vacuum cleaner always sucks up dirt, for example, but without adequate controls, the cellular equivalent might decide to in effect ingest dust bunnies one day and reheat frozen burritos the next.

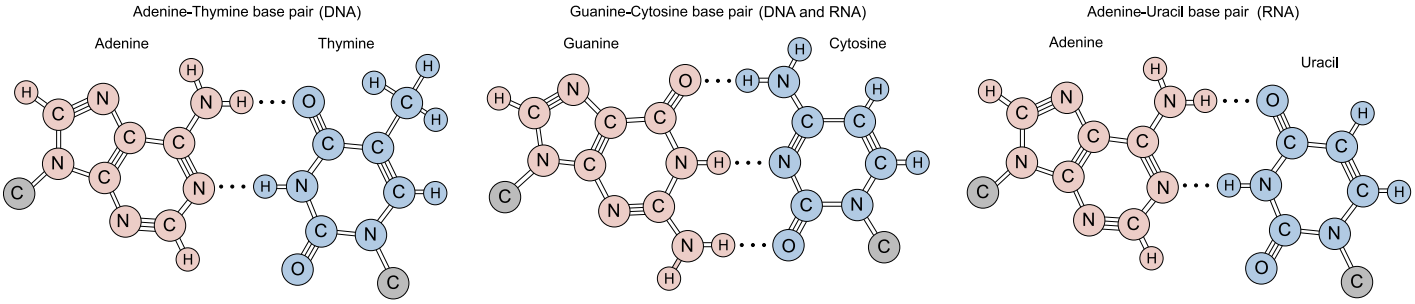
There are several kinds of tasks that we're interested in. In metabolic engineering, we reprogram a cell to produce a valuable compound, such as a pharmaceutical. Nature produces a wonderful array of medically useful molecules, but not always the ones we want in the quantity we'd like. It can also be very expensive and time-consuming to grow, harvest, and extract the natural product, but the molecules are frequently so complex that it is even more expensive and time-consuming to try to make them in factories. For example, the opium poppy produces morphine and codeine through a metabolic process that proceeds by way of several intermediate products, including (*S*)-reticuline, which is a molecule from which many potential anticancer and antimalarial drugs can easily be synthesized. But (*S*)-reticuline doesn't normally accumulate in the poppy, and shutting down a metabolic pathway partway through its course is a tricky proposition. For example, knocking out the gene for codeinone reductase, the final step in the path, actually shuts it down seven enzymatic steps upstream. Several intermediates accumulate, including (*S*)-reticuline, which then needs to be separated from the other intermediates and purified. This is in contrast to some simpler organisms, such as bacteria and yeast, where you can knock out a gene anywhere along a pathway and—assuming this action doesn't kill the cell—accumulate whatever substance was produced in the previous step. So if we could reconstruct the pathway of interest in some tractable microorganism, we might be able to make it produce bulk quantities of anything we want. With the sophisticated genetic-engineering tools at our disposal and an artificially constructed pathway, we'd be better able to control the production process and isolate our chosen substance.

We could also reprogram a diseased or problematic cell through "intelligent molecular therapeutics." That is, we'd design molecules that could identify the cell that they're in and then do something based on that identification. So, for example, if the molecule determines that it is in a cancer cell, it would rewire that cell's aberrant metabolism to make it behave like a normal cell.

Alternatively, we could design cellular biosensors, where the molecule would make the cell produce a detectable signal, such as fluorescence or luminescence, which would then be read by a machine. Such biosensor molecules could be used for enhancing our understanding of the key pathways that regulate important cellular functions (or make codeine!), or in the early detection and diagnosis of diseases.

And finally, we could reprogram a cell's entire life choice, not just some facet of its metabolism. There's been a lot of discussion about stem cells recently in the media. Stem cells are undifferentiated—that is, they have the potential to become any of many types of cells. So a completely undifferentiated stem cell first chooses to become one of several general types





Above: The bases in DNA and RNA recognize one another by forming hydrogen bonds (dotted lines). The gray carbon atoms are part of the backbone chain on which the bases are strung.

of cell—nerve cells, blood cells, liver cells, and so on—and then once it decides to be, say, a blood cell, it makes choices from progressively narrower sets of options until it reaches a particular sub-classification such as a T-lymphocyte, which is a specific type of white blood cell. Cells have natural preferences for certain choices at various forks in these pathways, so if we can figure out which molecules actually make those decisions, we could try to influence the choices. We could even make a cell decide to kill itself—programmed cell death, or apoptosis, is a choice that a surprising number of cells make in every developing embryo. If we found a cancer cell, for instance, that was too far gone to reprogram, we could simply shut it down altogether.

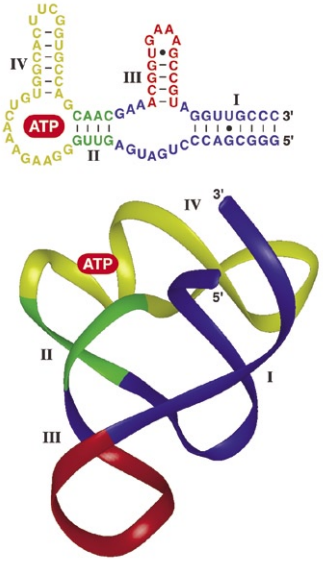
The cell's behavior is a property of the molecules that are within it at that moment, so the first thing to do is take inventory—what is the global set of proteins (and other biomolecules) that results in this particular behavior? And once we've identified all those proteins, what are the interactions, or the links, between them? One set of proteins will interact with another set of proteins that interacts with the next set of proteins which goes on to interact with other proteins, and eventually the cell winds up doing something. These interactions are the moving parts of the machine—the cogs, cams, and flywheels—and if we want to rebuild the machine to do something else, we need to trace their motions to determine what each part does.

But what controls the machine? For each protein, there is a gene, and when the gene is turned ON, the protein is produced. The gene is made up of DNA, which encodes the blueprint for that protein as well as instructions for when it should be produced, and in what quantity, depending on the cell's environment. These instructions are the buttons on the control panel, if you will, and their interplay is the wiring diagram. Once the gene is turned ON, the cell reads the blueprint through the medium of an intermediate molecule, called messenger RNA, via a process called transcription. And the messenger RNA instructs the cell's machinery to make the protein.

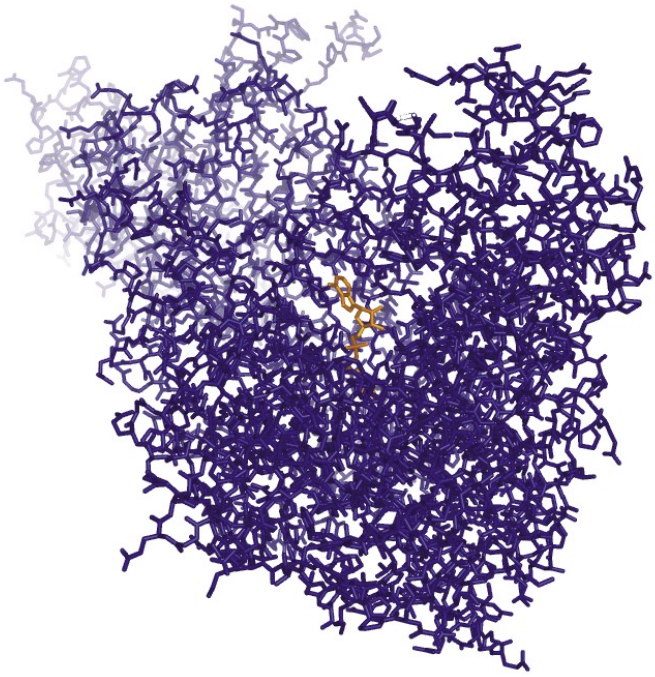
DNA and RNA are nucleic acids—a completely different type of molecule from proteins. They're made up of four different building blocks, called bases—adenine (A), guanine (G), thymine (T), and cytosine (C), with uracil (U) instead of thymine in RNA—strung together like pearls on a necklace. Inside cells, DNA normally exists as two strands that are bound to each other by interactions between the bases, like the meshing of teeth in a zipper. The bases recognize one another, so that T always bind to A, and G always pairs up with C. So, for instance, if I tell you that one strand of DNA has the sequence AGTC, you know immediately the complementary sequence—TCAG—that's going to bind to that strand. RNA is generally a single-stranded molecule, but its bases interact in the same way, with U being complementary to A. RNA molecules can bind to themselves, with parts of the molecule forming railroad-track structures called stems, often capped with little protruding knobs called loops. Ultimately, the whole molecule coils up, twisting and knotting like an unruly telephone cord, as does DNA. The sequence of bases in RNA or DNA is called the primary structure. The way the bases associate with one another forms the secondary structure, and the wadded-up tangle that results is called the tertiary structure.

Nucleic acids have traditionally been viewed as passive molecules within the cell. They stored genetic information, or they acted as intermediaries that transported it, but they didn't really *do* anything by themselves. But this turns out to be a very limited view. In the past couple of decades, nucleic

Far right: An RNA molecule's primary structure is its sequence of letters (top); some of the letters bind to one another to form its secondary structure of stems and loops. (The dots represent "wobble pairs"—slight mismatches that distort the molecular backbone.) The secondary structure kinks and twists to form the tertiary structure (bottom), shown as a ribbon. The colors and roman numerals mark various "domains" that actually do things—domain IV recognizes and binds to the adenosine triphosphate or ATP molecule, for example.



Adapted from Soukup and Breaker, *Trends in Biotechnology*, Vol. 17, No. 12, pp. 469-476, 1999.



There are about 300 amino acids in this protein, a tRNA synthetase, which recognizes and binds to phenylalanine, shown in orange.

acids have been found to have a number of very interesting functions. They really *do* do things, and we are exploiting these functions to design molecules to perform functions of our own choosing.

First, nucleic acids can exhibit catalytic activity: they can perform reactions, which is traditionally the province of proteins. RNA turns out to be very good at cutting apart other pieces of RNA. The reverse of a cleavage reaction is a ligation reaction, in which the RNA joins nucleic acids together, and RNA is very good at that as well. RNA has, in fact, been found to catalyze a large number of different types of reactions, leading some scientists to propose the existence of an "RNA world" on the early Earth, before the advent of DNA and proteins, in which RNA alone carried out all the business of life. So catalytic activity is a very powerful property with many uses.

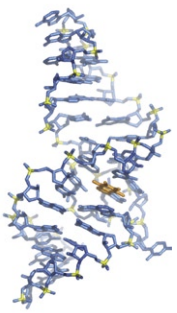
Second, nucleic acids can also act as regulatory elements. Remember, DNA encodes genetic information that is transcribed to messenger RNA, which is read, or "translated," by the cell. Meanwhile, scavenger proteins are destroying the RNA, preventing the cell's machinery from getting stuck in overdrive. So the amount of protein being produced at any given time is a balance between the competing rates of transcription, translation, and decay, and the cell modulates the fluxes between these different pathways to control the amount of protein that's produced.

Recently, it's been discovered that "trans-acting RNA" molecules—small RNA molecules that do not code for any protein—actually regulate protein production. They carry a complementary sequence of bases that allows them to bind to the messenger RNA. Because the messenger RNA makes sense to the cell's machinery, these strands of RNA are called "antisense" strands. Some antisense RNAs simply impede the translation of the messenger RNA—like trying to feed too many sheets of paper into a printer at once, they jam up the machinery. Others actually increase the messenger RNA's decay rate by flagging the molecule for destruction. Either way, less protein is produced.

There's another class of RNA regulatory elements called "cis-acting" molecules. These are actually parts of the messenger RNA molecule itself. They don't contain any of the code for the protein molecule, but they have a well-defined secondary structure—oftentimes some variety of stem-loop structure. This stem-loop structure forms a tertiary structure that interacts with other biomolecules in the cell to modulate the relative rates of transcription, decay, and translation of the messenger RNA to which it belongs. Each RNA molecule usually has several cis-acting regions that respond to different stimuli.

But the final type of activity is the most exciting, and is the basis for a lot of the engineering work in my laboratory. Nucleic acids can actually act as sensors to detect and identify other molecules, which is another property that was typically only associated with proteins. An RNA (or DNA) molecule can fold back onto itself to form a tertiary structure that creates a binding site for a protein molecule in a very specific manner—in other words, it will recognize and bind to the latter. Such pockets can also recognize small molecules, like caffeine and other drugs, and medium-sized molecules, such as the lipids in the cell membrane. Thus nucleic acids have enormous potential as molecular sensors, with specificities and affinities rivaling that of protein-based sensors.

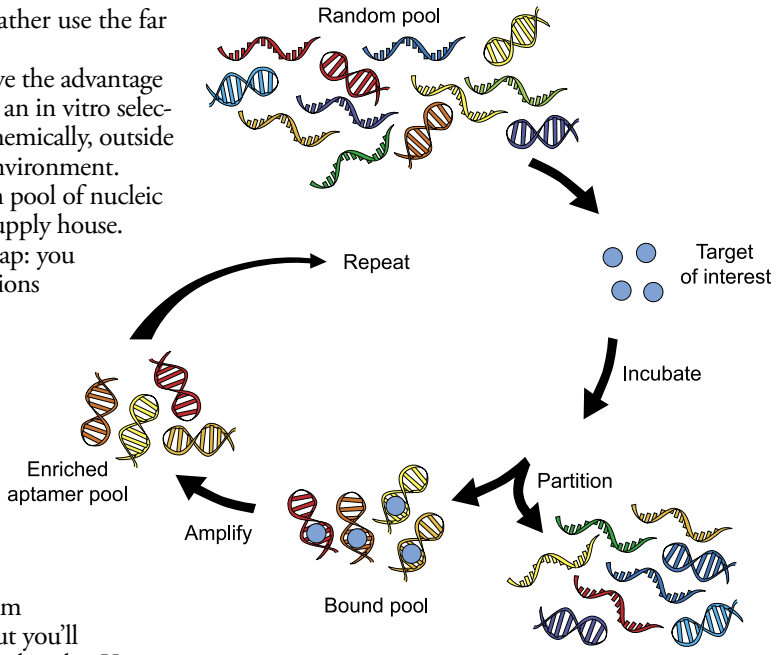
And nucleic acids have one huge advantage. Above left is the complete structure of a protein molecule, shown in blue, and a small biomolecule, shown in orange, to which it is binding. (The molecule being bound is called a ligand.) And below is a single-stranded RNA molecule that has twisted up to form a pocket that binds a ligand of similar size. The RNA likes to stack its base pairs in that famous double helix, and the ligand slips in between the pairs like a spatula sliding between flapjacks in a short stack. I don't know about you, but if I had to try to design one of these two



This much smaller strand of RNA contains about 30 nucleic acids, yet it recognizes and binds to theophylline, which is a molecule about the same size as phenylalanine.

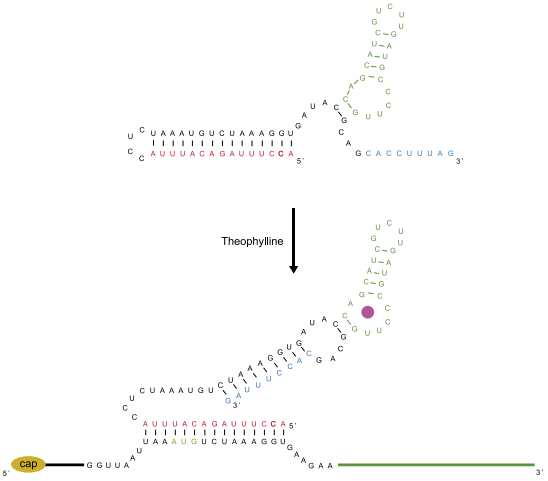
molecules from scratch, I'd much rather use the far simpler nucleic acid structure.

These nucleic acid sensors also have the advantage that you can generate them through an in vitro selection process. You can make them chemically, outside of cells, in a reasonably controlled environment. You basically start off with a random pool of nucleic acids, which you can order from a supply house. DNA synthesis is fairly easy and cheap: you just ask for all the possible permutations of, say, a sequence 40 base pairs in length. Using standard methods, you transform this random pool of DNA into the corresponding RNAs in your lab. Then you take whatever molecule you want the sensor to recognize—say, a viral protein that you want to use to detect infected cells—and you incubate it with this pool of random RNAs. Most of them won't bind to the target molecule, but you'll get a very small population of RNAs that do. You then fish those out, again by standard methods, and use them as the starting pool for the next cycle. Each cycle can take as long as a day—or at least several hours—to complete, and it usually takes eight to 15 cycles to get a good result. (My lab is working to get this down to one to three cycles of a couple of hours each.) In any case, you eventually wind up with a very selective, high-affinity pool of aptamers—nucleic acid structures that bind to the target. Then you decide which is the best one for your purposes and incorporate it into your molecule.



Far right: How to find the perfect aptamer: Lather. Rinse. Repeat.

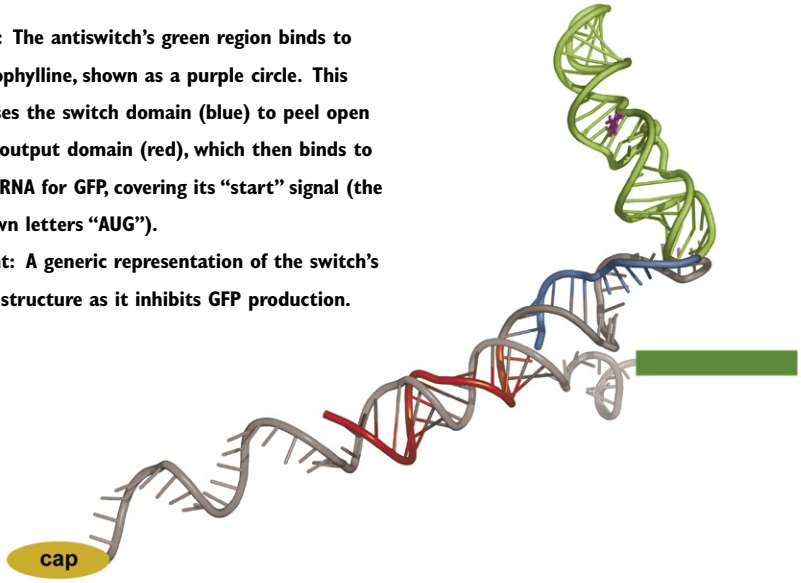
But a sensor is no good if you can't read its output. So we engineer RNAs that contain several different domains in each molecule, as you can see in the color-coded structures below. The sensor domain (green) is the winning aptamer from the talent search I described in the previous paragraph. This is linked through a switching domain (blue) to an output domain (red) that controls the production of some protein by the cell. This protein could generate a detectable signal—for instance, green fluorescent protein (GFP), which makes the cell emit green light when you excite it with a laser, is commonly used. GFP is popular because you don't have to disturb the cells in any way to sense its presence. You just hook up a video camera to your microscope, zap the cells, and watch them glow. Or the protein might direct the cell to change its behavior in some way—to stop dividing if it's a cancer cell, for example. Or the protein might stimulate the production of something we're interested in, like a pharmaceutical. The output can be digital—a very sharp response, basically ON/OFF or ONE/ZERO, meaning we've either detected the ligand or we haven't—or it can be analog, a graded response that increases



Adapted from Bayer and Smolke, *Nature Biotechnology*, Volume 23, Number 3, pp. 337-343, 2005.

Left: The antiswitch's green region binds to theophylline, shown as a purple circle. This causes the switch domain (blue) to peel open the output domain (red), which then binds to the RNA for GFP, covering its "start" signal (the brown letters "AUG").

Right: A generic representation of the switch's 3-D structure as it inhibits GFP production.



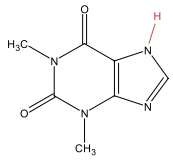
in proportion to the concentration of the molecule being sensed. We're working on both, but the one I'll describe here is the digital version.

We call it the antiswitch, because the output domain is an antisense, trans-acting RNA domain. But we could also call it the antiswitch because it works backward, in a way. The antisense domain is designed to bind to a messenger RNA and keep it from being read by the protein-producing machinery. But when there is no ligand present, the RNA loops back upon itself like a bobby pin and the antisense domain is actually bound to another part of the molecule containing the complementary sequence. So the antisense domain is all tied up, and can't bind to the messenger RNA and shut it down. The target messenger RNA I've shown here produces GFP, so that when the cell fluoresces, the detection value is ZERO. (In order for this to happen, of course, the cell must have been reprogrammed to produce GFP by default, but fortunately that's a well-known procedure.)

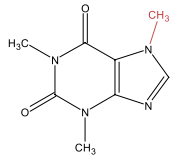
When the ligand slips into its binding pocket between the neatly stacked base pairs, something really interesting happens. The RNA molecule changes its tertiary structure, which actually forces a change in its secondary structure. The switching domain (blue) pivots inward and displaces the antisense domain (red), peeling it free from the other side of the hairpin. The liberated antisense domain then binds to the messenger RNA and shuts it down. The cell no longer fluoresces, and the detection value is ONE—the ligand is present. In digital terms, ONE is OFF and ZERO is ON—the opposite of computers.

So then, of course, we put this antiswitch in cells to see if it would actually work. Graduate student Travis Bayer created an antiswitch with an aptamer that recognizes theophylline, which is found in tea and is chemically very similar to caffeine. He then inserted instructions for making the theophylline antiswitch into the DNA of yeast, specifically *Saccharomyces cerevisiae*, using standard molecular-biology techniques, and grew a batch of yeast cells, which took several hours. At right is a plot of the cells' behavior. The blue line is the switch response. You can see that as the theophylline concentration increased, the GFP response was not affected until a threshold concentration was reached. Then the switch suddenly shifted its conformation as it bound the theophylline, letting the antisense domain bind to the GFP messenger RNA, and GFP production ceased. So these molecules really work, and they exhibit a sharp, binary, response. I also want to point out their specificity—when Travis grew the yeast in the presence of caffeine (the orange line), there was no switch effect. So these sensor domains really can differentiate between very similar molecules.

RNA aptamers can recognize both small molecules and big proteins, which is a really powerful property. To demonstrate this, Travis has developed switches that respond to such things as the



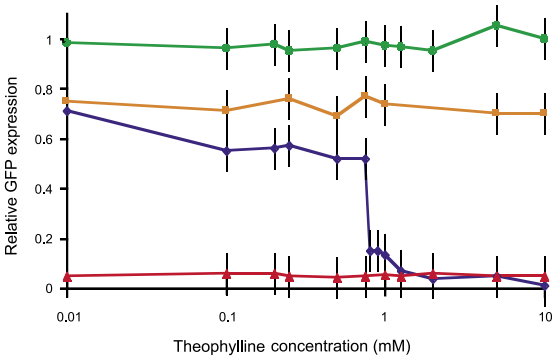
Theophylline



Caffeine

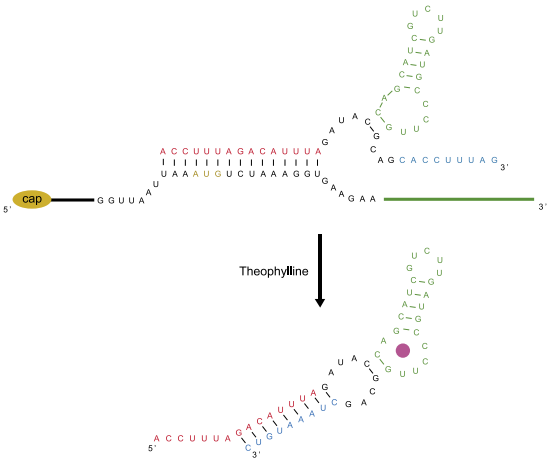
Theophylline and caffeine are very similar.

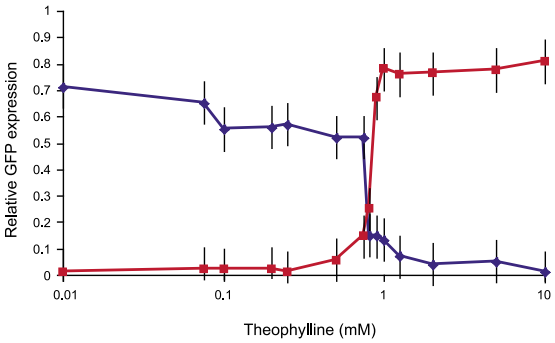
Adapted from Bayer and Smolke, *Nature Biotechnology*, Volume 23, Number 3, pp. 337-343, 2005.



Above: The blue line is the theophylline switch response. The green and red lines are control experiments. The green line represents a molecule with just the aptamer, so it never binds to the GFP messenger RNA to suppress production. The red line has the antisense domain but no aptamer, so it always suppresses GFP. And the orange line shows what happens when caffeine is added to the brew instead of theophylline, demonstrating that the response is specific to the latter. (The orange line is slightly lower than the green one because a few RNA molecules open their hairpins even with no ligand present, so GFP production is slightly inhibited.) The vertical black lines represent the error ranges in the measurements.

Right: This inverse switch turns GFP production ON when theophylline is present.





Above: In this plot, the blue line shows the response of the switch we saw before. The red line shows the behavior of the inverse design.

Both figures adapted from Bayer and Smolke, *Nature Biotechnology*, Volume 23, Number 3, pp. 337-343, 2005.

phosphorylated form of ERK2, which is a protein 250 amino acids in length that is involved in intracellular communication networks in human cells.

And we can adjust the threshold concentration by altering the relative binding energies of the antisense domain and the switching domain. Travis put some mutations in the antisense stem so that the nucleic acid sequences weren't a 100-percent match any more, and showed that this lowered the concentration at which the stem opened up. The two sides of the hairpin didn't stick together as tightly, so it didn't take as much effort—or, effectively, as many ligand molecules—to force them to let go of each other. On the other hand, when he elongated the stem (and kept all the matches perfect), it increased the stability of the closed state because it took additional energy to pry the longer sequences apart. This moved the switching response to a higher concentration of the ligand. So this is a really powerful platform, because not only can we sense a specific ligand by our choice of aptamer, but we can also program the concentration at which the switch senses that molecule.

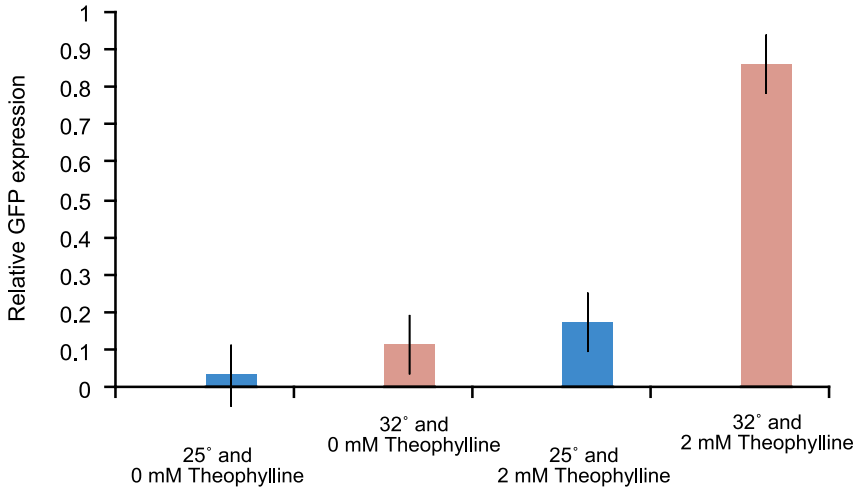
Jack went to the Red Door Café in Winnett Student Center and picked up decaf, regular, and espresso, and grew the gradient-filter yeast cells in them. Yeast cells do just fine in coffee as long as you add the standard culture medium, which is a broth of the sugars, amino acids, and other nutrients that they need to grow.

In real life, of course, you'd want to look at more than one ligand at a time. So Travis made a switch for tetracycline, an important antibiotic, which controlled the production of Yellow Fluorescent Protein, or YFP. When he put it and the theophylline GFP switch into the cell at the same time, the two switches retained their specificity. In the absence of the ligands, both were ON. In the presence of only one ligand, the respective messenger RNA got shut down as it should, and the other was unaffected. And in the presence of both ligands, both RNAs were shut down.

Travis also engineered the inverse design, where the antisense domain is bound to its messenger RNA in the absence of the ligand, and lets go when the ligand is present. He kept the same base-pairing energetics in the red and blue stems so that the switch would be triggered at the same concentration of theophylline, but in reverse.

We next asked whether we could make the molecular equivalent of electronic components within a cell, and we decided to start with a gradient filter. A concentration gradient is analog, varying smoothly from low to high, and a filter would translate it into, say, three discrete cellular states—LOW, MEDIUM, and HIGH—that you could then represent digitally. So in the summer of 2004, Jack Lee (BS '07) took our ON and OFF switches, altered their sensor domains to detect caffeine, and tuned their set points apart from each other. The ON switch, which controlled YFP, was tuned for a low concentration of caffeine (but higher than that found in decaf), and the OFF switch, which shuts down GFP, was tuned for a high concentration—higher than in regular coffee. Then Jack went to the Red Door Café in Winnett Student Center and picked up decaf, regular, and espresso, and grew the gradient-filter yeast cells in them. Yeast cells do just fine in coffee as long as you add the standard culture medium, which is a broth of the sugars, amino acids, and other nutrients that they need to grow. And behold, several hours later, GFP was found in the decaf. In the regular brew, he got GFP and YFP together, and the yeast in the espresso produced only YFP. So our caffeine sensor really works under field conditions, and we were very pleased by that.

Now we're looking at producing actual logic gates, which is the first step toward biocomputation. In practical terms, this means that the cell assays different biomarkers simultaneously—bio-

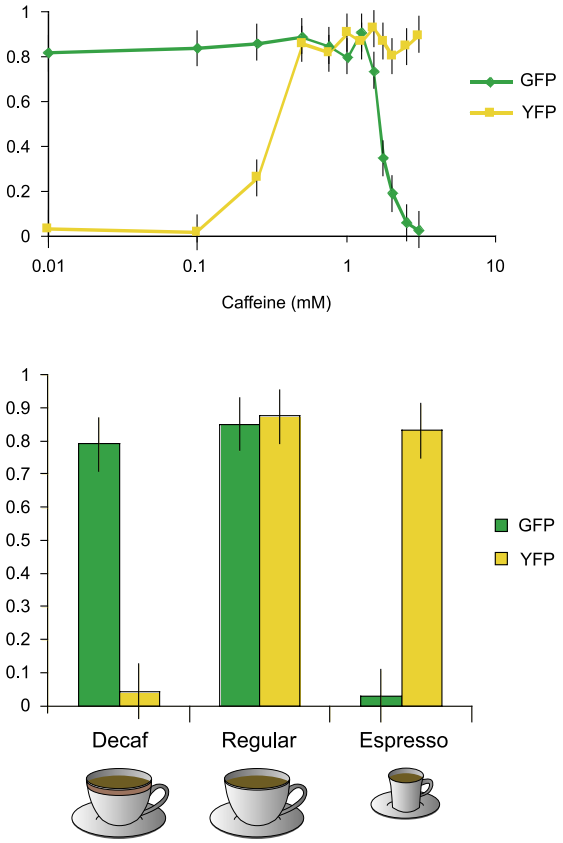


Right: The behavior of an AND gate sensitive to both high heat and theophylline.

markers being molecules that are indicative of certain conditions, such as a protein that is produced in cells only when they are dividing—and then, depending on the precise combination of biomarkers it finds, the cell performs a specific output. But the inputs to a biocomputer do not all necessarily have to be chemical in nature. For instance, Travis engineered a temperature sensor comprised of nucleic acids. I won't describe it in detail, but it's a cis-acting regulator inserted into the messenger RNA for GFP. The stem changes conformation with temperature, so that at low temperature, the protein is not produced, and as you increase the temperature, GFP production begins. So he put both the temperature-sensing GFP RNA and the inverse theophylline switch in the cell to create an AND gate that only fluoresces in the presence of high temperature and high concentrations of theophylline. Graduate student Maung Nyan Win (MS '05) is also working on the design of AND and OR gates that take two different biochemical inputs, but these are rather complex, and I won't go into them here. But in any case, these are our first steps toward performing logical functions within cells.

This logical capability will probably be really important in the design of intelligent molecular therapies. A properly chosen set of biomarkers would differentiate between normal cells and diseased or cancerous cells. That is, if and only if all the biomarkers are present, the cell performs some output, which might be metabolic reprogramming to make the diseased cells act like healthy cells, or targeted cell death, in which case we would *really* want to be sure that the cell is a diseased cell. Travis and grad student Chase Beisel are adapting our switches to function in mammalian cells. We are just getting started on this, but we are already seeing some very exciting results.

It's great to be able to identify all these biomarkers inside a cell, and maybe you wouldn't mind being injected with our switches as part of a cancer treatment, but if you're just going to the doctor's office for a checkup, you don't want to have all this stuff put in your body on the off chance that you might be coming down with something that it could detect. And you probably don't want to light up green, either. So the next logical step is to build some sort of chip-based diagnostic device that you could put a droplet of blood or urine or saliva into and get a rapid readout. Such a device would detect the presence of various critical proteins while also measuring the levels of important small molecules such as sugars, reliably pulling



Top: The set points for the caffeine-sensing GFP and YFP switches. Bottom: Their behavior in actual beverages.

Ultimately you'd be able to take, say, a blood sample, lyse the cells—split them open—and add their contents to a solution containing the switches and the templates . . . If testing several people revealed specific differences between normal, healthy subjects and people with a particular cancer, we could then use this as a diagnostic device for early detection.

all these diverse molecules out of a very complex mixture. So we're working toward a nanosensor based on our programmable switches and DNA amplification technology, the latter of which is the workhorse of biotech.

The polymerase chain reaction, or PCR, which won Kary Mullis one-half of the chemistry Nobel in 1993, allows you to start with one copy of a piece of DNA and turn it into millions of copies. The process basically takes the DNA-duplicating machinery out of the cell and puts it in a test tube. First you "denature" the DNA, pulling its two strands apart to reveal the bases. Then you add two short pieces of single-stranded DNA called primers that tell the polymerase enzyme where to start work. One primer binds to the strand of DNA that you want to copy, and the other one binds to the antisense strand of DNA that was pulled loose in the denaturing step. The PCR reaction uses both strands as templates, so that you wind up with two faithful copies of the original double helix. You denature those two and get four copies in the next cycle, then eight, and so on, increasing exponentially.

So Travis made an assortment of DNA templates some 100 to 200 base pairs long, and he made a unique switch for each of them, whose antisense domain acts as one of the primers. The other primer comes from the PCR kit. It's sort of like on a submarine, where it takes two officers, each with a different key, to launch a nuclear missile. When the ligand is present, the PCR reaction gets turned ON, and lots of copies of that particular DNA are cranked out. The switches recognize a substance called PDGF, for Platelet-Derived Growth Factor, which is one of many proteins that regulate cell growth and division, and he tuned the switches to respond to various concentrations of the factor. Then he put all the templates and all the switches and all the other primers into PCR reaction mixtures that contained varying amounts of PDGF, plus a complex stew of molecules that you get when you rupture cells—the sort of thing you'd find in a real medical specimen—which he added for background noise. And the switch-amplification combo not only successfully identified the PDGF, but it gave a digital readout of its concentration.

Other people in my lab are expanding on this work. We're moving toward a device that can detect multiple analytes, both proteins and small molecules, in a sample all at once. As a start, graduate students Arwen Brown and Maung Nyan Win are working on high-throughput technologies for generating and characterizing large numbers of switches and sensors. It would be nice, eventually, to be able to say, "I want a switch sensitive to growth factor X that stimulates the amplification of DNA template Y," and be able to make it more or less automatically. And the idea, of course, is that ultimately you'd be able to take, say, a blood sample, lyse the cells—split them open—and add their contents to a solution containing the switches and the templates and all that other PCR stuff. Then, once you've done the amplification reaction you'd pass the solution over a chip where the antisense DNA strands would be bound. The chip would be set up as a matrix, with each row being a different analyte, and the columns being various concentration thresholds. So we might be assaying for a whole set of growth factors, for example, and by reading the dots get easy, positive identification and quantification. And if testing several people revealed specific differences between normal, healthy subjects and people with a particular cancer, we could then use this as a diagnostic device for early detection.

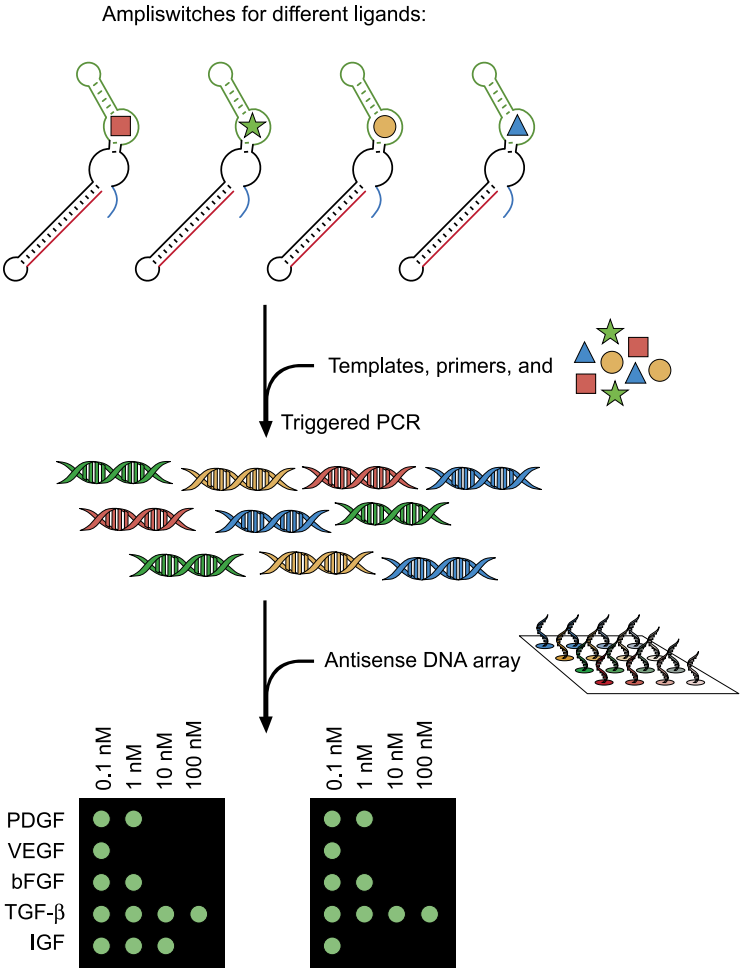
We've been using a similar scheme to pull out biomarkers for various diseases. We perform the reaction in a special way so that it outputs only the differences between, say, a regular cell and a diseased cell. We then identify those molecules with something like mass spectroscopy, which gives us biomarkers for different cellular states. And once we identify these biomarkers, we can use

PICTURE CREDITS: 28, 30, 32, 37 — Doug Cummings

Below: A schematic of a possible blood-test chip. This one is measuring the levels of various growth factors (Platelet-Derived Growth Factor, Vascular Endothelial Growth Factor, basic Fibroblast Growth Factor, Transforming Growth Factor-β, and Insulin-like Growth Factor). In general, growth factors direct the cell to change some aspect of its behavior in response to other cells or the environment. Altered levels relative to a healthy person could indicate, for example, the possibility of cancer.

them as targets for molecular engineering in their own right. But more importantly, we can use them to find other targets—by using each biomarker to find the next one, we can map the entire web of interactions that programs the cell to do whatever it's doing. This gets us back to the challenge I started with, of taking a global inventory of the cell's proteins and tracing the wiring diagram that connects them.

In summary, it's a very exciting time to be in this field. Nucleic acids present an inexpensive and robust platform for biomolecular science. These molecules exhibit impressive specificity and a staggering diversity of function. And because we understand so much about how their sequences of bases translate into structure and function, they are really a very powerful design paradigm. They're amenable to techniques that enable us to rapidly pull out functional molecules from randomized pools, and they're easily amplifiable, which is important for detection and diagnostic devices based on very small sample volumes. I've also been very fortunate, starting here at Caltech only a year and a half ago, to get great graduate students and undergraduate researchers. They come in with a lot of excitement, a lot of energy, and a lot of creativity, and that's really helped us make so much progress in this area so quickly. □



Assistant Professor of Chemical Engineering Christina Smolke earned her BS in chemical engineering, with an emphasis in biology, from the University of Southern California in 1997, and received her PhD in chemical engineering at Berkeley in 2001. After a postdoctoral fellowship in cell biology there, working on RNA decay pathways in S. cerevisiae*, she came to Caltech in 2003, returning to her southern California roots.*