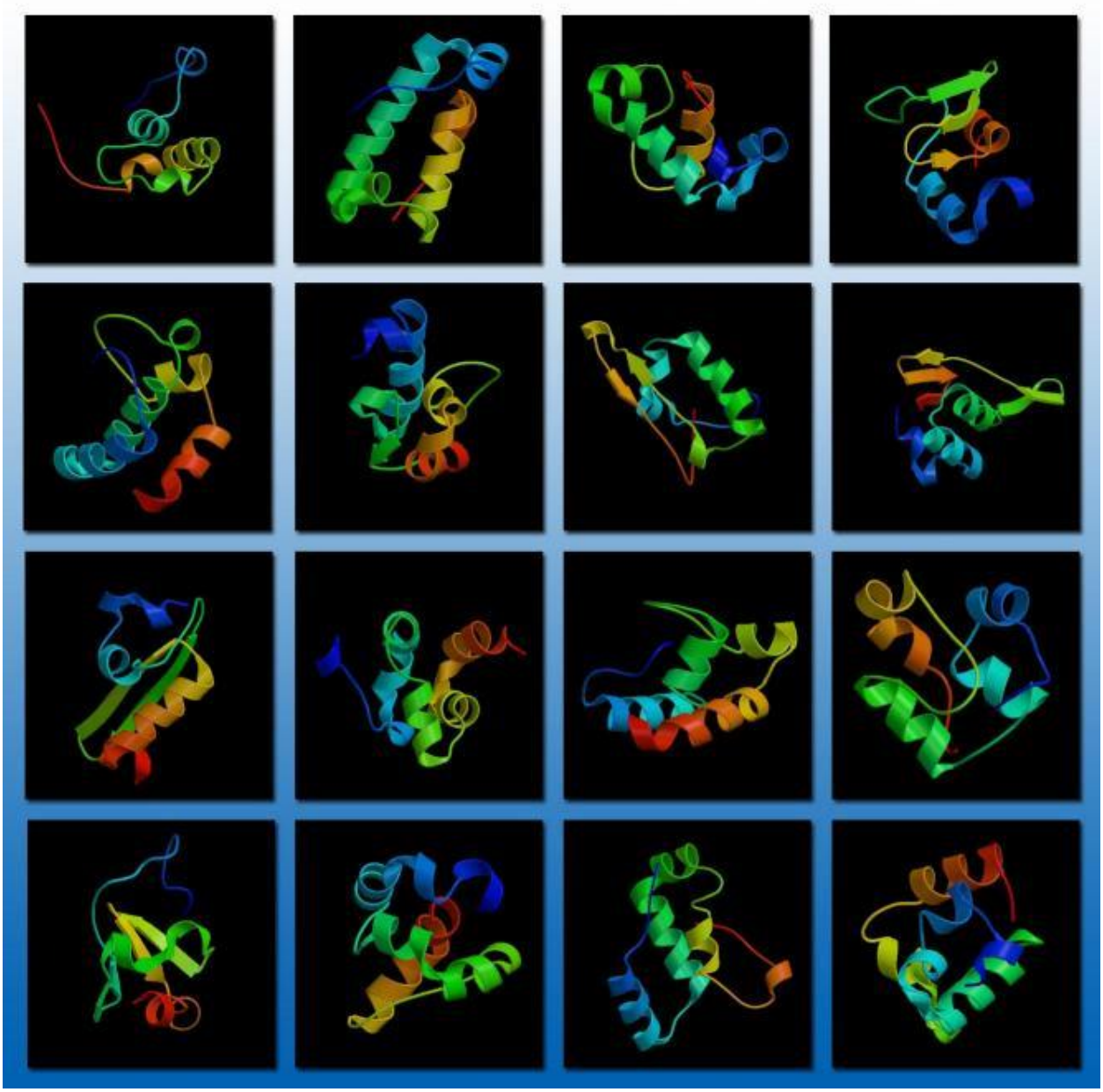
# DESIGN FOR LIFE

## Meredith Walsh

It is argued that design technologies developed for synthetic biology capture a surplus essential to life. [1] However, I suggest that life's emergent capacity is not being controlled by or inappropriate to design. Rather to explore the potential of life for novelty what is required is a new vocabulary of design. A vocabulary that can be drawn from architect and product designer Greg Lynn.



*A selection of proteins synthesized by Pier Luigi Luisi Synthetic Biology Laboratory, © Pier Luigi Luisi Synthetic Biology Laboratory.*

## Introduction

To design novel biological systems synthetic biology has created new software technologies such as *GeneDesigner, Rosetta*, *Foldit*, and molecular graphics programs such as *Chimera* that draw on areas of design including architecture, fashion, media production and electrical engineering to manipulate DNA and produce novel protein structures – the work horses of cellular function. Grouped under the heading of bioengineering, these new design technologies it is argued capture a surplus essential to life within a web of aesthetic and cultural practices aimed at mass producing novel life integral to the digital economy and biocapital. [2] In contrast to design, these arguments continue, novel life emerges chaotically or randomly. Therefore, the design of novel DNA and protein structures either attempts to control [3] or is inappropriate to this capacity. [4] However, I suggest that life's emergent capacity is not being controlled by or inappropriate to design, but the current language of design is limited in its ability to articulate the potential for life. Rather than opposing design and emergence synthetic biology requires is a new vocabulary of design. A vocabulary that can be drawn from architect and product designer Greg Lynn's use of geneticist William Bateson's idea of symmetry breaking.

To discuss developing a new vocabulary of design, because of the fundamental role proteins play in cell function and therefore synthetic biology, in this paper I will focus on the synthesis of novel proteins. As my motivation for exploring a new vocabulary of design is the opposition of design and the chaotic emergence of novel biological systems on the basis that it is an inherent capacity of life, I will look at the Baker lab's software Rosetta –as it is arguably the most common way to design new proteins –and the randomisation of DNA by the Luisi lab in Rome, where I was recently an ANAT (Australian Art and Technology) resident. To articulate the workings of these two techniques, I will first sketch out the very basics of protein coding and function.

## Design v. Randomisation

A relatively small segment of the genetic code, genes code for the synthesis of proteins. Through the production of an intermediary code via a process known as transcription, the four bases of DNA, GATC are translated into a sequence of amino acids. Comprised of 20 different discrete amino acid residues as they are called, these sequences are ordered in various lengths as specified by the gene code. After an amino acid sequence is formed, for a protein to work, generally speaking, it must fold into a structure specific to its function. Specified by the order of its amino acids, the structure of proteins plays a functional role in producing different types of cells: particular protein structures produce hair cells, liver cells, skin cells, etc.

Synthetic biologist David Baker describes protein design as the hunt for the amino acid sequence that will fold in such a way to create a protein structure that carries out a desired function. [5] The number of different ways a protein could fold, however, is astronomical. Known as the sequence space, for an amino acid sequence of only fifty residues, the possible combinations of sequences is 1065 (20x20 x20 etc. for 50 residues). Given that a particular sequence codes for the protein's structure, each one of these possible sequences can potentially fold into a different structure. This is a staggeringly large number: For structures that are nanometres in length, the combined weight of possible amino acid sequences for just fifty residues would be equivalent to that of the earth. What's more as most proteins are longer than fifty residues, the number of possible combinations of amino acids is even larger.

As the possible number of amino acid sequences is so staggeringly huge, some means had to be found to reduce the variables needed to be calculated, which would yield a novel protein. To reduce the number of calculations, Baker's protein prediction and design software Rosetta uses banks of known amino acids structures. Instead of using individual residues, Rosetta takes short fragments of about 9 residues from larger structures and variously assembles them into new structures.

When used in its predictive mode to provide the most likely structure into which an amino acid sequence will fold, Rosetta takes a known sequence and assembles nine hundred possible structures, from which it ranks the top ten. For each of these structures, Rosetta then gives the x, y, z, co-ordinates of every atom. Taking these co-ordinates the proteins are visualised as folded three dimensional static structures, using molecular graphics software.

When it is used in design mode to create a novel protein rather than predict the structure of an amino acid sequence, Rosetta uses a reiterative approach. To begin its reiterative approach, the designer inputs an amino acid sequence for which there is no known structure into Rosetta. Deploying its predictive feature, it then ranks the top ten most likely structures for this sequence. After ranking the most likely structure, this amino acid sequence is fed back into Rosetta and the process is repeated. Once again, sourcing fragments of known protein structures Rosetta assembles and ranks the top ten structures. Using this technique, the Baker lab created the novel protein Top 7.

However, despite successfully designing a novel protein using Rosetta, Baker's approach is criticised by both synthetic biologists and its commentators. Criticising the bioengineering style design claimed to characterise synthetic biology, both camps argue that the systematic and various assembly of extant genetic components found in databanks runs counter to the inherently emergent function of life. Commentator on synthetic biology Adrian Mackenzie for example suggests that the design of novel biological systems, using standard parts and assembled components in manner drawn from electrical engineering and other product design typified by drop down menus and drag and drop objects controls life's essential chaotic function. [6] Assembling and rearranging parts in an orderly fashion for an intended functional outcome, according to Mackenzie, by definition controls the chaotic capacity by which novel living systems emerge. From within the laboratory, the Pier Luigi Luisi laboratory similarly argues that bioengineering is unsuited to the emergence of novel life. Writing on the epistemology of life which informs his experimental research, Luisi argues that life did not occur nor changes by design. [7] As with Mackenzie, for Luisi bioengineering design is characterised by a modular means of modifying or producing novel biological systems developed from electrical engineering, which relies on the assembly of extant genetic components or biobricks stored in databanks. And similar to Mackenzie, Luisi argues that variously assembling banked genetic components does not take into account the emergence of life.

Drawing on systems biology and autopoeisis, novel biological functions he argues do not causally derive from the reassembly of single components; they randomly emerge and cannot be causally identified as the sum of their preceding parts. [8] On this basis that novel biological systems *emerge*, he further suggests the attempt to *design* novel living systems is unlikely to have little more than isolated success. Though, Baker has designed a novel protein, and Venter developed Synthia, using a cut and paste logic, the design of novel life he says assumes that when taken out of their systematic context components such as single enzymes and metabolic pathways will retain their specific functionality. [9, 10] Taking a top-down approach, as the lab refers to design, is therefore unlikely to be broadly effective. While the lab acknowledges that there have been some remarkable one-offs, using a bioengineering approach they argue, design has not provided synthetic biology with a roadmap. [11]

Though I empathise with the lab's argument for emergence, by defining the capacity of life as emergent, the Luisi lab's position resonates with the objections to design of Mackenzie and others. While the lab does not overtly argue that the chaotic capacity of life for novelty is being controlled by design, they do argue that design cannot emerge novel proteins because it is intentionally directed at causal outcomes, unaware of life's emergent capacity for novelty. [12] Despite their differences, both Mackenzie and the Luisi lab effectively insist that emergence is an inherent capacity of life.

In keeping with their assertion that design is at odds with life's emergence capacity, the lab has developed a technique to randomise DNA as a way to create novel proteins. Rather than relying on the assemblage of extant components or biobricks to design a protein with a preconceived function, they have developed a means to randomise the order of the four DNA bases GATC, in an attempt to code for amino acid sequences that may fold into novel protein structures. As the possible order of amino acids for even a fifty residue sequence is astronomical, randomising the order of the four DNA bases GATC even a restricted amount of times generates a significantly large number of amino acid sequences. Since not all amino acid sequences will necessarily fold into functional protein structures, the large number of sequences generated constitute a pool (approximately 109) in which to fish for those that may fold, and do so in novel ways. (While the lab has generated a significant number of folded structures, the lab is still in the process of exploring the function of the proteins they have synthesised in comparison to already known proteins. As such, it is not entirely clear whether they have created any which are novel. [13] Given the large number of sequences, to fish for novel amino acid sequences which may fold into functional protein structures requires a means by which to select possible candidates which will structure when inserted and expressed in a living cell. It is not practically possible to test them all. Since it is not possible to test all sequences, the lab must use some form of protein prediction software. Enter Rosetta. Despite their attempt not to use extant components of any sort, the Luisi lab's random technique relies on Rosetta's use of statistically averaged structures assembled from existing protein fragments to predict the likely structure of the amino acid sequences they have generated. While they assert that novel systems emerge randomly, their technique remains reliant on banks of extant components. And their approach snared in the tenets of protein design.

Rather than design either controlling or being an ineffective means to create novel proteins based on life's inherent emergent capacity for novelty, I suggest that the current method of protein design is limited in its ability to experiment with creating novel proteins. Importantly, what I mean by 'limited' differs from any claim about the affectivity of design or otherwise in regard to the inherent capacity of life. In contrast to claiming that design is ineffective on the basis that life is inherently emergent, a limit as I propose it is coupled to the idea of potential. Potential does not inhere in life. Rather, potential is an opportunity for novelty that may occur through the temporal interaction of the components of a system which may themselves change. Though I will detail this idea below, simply put a potential for change is not an inherent capacity of life because it does not pre-exist in life *prior* to the temporal interaction of its components; it occurs *during* the interaction of its components, which may themselves change. Taking into account that the potential for novel life occurs temporally, bioengineering design cannot be argued to either control life or lack efficiency because it does not acknowledge its inherently emergent capacity. Instead a bioengineering approach can be seen to be limited in its capacity to explore life's potential because it does not factor time into design. Static structures assembled out of given components which are statistically ranked have no dynamic temporal dimension, and therefore no opportunity to interactively change.

Rather than eschewing design on the basis of life's inherent capacity for novelty, I suggest a new vocabulary of design is required to experiment with the potential for life to produce novel proteins. A vocabulary that can be fruitfully explored by drawing on the language of architectural and product designer Greg Lynn.

## Design for Life

Rather than Darwin geneticist William Bateson is Lynn's hero. Coining the term genetics in 1905, Bateson inspired Lynn's design vocabulary for which he is renowned. [14] Drawing on Bateson, Lynn broke with the static design of organic structures to reconceive architectural design through animated forms. Arguing that static approaches to architecture were limited by structure being the determining force of design, Lynn turned to Bateson's interest in the exception rather than the ideal form. Conceptualised in terms of symmetry breaking, exceptions occur whenever there is change in the symmetry of an organic form to asymmetry. Contrary to the idea that a symmetric form is a source of information about its structure, according to Bateson when organic forms become asymmetrical information is generated and when they return to symmetry it is lost. To apply Bateson's idea of symmetry breaking to architectural and product design, using calculus Lynn developed his vocabulary of animated structural change.

According to Lynn, the calculation of time is possible through the mathematical language of curvature. And calculus is the language of curves. In contrast to static structures and ideal forms, Lynn's interest in calculus is "... the creative structural role of time and force …" [15] A creative role that he marries to Bateson's idea of symmetry breaking to generate a design language of continuous form and exceptions, information and its loss. Using calculus to generate exceptions, Lynn reconceived the architectural idea of a static form that exists in empty space to a manifestation of dynamics forces that are temporally shaped. "Continuous curvature", Lynn says "is the graphical and mathematical model of the imbrication of multiple forces in time." [16] Dimension in Lynn's design vocabulary is not conceived in terms of ideal units and discrete components, rearrangeable in empty space. Instead, the assembly of a large number of components loses its modular quality: Wholes and parts are no longer discrete points, but a continuous stream of relative values inseparable from the creation of their form. [17] Conceived as a current of forces, the subdivision of the components of a form is more complex than in empty space. Instead of a neutral abstract space which is an empty container in which given components can be discretely located according to their x, y, and z co-ordinates, space is an active force of design. Significantly, in terms of the difference between Lynn's vocabulary of design and a classic model of empty space and ideal structures there is no essential structure to the forms that manifest which exist relative to the shapes which occur. There is no deviation from an ideal. Forms occur according to their own logic of differentiation and exchange of which active space is an irreducible part. [18] Occurring according to their logic of differentiation and exchange, symmetry breaking is not necessarily arbitrary or chaotic but is co-extensive with the logic of interactions as they occur. Randomness or chaos is therefore not in simple opposition to a logic of forms.

As proteins are curved structures there is an obvious applicability of Lynn's vocabulary to protein design. However, while there appears to a reflexive fortuitousness in proposing to apply an architectural vocabulary inspired by an early geneticist, I am not interested in Lynn's language of design because it is underpinned by an essential definition of biological life: A definition that will replace the error of others on which the opposition of design and random emergence is based. I am not suggesting any sort of corrective correspondence between Lynn's vocabulary and a definition of life drawn from Bateson. On the

contrary, it is the shift from a language of given components and ideal forms to a language of differentiation and exchange which does not correspond to any pre-existing capacity that I suggest offers the opportunity to develop a language of design articulating the potential for novel proteins to occur, and indeed novel biological systems generally. When emergence is asserted to be a pre-existing capacity it is equated with life itself. Equated with life, both design and the randomisation of DNA are placed in relation to its essential capacity: *the design of life, the randomisation of life (*even as the latter is asserted to correspond to its inherent definition) and their opposition is drawn. Exploring the potential of structures to form during temporal interaction on the other hand, shifts the language of design understood as a manipulation *of* life, to designing *for* life, for the way it may dynamically occur in future.

With Hugh Fisher of the Australian National University, I have begun to explore just such a language of change. Drawing on the structurally creative aspect of time and force we have begun to address the possibility of developing a language of protein design which steps outside of the opposition between design and randomisation on the basis that life is defined through its inherently random capacity for emergence.

While I'm unable to discuss our attempt so far in this context, and I've only offered a sketch of Lynn's work as it might be applied to protein design, seriously entertaining the idea of such a vocabulary I suggest generally offers the opportunity to shift arguments in synthetic biology away from the design *of* life to the design *for* life.

***References and Notes:***


1. *Adrian Mackenzie, "Design in Synthetic Biology," in BioSocieties 5, no. 2 (2010): 192-193*
2. *Ibid., 195.*
3. *Ibid., 186-193, 195.*
4. *Pier Luigi Luisi, "The Synthetic Approach in Biology: Epistemological Notes for Synthetic Biology," in Chemical Synthetic Biology, eds. Pier Luigi Luisi and Christiano Chiarabelli, 352 (Hoboken, NJ: Wiley Scientific, 2011).*
5. *Baker, David, quoted in Ewen Callaway, "The Shape of Protein Structures to Come," Nature News, November 16, 2007, http://www.nature.com/news/2007/071016/full/449765a.html (accessed November 3 2010).*
6. *Adrian Mackenzie, "Design in Synthetic Biology," in BioSocieties 5, no. 2 (2010): 187.*
7. *Pier Luigi Luisi, "The Synthetic Approach in Biology: Epistemological Notes for Synthetic Biology," in Chemical Synthetic Biology, eds. Pier Luigi Luisi and Christiano Chiarabelli, 352 (Hoboken, NJ: Wiley Scientific, 2011).*
8. *Ibid., 351.*
9. *Ibid., 352.*
10. *Christiano Chiarabelli, "Searching for de novo Totally Random Amino Acid Sequences" in Chemical Synthetic Biology, eds. Pier Luigi Luisi and Christiano Chiarabelli, 157 (Hoboken, NJ: Wiley Scientific, 2011).*
11. *Ibid., 156-157.*
12. *Pier Luigi Luisi, "The Synthetic Approach in Biology: Epistemological Notes for Synthetic Biology," in Chemical Synthetic Biology, eds. Pier Luigi Luisi and Christiano Chiarabelli, 343, 352 (Hoboken, NJ: Wiley Scientific, 2011).*
13. *Christiano Chiarabelli, "Searching for de novo Totally Random Amino Acid Sequences" in Chemical Synthetic Biology, eds. Pier Luigi Luisi and Christiano Chiarabelli, 172 (Hoboken, NJ: Wiley Scientific, 2011).*
14. *Greg Lynn, "Greg Lynn on Calculus in Architecture," TED official Web Site, filmed in 2005, posted in 2009, http://www.ted.com/talks/greg_lynn_on_organic_design.html (accessed June 10 2010).*
15. *Greg Lynn, Animate Form (New York: Princeton Architectural Press, 1999), 18.*
16. *Ibid., 23.*
17. *Ibid.*
18. *Ibid.*