C. R. Chimie 14 (2011) 424–428

Contents lists available at ScienceDirect

Comptes Rendus Chimie

www.sciencedirect.com

Full paper/Mémoire

Building a new biology☆

Drew Endy

Stanford Bioengineering, 473 Via Ortega, Stanford, CA 94305, USA

ARTICLE INFO

Article history:
Received 1 November 2009
Accepted after revision 15 November 2010

I hope to share some tools and perspective from the engineering community on the subject of synthetic biology. My views will reflect a lot of constructive input from many colleagues, who have been working now for about eight years, coming to biology as technology, not so much as a science, to think about what we might do next. So, to give you a sense of context, imagine that you like to build stuff, and you look at the living world as one of the most impressive technologies that's ever existed. It's already taken over the planet, it replicates, it's very tiny and it's very big, it makes all sorts of things. What if we could get really good at engineering biology, what would that look like?

So that's what this boy is asking. He decides to try. He asks some questions, such as how does the natural living world work? You could go inside a cell and look at the master programme, if you believed these metaphors, and as we heard this morning these beliefs get to be tested quite directly by building. Nobody really knows how this works exactly, but maybe we could try and make stuff and see what happens. This young man decides he wants to make balloons, in his laboratory, filled with hydrogen. All he needs to do is get the components that you would put inside the cell that would operate a genetic programme. Of course, you could get the components off the shelf, right? But we're going to programme the DNA. And so you look up the different components from the hydrogen production module to the balloon production module and so on. And then you get the DNA that encodes those components. This is a comic book, it's not a real laboratory, but you could imagine that somebody might want this to come true. The DNA is not to scale, it's a little bit smaller in reality. You take the DNA programme, you put it into the cell, the cell starts running the programme... it looks pretty interesting. It looks a little too interesting! The boy forgot to tell his programme to stop. Maybe he should go learn more before he tries to programme DNA.

Well it turns out that these fantasies in a comic book are representative of desires, and the engineering community would like to make some of these fantasies come true. So for example, if you go to this website, you can find a collection of standard biological parts. Our inspiration might be from mechanics, or electronical engineering, we don't really know how to make biological parts standards. But we want to try. So today, there are over 3000 BioBrick parts, freely available via this collection. A quarter of a million of them were shipped to students around the world last year. These are the sorts of students who make the BioBrick parts. I'll give you an example of one. Students from Melbourne, Australia, having read the comic book perhaps, made the BioBrick part 750016. It's an adaptation from a natural piece of DNA, from a microscopic organism that lives in the soil, and it's about 6000 base pairs of DNA long, and the DNA encodes 11 different proteins that are expressed inside the cell, they self-assemble into a little protein balloon. So you could take this DNA, put it into a cell, and this cell would make a balloon. Not on the outside but inside. Students in Melbourne Australia, which is the

☆ This article is a transcription of an oral presentation made by the author at the *Collège de France*, during an International Colloquium organised by the Chair of Philosophy of Science, Professor Anne Fagot-Largeault: *De la chimie de synthèse à la biologie de synthèse*.

E-mail address: endy@stanford.edu.

1631-0748/$ – see front matter
doi:10.1016/j.crci.2010.11.013

antipode practically of Cambridge Massachusetts, were able to engineer a piece of DNA making 12 changes to it so they could give it back to the world, a protein balloon DNA, such that anybody else could quickly try it out for themselves. So you could make cells that float, or sink, or are neutral in their buoyancy, which if you're interested in fermentation might be a useful property to programme.

There are other parts you can find; as we heard about this morning there's a banana odour generator that catalyses the synthesis of isoamyl acetate. If you put this DNA into *E. coli*, the *E. coli* might smell like bananas. This part was developed by these students: first and second year undergraduates, teenagers, who upon coming into a laboratory for the first time, not having taken genetics, or a microbiology lab, were immediately dissatisfied with the bouquet of *E. coli*, and so they implemented a project which they called Eau d'Coli. And practically speaking, what they really wanted to do what encode this genetic programme: if the cells are growing call wintergreens smell like mint, else call like bananas smell like bananas. Their contribution would be thus not only to improve the working conditions of the lab, but also to have an odour reporter of the growth state of the culture. In the period of a summer, they were able to sketch out the chemical conversions from chorismate to methyl salicylate and leucine to isoamyl acetate. This is about 24 different BioBrick parts to implement this system. When they started the summer about 12 of the parts had already existed, they had to make the other 12. In 3 months they were not able to get this to work, they were only able to get the right half of the picture to work, from the intermediates to the final products. That was enough to do a demo, and a smell test. And it turns out that most people can identify the laboratory strain of *E. coli* as stinky in a blind smell test. The mint's is not so easy to detect and the banana's is very powerful. One of the students was so excited, he came back the next summer and did a full biosynthesis of the methyl salicylate.

Why is this important? It's important for things that have to do with our economy and our security, but there's an element of this which is, we have opposable thumbs and we like to make things, and that's part of who we are often times as human beings, so I just want to note that. But then there's also the world we live in with all the challenges and opportunities and if you bin the world into sectors that need help, or where you look for opportunities there's lots. So you could look in food, and imagine the time it took to go from this type of corn to what we've got today, and then you could note that most of the things we might eat we can't, because we can't produce enough, and if we wanted or needed to have agile domestication of crops, how could we do that more quickly if we had to? Or you could open up another bin, say chemicals, and this comes from the US department of energy: they would like to produce over a hundred different chemicals from renewable feedstocks as opposed to petroleum. We actually can't go do this, with any economic efficiency.

For example, here's a famous project, it's an important project. A colleague and friend Jay Keasling at Berkeley, recognised that a lot of folks who have malaria can't afford the drug. It's wonderful to have a drug that might work, but if your healthcare budget is quite tiny, you probably need to get to a price point that's about a dime a dose. The drug, artemisinin in this case, comes from the bark of a wormwood tree, which is extracted therefrom, what if you could reprogramme the metabolism of *Escherichia coli* or yeast to make this at a cheaper price point? Turns out they've been successful in getting that out of the labs and into the industrial scale-up. It cost them 25 million dollars to do the metabolic engineering work. Twenty-five million dollars paid for about 150 post-doc years, experts doing the genetic engineering and testing to do biosynthesis of one drug. In part, due to reasons of cost and supply, you have resistance to artemisinin, this project is a herculean project, is it always going to cost us 25 million dollars to do this? If the answer to the question yes, because it's hard, or yes because we can't make biology easier to engineer, okay, but if there's an opportunity to get much better, to make this a 25,000 dollar project or 2500 dollar project, that's an interesting opportunity to explore. Could we make this part of the natural world, the living part of the natural world, much easier to engineer? Would things like this be possible, or is it simply a fantasy: a gigantic gourde that grows into your own house, right? Everybody talks about medical applications, so I just want to show you what architects imagine you could do with biotechnology.

There's a great slide from earlier this morning about stealing, I steal a lot, and I'm a naive felon in many respects, and so I want to show you some points of departure for theft not to be directly linking what I'm showing as something that's relevant to biotechnology but it's how an engineer looks at the process of the work that we set out to do. For example, a long time ago our ancestors decided to build with these natural objects. It's hard to build with rocks that look like this because they're all different. And so people decided to work to regularise rocks, as much as you might be able to. Thus other sorts of artefacts become possible. The substrate is being simplified but the artefact so produced from the simple objects becomes more complicated. Later on, we become dissatisfied with rock, so we grind up rock, and we make new synthetic rock, in the form of concrete, and that allows all sorts of new things to become possible. The Stone Age as it exists today... we can compile cities like Hong Kong, typically not this clean air, but the cities in decades as opposed to generations, because of the tools, the change in the process of engineering and also the modifications to the substrates from..... engineering.

Another point of departure for theft, if you will, is the transition over the last 70 years or thereabouts in computing, so if this is what a computer looks like in 1952, this is less than a human generation later: 25 years later... and this is 30 years later. How come this is true? The point I want to bring up is, how did we as a civilisation take natural objects and make them easier to engineer, or how did we have a set of things that we wanted to produce, and become qualitatively better at doing so? When I go and buy my phone or my computer, or a biotechnology application I'm almost all the time driven by the application of the phone: what can I do with it, what disease will I cure? Biotechnologies I would argue over the last 35 years have been overwhelmingly driven by applications. There's been very little support for things

below that. What's below applications? Tools, things that impact the process of the work, things that change the substrates from which your applications are built upon. So if you take computing as an example and look at the last 70 years of computing, you have whole categories of tools that I don't typically think about when I buy a computer. I don't think about the ongoing and past of manufacturing of silicon waveforms. I don't think about the work that went into inventing computer programming languages. I don't think about computerated design industries, electronic design automations. These two right here, for example, CAD and EDA, these are representative today of billion dollar markets just in design tools, for figuring out how to manage information going onto silicon waveforms. This is all hidden to the consumer of a technology.

Well, how about biotech? How about genetic engineering? Again I'm trying to reflect on the content and the tools that the engineering community brings to this new world of synthetic biology, such that it might exist. Here's a paper, from 1973, when the first plasmid was constructed, using recombinated DNA. This led to the production of insulin and bacteria for treating diabetes. And you can see the methods sections, it's quite detailed, and it's expert driven work, where folks are looking at the different restrictions under nucleus sites and... figuring out how to build and test this. Twelve years later, here's another paper, leading to the cloning of erythropoietin, around the birth if you will or err, one – early stage of –, as a company. Here's the method section from the paper, and the method section is... in detail different, very different, but in process, if you abstract out the work, the workflow is not different. In a period of 12 years, people still ended up manually figuring out how to manipulate and assemble genetic material and test it out, just to produce a particular product. Well, okay, that's 1985. How about now? Okay, this is as good as I can get: 2006. This is the paper, leading to the production of artemisinin, in the Keasling lab. The drug for treating malaria..... and I'm sorry the text is so fine, the big difference it turns out of the last 35 years is the methods have moved from the paper to the online supplement, and so you have to go look them up separately. But if you read this, the details are different but the process has not changed. Over my entire lifetime, some of the most basic steps of the workflow in genetic engineering is qualitatively unchanged. Compare and contrast this to microprocessor design, where in the 1970s people are drafting manually by hand the layout schematics for chips, versus what we have today. Now it could be that you can't do any better than this. People have been working very very hard to replace this with a superior process. Or, it might be that people have been so driven by applications in biotechnology and many other factors, that there's a lingering, now looming opportunity to get in down here at the tools level, and change things around. Or at least experiment. We've split this up into four problems, and I'll talk about some of them in combination.

So this will be synthetic biology as a tool of revolution not to be exclusive, but as the engineer's perspective. So we're familiar with these sorts of tools: recombinant DNA for cutting and pasting, PCR for amplification it's very interesting, we could have had PCR and recombinant DNA at the same time, but they were separated by a decade, and then sequencing for reading things out. Are those the only tools? Probably not. Maybe you could get really good at automating the construction of DNA that would allow you to separate design and fab; architects, engineers, constructors. Maybe we could abstract genetic material so that it's not always so complex, and I'll give you an example of that, and maybe we could standardise what we're working with so that it could be reusable. What does this mean? So here are four bottles of chemicals, phosphoramidites derives from sugarcane, they cost about 250 dollars a pop. There's enough of the four bases of DNA, A, T, C and G in these bottles to build 60 copies of every human genome on the planet. We don't have DNA synthesisers that are capable of doing that compilation at more efficiency. But it's interesting just to imagine sitting with that sort of potential there. You plug these chemicals into a machine called a DNA synthesiser, it's hooked up to a computer network, information comes in, and a strand of DNA as physical material is compiled. We have matter compilers, the chemist have worked out matter compilers for genetic material. As the naive engineer, this is the most amazing technology I've encountered.

What's also interesting is the pace at which these tools have been improving in part, some of them. So this is... across the last 40 years, an algorithmic scale, features per chip on microprocessors, ... for computing, where you go from the 1000 and 10,000 to 100,000 and so on. Computers get better by a factor of two every year and a half. This is capacity to sequence DNA in base pairs read per unit time. So Sanger is around 1977 and the genome projects get going, and 1995 you sequence bacteria, the scientific community, only six years later we have a draft of a human genome, That's because there was geometric improvement in sequencing capacity during the Clinton administration, let's say. This is synthesis of DNA in masse as oligonuclear tides, short fragments up to a couple hundred units long, and that's improving geometrically too. And last year was the year that we had the first compilation of a bacterial genome from scratch. What will the world look like in six years? How big a piece of DNA would we be able to construct? Well the take-away is, at some levels, short fragments, not big genome fragment scale yet. Our ability to build DNA from scratch is improving quite quickly.

What does this lead to? Now I'm going to take a little detour here and then come back to the tools and context. All natural biological systems that we tend to celebrate exist via this process of direct descent and replication with error that allows for selection and evolution. Everybody talks about how cool evolution is, but an engineer can naively say evolution is not cool, it's a tyrant: it gives us mutation without representation. Who ordered all those changes? Well, what if we could do this? What if we could take this process of direct descent, decompile it up to information, and then change the information as we wish, and recompile back down? From matter to information to matter, this allows a new path forward in time for the propagation of living systems. I'm happy to call them living; it also means we have to figure out what this might mean and if we're comfortable taking responsibility for

such work. It has a lot of impacts that are practical. Let's say you're concerned about property rights, what if someone just gets the information and compiles the material as opposed to... waiting for you to sign a piece of paper? What if you're concerned about biosecurity and somebody can get access to pathogenic sequences online and compile their own pathogens? What if you're concerned about biodiversity? The transmission of organisms across borders, either via biopiracy concerns, or environmental preservation concerns? Probably some other things too. This decoupling of information and material is a very interesting access, when it propagates. You can think about digital music, digital entertainment as that gets distributed over networks.

It's been touched on in passing and I'll just have a shout out for it here, there's a tremendous opportunity here to change the science of biology, as has been hinted on. So I'll show one example just for a couple of seconds, where five years ago now, we took the genome of a natural organism and we made 600 engineer changes to it, just to test its architecture, to see what we could throw out, and then went to see whether the organism after it had 600 changes to it was it still viable? It turns out that, yes it was still viable, but it wasn't very happy and if you let it evolve, it would throw out the changes we had made. This idea of learning by making is incredibly powerful in biology and deserves more attention I think in many communities.

Back to the engineering. This is to help me remember to basically highlight the obvious problem. If you could build whatever piece of DNA you want, and your capacity for building DNA is improving geometrically, what do you want to say? How do you manage the knowledge and information going into the synthesisers? How do you scale that? In silicon manufacturing, it's called the technology gap, so mind the gap! I really appreciate how, we heard this morning the simplicity if you will, of some of the chemistry of DNA. The information that's often coded on DNA to an engineer is complicated. It's a lot, and we don't have to deal with a lot of complexity to feel like we're overwhelmed with complication, so TAATACGACTCACTATAGGGAGA is a consensus promoter for the ..... preliminaries: it's a little start sequence. If I had to do all my genetic programming, at this level, I would be overwhelmed with the sequences. It would be like teaching students how to programme computers in machine language. Maybe it would be like that, right? We have different approaches to programming computers now that have been invented over decades.

Could we invent approaches to programming in DNA? We don't know how to do this, but we're trying, and what we're starting with is this naive idea that we could implement an abstraction hierarchy for managing biological complexity. Where somebody could be assistant engineer, let's say they want to reprogramme the odour of *E. coli*, and they wouldn't need to know that DNA is made up of 4, or 6, or 8 bases, let alone anything about how to synthesise it. Because they'd be able to call down to lower levels to get access to the functions, and it would work. Could we make this a reliable technology process? Don't know. It is a research question, and in the engineering community it's a pressing research question. For example, could you take the biology of a squid, that lives in the Pacific Ocean and has bacteria that produce light at dusk, when the squid likes to hunt, that communicate to one another using a modified sugar, and via this cell–cell communication system extract from the natural genetic components a sender and receiver; a transmitter and a receiver if you will, that other genetic engineers could redeploy without having to know the scientist, who did the wonderful work, to figure out how this natural system might work? Right, without having to know all the details of the squid. Well you could begin to prototype a system, you could try and standardise the components, you could naively try and blackbox them, abstract them, you could try to describe the behaviour of this device, such as it is on a data sheet, so that you could easily present the information to somebody else or perhaps even just to a computer, thus when somebody comes in and wants to build an engineer biological system, they have access to the object.

We published a first example, a very naive example of trying to do this last summer, this is stolen from the transistor-transistor logic databook that Texas Instruments produced in format this is a datasheet, not for a piece of electronics but for a piece of genetics. A cell–cell communication device. Here are the three tools over the last generation that allowed for manipulation and reading out of sequences. We could get better at construction, we could get better at abstraction, we could get better at standardisation. And by getting better, I don't mean solve the problems like what we see today in nuts and bolts, although that would be great, or electronics, that would certainly be great. Just make an incremental contribution to the process of engineering. Turns out we've been able to do that, and since we're competing practically with nothing, any incremental utility has a relative infinite return.

Students, given the access to these tools, start to design and build genetic systems. This is now what's called over the years iGEM, the genetic engineering, where teams of students each summer design and build a genetic system of their own specification. The Eau d'Coli project, bacteria that detect arsenic in well-water in Bangladesh, and so on. They recognise where they come from, they go out and have tours to recruit new students, in this case from China, and they grow geometrically, because biology is an exciting technology platform. This is last year. So this is now probably up to about 1200 or 1300 students each summer. To give you an example real quickly of what these students are capable of, here's a project from UCSF: UC San Francisco. They don't have undergraduates at the University of California San Francisco, so they took high school students from a public school in San Francisco. These students noticed on the online videos that many of the other projects that students worked on failed. And the explanations for the failures were not really satisfying. And so their contribution was to... sketch out an approach that would allow other teams to separate the modes of failure. And if they could help other students figure out, well did it not work because when I put DNA into the cell, it had weird interactions, or did it not work because my design was very poor? Maybe if you could separate those two modes of failure, that would help other engineers debug. Their inspiration, to start off solving this problem, was a very

sophisticated idea from computer programming. There's a programming language called Java, from Sun Microsystems, now acquired most recently, Java can run on a Mac, an Apple, or a Linux machine, or a Windows machine. And the reason this is true is that Java software does not run directly on the computer, it runs on something called a virtual machine that is installed on the computer and then that provides a common sandbox, or operating environment, for the genetic programme. What would a common virtual machine, a common environment, look like for genetic engineering? So that you could take the same piece of DNA, and operate it inside a yeast cell, inside a newt, or a plant, or a person? With less complication. Their idea was to build a new compartment, a new organelle, a synthetic organelle, which they called a synthecell, and if you could boot up from scratch inside any organism of your choosing a new compartment, then maybe you could operate your genetic programmes inside that compartment, as opposed to throwing the DNA willy-nilly, into some uncharacterised or relatively poorly characterised environment. And the idea here, note this would be a tremendous amount of toolwork, would be to install and be able to boot up this organelle across many different types of organisms.

High school students: they were able to think about doing this by taking these computer science ideas and mapping them into organelle trafficking and biocellular chemistry specifically, phospholipids that are responsible for rouding vessicals that get internalised during receptor and disitosis and signaling response. They brought into yeast a modified phospholipid from human and were able to specifically target that, and then in the period of a summer, were able to show that before and after treatment of a stimulus, you could get production of what looks to be, vessicals that are targeted with a fluorescent protein specifically binding the orthogonal phospholipid.

This leads to a couple of questions. One question has to do with safety. How do we turn over biotechnology to the next generation of biotechnologists? Do we have to worry about that, or do we get that for free? These little things on the right are my opinions, but they're all really points of discussion. In the last century, we had very direct and strategic, uhm some might say, not strategic, militarisation of biotechnology. Could we afford that again, and going forward. Are folks going to use this technology outside of institutional frameworks? Will there be hack-hackers of biology? Will these folks be good or bad? Who is the community who's interested in biotechnology? Remember the transition from 1952, the computer to the computer of 1977? That's in part a tools evolution, but it's also driven by people being very excited about computing, and being disappointed about limited access to computing. Biotechnology is more exciting than computing, in my opinion. These parts, should they be free, or patented or what? So we had 1500 new parts come into the collection last year. At 25,000 dollars, a pop to patent worldwide, per part, that's 37,5 million dollars in attorney fees. The whole cost of the competition is 3 to 4 million dollars. So already our legal bill in ten-fold. Our actual operating budget. And if you gave us 40 million dollars, we would go try and make better parts. I suspect we're heading towards a new type of law, recognising that genetic material and information are becoming interconvertible, via sequencing and synthesis, and we have to somehow update things.

Real surprise, I studied civil engineering as an undergraduate, reinforced concrete and steel, and we had to take professional engineering courses, and I had to, if I wanted to go practice, be licensed by the state. In the United States, I don't know how it is in France, but in the United States there's the American Society of Civil Engineers. There is not the American Society of Genetic engineers. There's no code of profession. Should there be one? Should we sign our work in the DNA? Perhaps. How much..... can we make with biology? I'll end with this: the excitement in the US and elsewhere right now for example around biofuels, is quite something, but I can't figure out what it's... really going to end up with. I hear estimates that we might be able to get 5 TW worth of energy via biology. Others suggest it might be 90 TW. The reason this spread between 5 and 90 is significant, is our civilisation runs on about 15 to 18 TW. So either biology is going to be scarce, biology-based power, biology-based manufacturing is going to be limited, or it's going to be in surplus, and I don't know how you would make these predictions accurately. But the reason it's significant, is if you think about planning, if you think about land-use politics, whether we have an excess of power and manufacturing capacity, or gross limits, it makes a big difference.