REVIEW

Building a Parallel Metabolism within the Cell

Aleksandra Filipovska[†,‡] and Oliver Rackham[†,*]

[†]Western Australian Institute for Medical Research and Centre for Medical Research, University of Western Australia, Perth, Western Australia, Australia, and [‡]School of Biomedical, Biomolecular and Chemical Sciences, University of Western Australia, Crawley, Western Australia, Australia

ABSTRACT One of the key aims of synthetic biology is to engineer artificial processes inside living cells. This requires components that interact in a predictable manner, both with each other and with existing cellular systems. However, the activity of many components is constrained by their interactions with other cellular molecules and often their roles in maintaining cell health. To escape this limitation, researchers are pursuing an "orthogonal" approach, building a parallel metabolism within the cell. Components of this parallel metabolism can be sourced from evolutionarily distant species or reengineered from existing cellular molecules by using rational design and directed evolution. These approaches allow the study of basic principles in cell biology and the engineering of cells that can function as environmental sensors, simple computers, and drug factories.

According to Francis Crick, "The ultimate aim of the modern movement in biology is to explain all biology in terms of physics and chemistry" (*1*). This reductive approach has proven especially powerful in determining the structures and functions of the molecules that make up cells and organisms. The development of large-scale genomic and proteomic technologies has further revolutionized our ability to study biological problems (*2, 3*). The resulting data sets of molecular components, interaction networks, and phenotypes, combined with mathematical modeling, provide fertile ground for illuminating how biological systems function. These studies have revealed remarkable complexity in signaling pathways (*4*), new principles in gene expression (*5*), and new ideas regarding the origin of complexity in multicellular organisms (*6*). However, they have also provided new challenges, such as the difficulty in explaining observed phenotypes in terms of modifications of the cell's components, a lack of concordance between network structure and function, and the recalcitrance of most systems to rational redesign. These challenges have led to a rise in holistic concepts such as emergence. Synthetic biology combines aspects of chemistry, molecular biology, and engineering to address these issues, adding synthesis to the traditional tools of observation and analysis in biology. Synthetic biology efforts follow a strategy of constructing deliberately simplified systems to elucidate molecular and cellular processes from first principles (*7–10*). The synthesis of engineered systems provides a tool to test theories developed from the observation and analysis of natural systems and offers an opportunity to create systems with practical technological applications.

Properties of Natural and Engineered Cellular Systems. Cellular metabolic pathways have adapted to efficiently tolerate wide variation in their environment (*11, 12*). This is generally achieved by breaking pathways into modules with discrete functions. A molecule

*Corresponding author, rackham@waimr.uwa.edu.au.

Received for review September 5, 2007 and accepted November 9, 2007.

Published online January 4, 2008

10.1021/cb700185e CCC: $40.75

TABLE 1. Features of engineered pathways in cells and *in vitro*

Cellular pathways	*In vitro* pathways
Large variety of small molecules available for use as precursors	All components must be added individually
Stochastic effects due to low concentrations of individual components in small volumes	Bulk effects due to large volumes with high concentrations of individual components
Components can be added using generic transgenic approaches	Components must be isolated using empirically optimized techniques
Systems typically directed toward homeostasis	Optimized for product formation
Predicting quantitative behavior is difficult	Knowledge of rates and interaction affinities allows quantitative simulation
Additions can cause pleiotropic effects	Additions have predictable and easily measured effects
Energy generation systems hardwired in system	Energy generation systems must be added
Cells must be maintained within a tight temperature and pH range	Reaction environment can be optimized by varying pH and temperature
Cells grow, divide, and regenerate	Components must be supplemented
Artificial pathway can be linked to many, often complex, cellular phenotypes	All functions must be added individually
Context enables accurate modeling of natural pathways	Reconstruction of authentic cellular conditions is extremely difficult

that has only one function that cannot be divided (such as the enzyme glucose-6-phosphate dehydrogenase) is defined as a module. Two discrete activities of a molecule that can function in isolation (such as the DNA binding and ligand recognition domains of some transcription factors) are defined as two separate modules. Modules within pathways exploit weak interactions, enabling them to rearrange into structures with activities better suited to the new environmental conditions. Cross-talk between modules must be maintained to allow feedback between related pathways and prevent futile metabolic cycles. Investigating pathways within the cell is challenging because of the many interactions within these networks. *In vitro* studies avoid these challenges by reconstituting a pathway of interest by using purified components; however, they suffer from some important drawbacks (Table 1). Most notably, they cannot accurately model the behavior of cellular pathways, cannot use the many building blocks for biosynthesis, and cannot induce complex phenotypic responses.

The modularity of cellular systems provides a way to reengineer them to perform modified or new functions. However, the network of interactions between natural cellular modules that enables adaptive rearrangements and cross-talk imposes unwanted constraints on these efforts. Lessons from nature show how engineered pathways can escape these limitations. Naturally occurring cellular modules avoid deleterious interactions between unrelated systems by temporal, spatial, and functional compartmentalization. Spatial compartmentalization can be achieved by using membrane barriers (typical of membrane-bound organelles such as nuclei) or *via* complex structures (such as the protease active sites buried within the proteinaceous barrel of the proteasome). Temporal compartmentalization is generally achieved by regulating expression of modules so that they are never found together at the same time (*e.g.*, certain proteins expressed during either glycolytic or respiratory metabolism in yeast (*13*)). Functional compartmentalization can be achieved *via* structural modifications that prevent interactions between modules, despite the modules being present at the same sites at the same times (exemplified by the aminoacyl-tRNA synthetases, each of which can load a specific amino acid onto a designated tRNA from a pool of superficially similar tRNAs (*14*)). Molecules and pathways with this property are said to be orthogonal. For engineering purposes, orthogonality provides the most flexible solution to iso-

late distinct activities within the cell given the paucity of membrane-bound compartments in individual cells and the challenges involved in modifying the temporal expression of multiple components simultaneously.

Orthogonality in Cellular Systems. The synthesis of engineered pathways by assembling existing non-orthogonal modules in novel combinations has provided important insights into the functions of cellular pathways (*15*). However, the use of orthogonal modules provides a number of advantages for studying the behavior of model systems and the engineering of new functions. Building a variety of synthetic systems requires that the components are truly modular so that their behavior can, to some degree, be predicted. This enables existing pathways to be reconfigured and new pathways to be built with little empirical optimization. Furthermore, the accurate modeling of simple cellular systems requires the new modules to function without any interference from other cellular processes. Such interference would confound computational modeling and necessitate that the influence of changes in the cellular network be quantitated whenever the cell physiology was altered. The engineering of new function at the expense of a molecule's original function is not possible when its function is linked to essential metabolic pathways, unless an orthogonal approach is taken.

The availability of orthogonal parts generated by natural evolution is limited, and those that have been identified to date are typically derived from regulatory systems that must distinguish between different environmental and cell states. Only genes that are present in multiple copies in the cell can be tailored for functional compartmentalization by evolution, making it unlikely that orthogonal modules with essential cellular activities will be present in the same cell. In some cases, this can be circumvented by the accumulation of compensatory mutations over time, which can render natural interactions orthogonal if transferred from a distantly related organism into the organism of interest. There are several advantages that make the engineering orthogonal function in endogenous molecules attractive. Engineered endogenous molecules can maintain productive upstream or downstream interactions that link the engineered pathway to signals, make use of cellular building blocks, or enable complex phenotypic responses (*16*). The endogenous parts of these engineered molecules allow them to retain their known structural properties, whereas the engineered parts carry out the orthogonal function of the component, which makes their activity easier to predict. Most importantly, engineered components can be created with desirable new functions not found in nature (*17, 18*).

Engineering Orthogonal Function. In the process of natural evolution, new function is thought to have arisen by the duplication of existing genes, followed by mutation and selection of the new copies to perform new functions (*19*). One advantage of working with microorganisms is that very rare mutants can be identified within large populations given an environment in which only the mutant of interest can survive. Therefore, an analogous process of mutation and selection can be achieved in the laboratory if the activity of the mutant of interest provides a selectable phenotype. This can most easily be achieved by linking the activity of the mutant of interest to the activation or repression of gene expression, to take advantage of a number of well-characterized genes that confer prototrophy on defined growth media or resistance to antibiotics and metabolic analogues (*20, 21*). In addition to selections, screening can also be used to engineer orthogonal properties in a molecule. Unlike selections, screens do not require the activity of interest to be linked to cell survival, allowing a much greater variety of methods to be used in the analysis of mutants and providing the opportunity to measure protein and nucleic acid activities directly. However, because each mutant must be individually evaluated, screens are generally laborious, limiting the diversity of mutants that can be tested (*22*).

Four general strategies are available to generate a pool, or library, of mutant molecules: rational design, random mutagenesis, recombination, and targeted mutagenesis (*22*). In rational design, individual sequence changes are made on the basis of structural and mechanistic knowledge (*23, 24*). Rational design is primarily limited by our incomplete understanding of the relationship between the primary sequence of macromolecules and their structure and activity (*25*). Despite some success, rational design cannot be applied in any generic fashion and relies on detailed structural knowledge and computational modeling approaches that are specific to the molecule of interest (*26, 27*). In contrast, random mutagenesis uses error-prone polymerase chain reaction (PCR), chemical mutagens, or replication by a mutator bacterial strain to introduce small numbers of point mutations in the gene of interest and requires no knowledge of a molecule's structure or mechanism (*28*). Al-

though random mutagenesis has been used with great success to improve the existing qualities of enzymes, it is less suited to the engineering of orthogonal modules. This is because reassignment of active site recognition and binding interfaces usually requires multiple changes within a confined sequence space, and such mutants are rarely generated by error-prone PCR (*22, 29*). Recombination generates diversity by shuffling a set of related genes to yield new combinations not found in nature (*30*). This approach can generate more substantial changes than random mutagenesis but is limited by the availability of gene variants. Targeted mutagenesis represents the middle ground between knowledge-based and assumption-free approaches (*22, 24*). Mechanistic and structural information is used to choose regions of interest that are all mutated while leaving the rest of the gene unchanged (*31*). These libraries allow multiple changes to be generated simultaneously in an active site or binding interface and are well suited to engineer mutants with orthogonal or reassigned functions. In some cases, a combination of these approaches has proven to be effective (*32*). Ultimately, the choice of a mutagenesis approach will depend on the target gene and the desired orthogonal activity.

Selections for orthogonal activity typically require two steps, each of which filters a population of cells to eventually isolate those containing the orthogonal mutants (Figure 1). A negative selection eliminates the original function or interaction. A positive selection enriches for molecules with the new orthogonal function or interaction. In some cases, a screen, rather than a selection, can be used in one of these steps. These steps may need to be reiterated to improve the selection process. Mutants that survive the selection are characterized in detail to confirm that they have the desired properties. These new properties are usually further tested in simple synthetic pathways to validate their use for building new pathways in cells. In the following examples, we hope to illustrate the current approaches for engineering orthogonal cellular pathways and the insights that can be gained from such endeavors.

Building Proteins That Use Unnatural Energy Sources. Because ATP and GTP are the most widely used energy currencies in the cell, it is challenging to determine their roles in protein function or to isolate the activities of these proteins within an engineered cell. In one of the first examples of a protein engineered to recognize an unnatural substrate, Hwang and Miller cre-

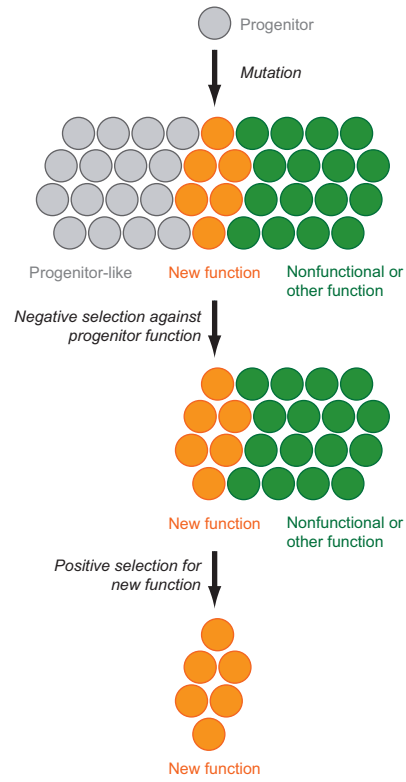

Figure 1. A generalized scheme for the selection of unnatural orthogonal activities. A large, diverse pool of mutants is first generated from the progenitor molecule of interest. Within this pool, there can be subpopulations of three types: those with activities similar to their progenitor, those with the new activity of interest, and those that are nonfunctional or possess another function not of interest in the context of the experiment. The first selection step eliminates cells containing progenitor-like activities, a negative selection. The second step eliminates cells containing nonfunctional molecules, a positive selection.

ated a bacterial elongation factor Tu (EF-Tu) mutant able to accept xanthine triphosphate (XTP) as a substrate (*33*). The crystal structure of guanosine 5′-diphosphate (GDP) bound to EF-Tu was used as a guide to engineer a mutant that lacked a hydrogen bond donor to coordinate GTP. Instead, XTP provided a hydrogen bond acceptor that functionally complemented the mutation enabling the mutant EF-Tu to use XTP. This mutant EF-Tu·XTP pair is truly orthogonal in that normal cellular GTPases cannot use XTP while the engineered protein uses XTP but not GTP. The orthogonal mutant EF-Tu·XTP pair allowed the number of GTP molecules used by EF-Tu

per amino acid coupling on the ribosome to be determined (*34*), and this approach has been used generically to examine the function of other GTPases (*35*). An analogous approach combining protein engineering and unnatural chemical analogues has been pioneered by Shokat and coworkers to dissect the roles of specific protein kinases in cells. Protein kinases regulate signaling pathways by using ATP as a phospho donor to phosphorylate serine, threonine, or tyrosine residues on specific proteins. Mutations that increase the size of the ATP-binding site of protein kinases (by replacing amino acids with bulky side chains with those carrying smaller side chains) enable these mutants to efficiently use ATP analogues with aromatic substituents instead of natural ATP (*36*). The substrates of these orthogonal protein kinases with ATP binding site mutations have been identified *via* radiolabeled ATP analogues or analogues that facilitate the phospho transfer of an affinity label (*37, 38*). In future applications, orthogonal energy currencies such as ATP analogues and XTP may allow engineered networks to function independently of the cell's endogenous energy sources.

Controlling the Function of Engineered Transcription Factors with Small Molecules. Transcription factors are essential for gene expression, and extensive research has revealed many details of their structure and function (*39*). Transcription factors were used in the first efforts to construct unnatural regulatory circuits in living cells (*40–42*) because they are highly modular in structure and activity. For example, it was recognized early on that the DNA-binding and transcription regulation activities of many transcription factors could function independently of each other, which led to the invention of the widely used yeast two-hybrid system for the identification and analysis of protein–protein interactions (*43*). In addition, the transcription activation domains of nuclear hormone receptor transcription factors are only active upon binding of a ligand (*44*). The use of nuclear hormone receptors in synthetic circuits is limited because the natural ligands will activate the endogenous ligand responsive pathways in the cell at the same time as the engineered pathway. To overcome this, a number of research groups have modified the ligand-binding domains of nuclear hormone receptors with the goal of engineering an orthogonal response to an unnatural ligand (*45*). Doyle and colleagues (*46*) began with a series of synthetic retinoids that were designed to target the retinoid X receptor (RXR) but failed to act as agonists. A rational design approach was employed to modify individual amino acids within the active site of the RXR on the basis of the crystal structure of a related receptor. Mutations that decreased activation in the presence of the natural ligand (9-*cis*-retinoic acid) or enabled a response to an unnatural ligand (LG335) were combined in a new set of mutants. A triple mutant receptor was identified that had significantly increased activity to LG335 and that was also relatively insensitive to 9-*cis*-retinoic acid. However, the ~100-fold change in specificity was not sufficient to provide a functionally orthogonal receptor in living cells.

To engineer a truly orthogonal nuclear hormone receptor, a modified yeast two-hybrid system was used to evolve RXRs that were more selective for LG335 (*47*). In this system, the RXR ligand-binding domain is tethered to the promoter of the *ade2* purine biosynthetic gene *via* the well-characterized Gal4 DNA-binding domain and its cognate DNA response element (Figure 2). Upon ligand binding, a conformational change in the ligand-binding domain facilitates the recruitment of a protein fusion between the Gal4 activation domain and RXR's natural effector molecule (ACTR), turning on transcription of the *ade2* gene and allowing survival on media without adenine. A library of RXRs mutants was generated in which six key residues in the ligand-binding pocket were mutated and was introduced into the modified yeast two-hybrid strain to select for domains that bind LG335. The surviving mutants were then analyzed by using a β-galactosidase reporter gene to eliminate receptors that were activated by the natural ligand. This enabled a further 25-fold increase in specificity over the previous rationally engineered near-orthogonal receptor. The applicability of this approach has been corroborated in a recent study in which the estrogen receptor was engineered to recognize the unnatural ligand 4,4′-dihydroxybenzil using a modified yeast two-hybrid system (*48*). In this work, reiterated targeted mutagenesis of the ligand-binding pocket was

KEYWORDS

Cellular network: Description of the functional connections between cellular modules.

Compartmentalization: A phenomenon used to isolate different molecules within the cell *via* physical barriers, temporally distinct synthesis, or an absence of molecular interactions. For example, membrane-bound organelles can have different pH, availability of molecular precursors, and enzyme systems. This enables the cell to carry out different metabolic activities without interference.

Directed evolution: A method that harnesses the power of Darwinian selection to evolve proteins or nucleic acids with desirable properties not found in nature. A typical directed evolution experiment involves two steps. First, the gene of interest is mutated and/or recombined at random to create a large library of gene variants. Then, the library is screened by a high-throughput assay or subjected to selection to identify variants with desirable properties.

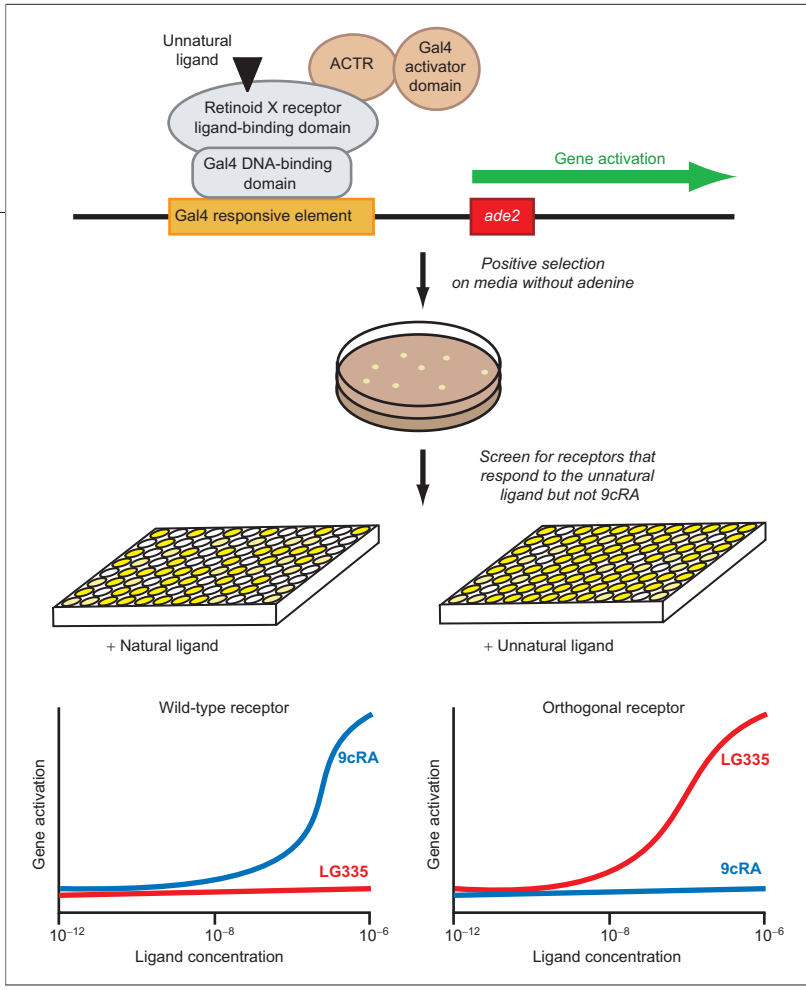

Figure 2. Engineering an orthogonal nuclear hormone receptor that responds to an unnatural ligand. The ligand-binding domain of the retinoid X receptor is attached to the promoter of the *ade2* biosynthetic gene by the Gal4 DNA-binding domain–DNA interaction. The effector protein for the ligand-binding domain (ACTR) is fused to the Gal4 activation domain. If the unnatural ligand (LG335, in this case) interacts with an engineered ligand-binding domain, transcription of the *ade2* gene allows the cell to grow to a colony on media without adenine. Colonies are subsequently screened using a colorimetric assay to identify those that respond specifically to the unnatural but not the natural ligand (9-*cis*-retinoic acid, 9cRA).

combined with random mutagenesis of the entire domain to generate an orthogonal receptor with an impressive shift in specificity of $>10^7$-fold for the new ligand. These studies provide a generic approach for engineering orthogonal nuclear hormone receptor–ligand pairs and demonstrate the power of directed evolution approaches compared with rational design.

A number of natural small-molecule-responsive transcription factors of bacterial origin have been shown to function in a mutually orthogonal manner (lacI, cI, tetR, and luxR). This has facilitated the construction of the vast majority of synthetic pathways to date (*49–51*). However, the design of more complex network architectures will require a substantially wider selection of orthogonal modules (*18*). The modular nature of base recognition by DNA-binding proteins such as zinc fingers has enabled the rational engineering of new DNA-binding properties (*52*), although creating transcription factors that respond to new ligands has proven challenging. Evolutionarily related proteins that recognize distinct ligands have been studied, but these show significant cross-reactivity with ligands from their homologous family members (*53*). Furthermore, it would be advantageous to have a set of factors that differed in their ligand binding specificities but otherwise shared near-identical properties, such as activation profiles in response to ligand concentration, DNA-binding affinity, and protein stability. This could be achieved if a common scaffold could be reengineered to respond in an orthogonal manner to a variety of different ligands. Arnold and coworkers (*54, 55*) have addressed this challenge in their studies of the luxR transcription factor from the marine bacterium *Vibrio fischeri*. *V. fischeri* is able to colonize fish and squid, which exploit the bacterium's bioluminescent properties to attract prey and even as a means of camouflage. *V. fischeri* uses luxR to detect the levels of 3-oxy-hexanoyl-homoserine lactone (3OC6HSL), which it synthesizes itself. When the population density of cells increases upon successful colonization of the light-emitting organs of its hosts, levels of 3OC6HSL rise and luxR is activated to turn on the expression of genes encoding enzymes required for light production. Natural luxR has a very strict specificity for 3OC6HSL and does not recognize related straight chain HSLs such as decanoyl-homoserine lactone (C10HSL). To reengineer luxR to specifically recognize C10HSL but not 3OC6HSL, error-prone PCR was used to generate a library of luxR mutants that were introduced into cells containing a chloramphenicol resistance gene under the control of a luxR-responsive promoter (Figure 3). Plating these cells on media with both C10HSL and chloramphenicol selected for luxR variants that could now recognize the unnatural ligand. One such mutant was found to have a broader, or promiscuous, function, recognizing 3OC6HSL, C10HSL, and a number of other straight chain HSLs (*55*). Mutant luxR genes that survived the first selection were introduced into cells containing a constitutively expressed β-lactamase along with a β-lactamase inhibitory gene under the control of a luxR-responsive promoter. Plating these cells on media with the natural ligand 3OC6HSL and the β-lactam antibiotic carbenicillin selected for luxR variants that no

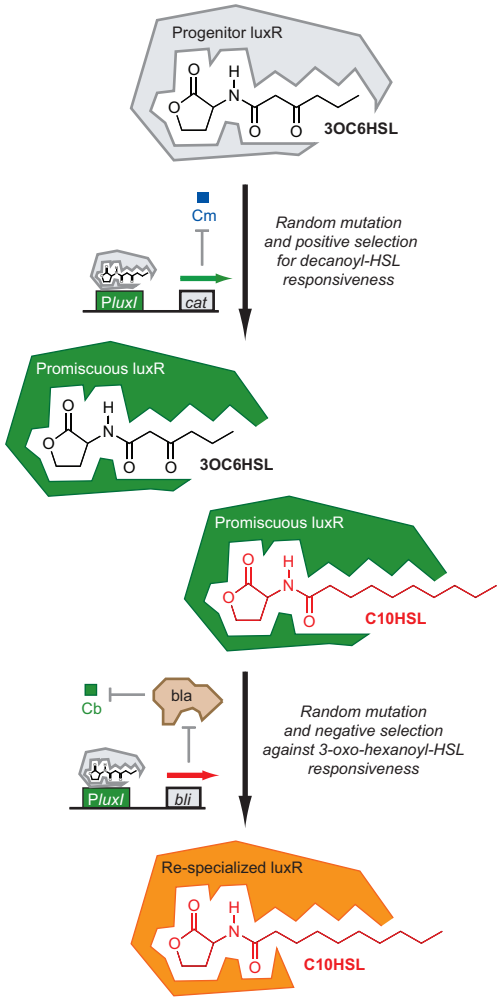

Figure 3. An engineered orthogonal quorum-sensing transcriptional activator. The progenitor luxR activates transcription in response to 3OC6HSL but not C10HSL. Random mutagenesis followed by selection for activation of the *cat* gene in the presence of C10HSL by plating on chloramphenicol (*Cm*) generated a promiscuous mutant that bound both 3OC6HSL and C10HSL. Further mutagenesis and selection against activation of the *bli* gene by growing β-lactamase (bla) expressing cells in the presence of carbenicillin (Cb) and 3OC6HSL generated a respecialized luxR with specificity for C10HSL at the expense of 3OC6HSL recognition.

longer recognize the unnatural ligand. The final products of these selections were found to recognize C10HSL but not 3OC6HSL, with an overall 50,000-fold change in specificity. All of these reengineered clones had the same active site substitution, in addition to the mutations found to confer promiscuous function to the clone obtained from the positive selection only. This observation concurs with a widely held theory that respecialized function must be obtained *via* an evolutionary intermediate with broad activity (*56*). The ability to engineer transcription factors that respond to new ligands will be vital for the development of cell-based sensors.

Reengineering Cellular Translation. In response to the intense transcriptional focus of early synthetic biology efforts, a number of research groups are now investigating the creation of synthetic circuits based on the regulated translation of specific genes (*57*, *58*). Small RNA "riboregulators" have been engineered that activate the expression of specific mRNAs (mRNAs) by allowing the ribosome to access them (*59*). The mRNA target of the riboregulator is a short *cis*-repressed (CR) sequence upstream of the start codon of a gene. In the absence of a cognate riboregulator, the CR sequence folds into a stable structure that obscures the start codon. The riboregulator is designed to form a stable intermolecular RNA complex with the CR, revealing the start codon and allowing translation. These elements are truly modular and can be used in combination with any promoter or gene to activate gene expression when the riboregulator is present. Although RNA can adopt a variety of tertiary structures that are difficult to predict, simple riboregulators can be rationally engineered following the basic principles of base pairing (*60*). RNA can be engineered with more complex functions, such as the ability to bind molecules other than nucleic acids, using *in vitro* selections based on binding to an immobilized ligand (known as the systematic evolution of ligands by exponential enrichment or SELEX (*61*, *62*)). Bayer and Smolke (*63*) recently incorporated such an RNA sequence, which can selectively bind the xanthine derivative theophylline, into a riboregulator that consequently only bound to its target mRNA in the presence of theophylline. This work demonstrates that RNA molecules generated by rational design and directed evolution can be successfully integrated to create allosteric riboregulators, providing opportunities to build RNA networks with increasingly complex characteristics.

In an alternative approach to regulate translation, recent research has involved reengineering the ribosome. In bacterial cells, the ribosome identifies the start sites for protein synthesis on mRNAs *via* base pairing between an mRNA-binding sequence (MBS) in the 16S

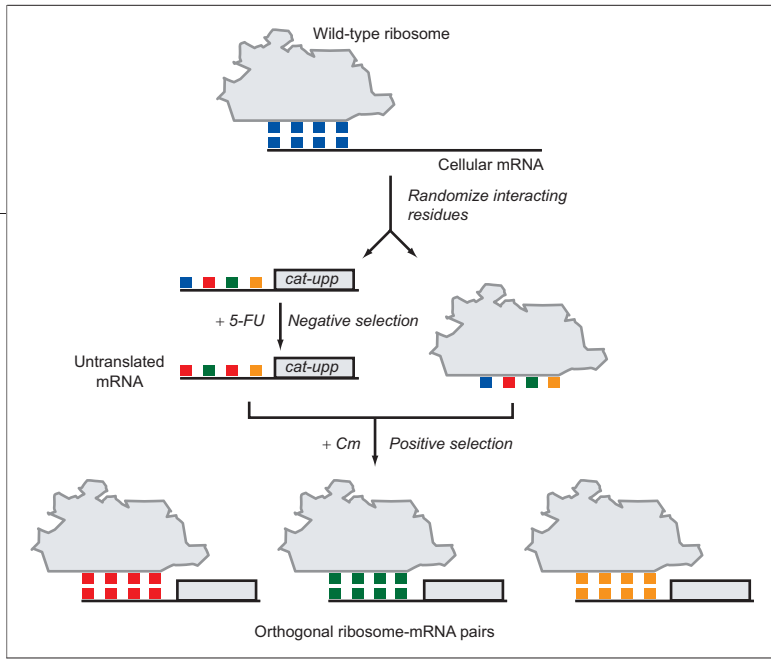

Figure 4. Evolution of orthogonal ribosome—mRNA pairs. Ribosome—mRNA pairs that work together without cross-talk from their naturally occurring counterparts in the cell were engineered *via* a two-step selection process. First, a pool of mRNAs with randomized RBS sequences were assembled on a *cat-upp* fusion gene template. Those that are not translated by endogenous ribosomes were selected by growth on 5-FU, which results in cell death if the *upp* gene product is synthesized. In the second step, a library of mutant ribosomes was combined with the surviving mRNAs, and restoration of *cat-upp* translation was selected for by growth on chloramphenicol (*Cm*).

rRNA (rRNA) and a ribosome-binding site (RBS) upstream of the authentic start codon (*64, 65*). This system was reengineered to create new mRNAs and ribosomes with orthogonal properties, consisting of mRNAs containing RBS sequences that are not translated by endogenous ribosomes (orthogonal mRNAs, O-mRNAs), and ribosomes that translate the O-mRNA but not any of the thousands of cellular mRNAs (orthogonal ribosomes, O-ribosomes). To achieve this result, a library of new potential RBSs was produced by creating all of the possible combinations of residues at the nucleotide positions of the existing RBS upstream of a gene fusion between the chloramphenicol resistance gene and a uracil phosphoribosyltransferase gene (*cat-upp*). The *upp* gene product converts 5-fluorouracil (5-FU) to 5-fluorodUMP, which strongly inhibits the essential thymidylate synthase enzyme and causes cell death. O-mRNA sequences that are not translated by the endogenous ribosome were selected by plating the library of RBS mutants on 5-FU (Figure 4) (*66*). These O-mRNAs were then combined with ribosomes containing randomized 16S rRNAs and grown on chloramphenicol to find mutant ribosomes that specifically translate O-mRNAs. These ribosomes do not significantly translate cellular transcripts nor significantly alter the growth rate of the cells that contain them, confirming their orthogonal properties (*66*). This study provides evidence that cellular modules that are essential for life can nonetheless be duplicated and respecialized given an appropriate experimental approach.

The O-mRNA and O-ribosome system enables a fundamentally new way to regulate gene expression. Production of O-ribosomes within the cell allows the selective translation of a single, previously silent, mRNA. Because multiple O-ribosomes and O-mRNAs can be functionally expressed in a single cell, this approach was able to create cells capable of performing simple Boolean logic (*64, 67*). In one example, an "AND" function was created, in which the synthesis of a functional enzyme depends on the presence of two O-ribosome·O-mRNA pairs. The endogenous ribosome is essential to cell function, and many mutations are lethal. However, O-30S ribosomes are not essential, and it is possible to further specialize their function. This capacity makes O-ribosomes a good tool for investigating structure–function relationships in rRNA. Recently, mutagenesis of O-ribosomes allowed the large-scale identification of residues required for ribosome subunit association (*68*). The use of an analogous approach for engineering O-ribosomes in nonprokaryotic cells is complicated by the complex, factor-dependent mechanisms of translation initiation in eukaryotes and archaea.

Reengineering the Genetic Code. Despite advances in solid-phase peptide synthesis and semisynthetic methodologies, our ability to engineer the structure and function of proteins is still in its infancy. In particular, the incorporation of unnatural building blocks at specific sites using these methods is only applicable to peptides and very small proteins (*69, 70*). By contrast, the ribosome synthesizes protein with remarkable processivity and fidelity (*65*). Furthermore, the ribosome has been found to exhibit a relatively wide tolerance for a variety of unnatural substrates, when tested *in vitro* with chemically acylated tRNAs (*71*). Despite these advantageous properties, ribosomal protein synthesis is limited to the common 20 amino acid building blocks in all organisms, with the rare exceptions of pyrrolysine and selenocysteine (*72*). The development of chemical methods for the *in vitro* misacylation of tRNAs with unnatural amino acids has enabled the ribosome to be hijacked to synthesize proteins containing unnatural amino acids *in vitro* and in microinjected oocytes (*73–75*). More recently, Schultz and co-workers have created systems that allow aminoacyl-tRNA synthetases, the enzymes that load amino acids onto tRNAs, to be engineered to specifically incorporate unnatural amino acids into pro-

teins expressed in both bacteria (*76*) and yeast (*77*). Although the approaches use different selection markers in either organism, they are conceptually analogous (Figure 5). Adding an unnatural amino acid to the genetic code of a living cell requires three orthogonal modules: an orthogonal codon, an orthogonal tRNA that recognizes that codon, and an orthogonal aminoacyl-tRNA synthetase that can charge that tRNA. Although all 64 possible three-base codons are used in the genetic code, only 61 are recognized by tRNAs; the remaining three (UGA, UAG, and UAA) are recognized by proteins known as release factors, which stimulate hydrolysis of the newly synthesized protein and function as stop codons (*78*). The UAG stop codon is rarely used and can be reappropriated as a pseudo-orthogonal codon, given that suppressor tRNAs can compete with the endogenous release factor for UAG recognition and that endogenous tRNAs do not recognize the codon to compromise the fidelity of proteins that are synthesized by suppression of UAG codons. To find orthogonal aminoacyl-tRNA synthetase-tRNA pairs, modules were sourced from organisms from different superkingdoms, taking advantage of fortuitous divergence in tRNA-synthetase recognition elements.

The amino acid specificities of the orthogonal synthetases are reassigned by targeted mutagenesis of the synthetase active site followed by iterative positive and negative selections (*76, 77*). These selections are based on well-characterized genes that are modified to incorporate UAG codons so that their activity is only observed if a mutant synthetase is able to load the orthogonal tRNA. In bacteria the positive selection is based on resistance to chloramphenicol in the presence of the unnatural amino acid, selecting for functional synthetases; the negative selection is based on expression of the toxic barnase protein in the absence of the unnatural amino acid, so that synthetases that are able to use any of the endogenous amino acids within the cell are eliminated. Whereas in yeast the positive selection is based on expression of a gene that enables yeast to survive in the absence of histidine in the presence of the unnatural amino acid, the negative selection is based on expression of a gene that converts a metabolic analogue to a toxic product. These approaches have proved to be remarkably general, allowing >30 unnatural amino acids with diverse structures to be incorporated into bacterial and yeast proteins with high fidelity and good yields (*79*).

The ability to introduce unnatural amino acids to the genetic code of living organisms has enabled an array of new approaches for manipulating protein function within cells. For example, the incorporation of the photoisomerizable amino acid *p*-azophenyl-phenylalanine into the catabolite activator protein enabled its DNA-binding activity to be photoregulated (*80*). Incorporation of photocaged cysteine into the proapoptotic caspase 3 enzyme allowed its protease activity to be specifically turned on by irradiation with UV light (*81*). Amino acids with azide and alkyne side chains provide useful reagents for subsequent protein modification (*82–84*). Azides and alkynes are not found in native biological systems, and despite being relatively unreactive in these systems, they can nonetheless take part in highly selective reactions with exogenously delivered small molecules (containing alkyne or azides functionalities, respectively), allowing "bio-orthogonal" chemical reactions within living cells (*85, 86*).

The key problem limiting the expansion of the genetic code is that cells contain a full complement of tRNAs and aminoacyl-tRNA synthetases that incorporate natural amino acids in response to all 61 sense codons and release factors for the termination of protein synthesis upon recognition of the three codons. In recent work, Chin and coworkers engineered O-ribosomes that did not recognize release factor-1 efficiently and used them to incorporate amino acids to efficiencies of >60% at a single UAG codon (*87*). This increase in suppression efficiency will allow a wider variety of experimental approaches using unnatural amino acids; however, in some cases the low levels of truncated protein that are still produced could complicate the engineering of new functions in living cells. Furthermore, reliance on the single UAG codon limits the extension of this approach to incorporate multiple different unnatural amino acids in a single protein or cell. An elegant solution to this problem is to pro-

KEYWORDS

Modularity: Characteristic of complex systems that consist of multiple units with discrete functions. Modules have the following features: they have identifiable interfaces with other modules; they can be modified and evolved with some degree of independence; and they maintain their function when isolated and rearranged.

Orthogonality: Characteristic of chemical reactions or molecular interactions that can operate in parallel with, but independent of, each other.

Screen: Evaluation of phenotypic variants within a population. An appropriate assay ensures that all members of the population are examined.

Selection: Placing organisms under conditions where the growth of only those with a particular genotype is possible.

Specificity: The ability of a molecule to distinguish one substrate from other substrates.

Synthetic biology: The engineering or design of biological systems with either improved properties or new functions.

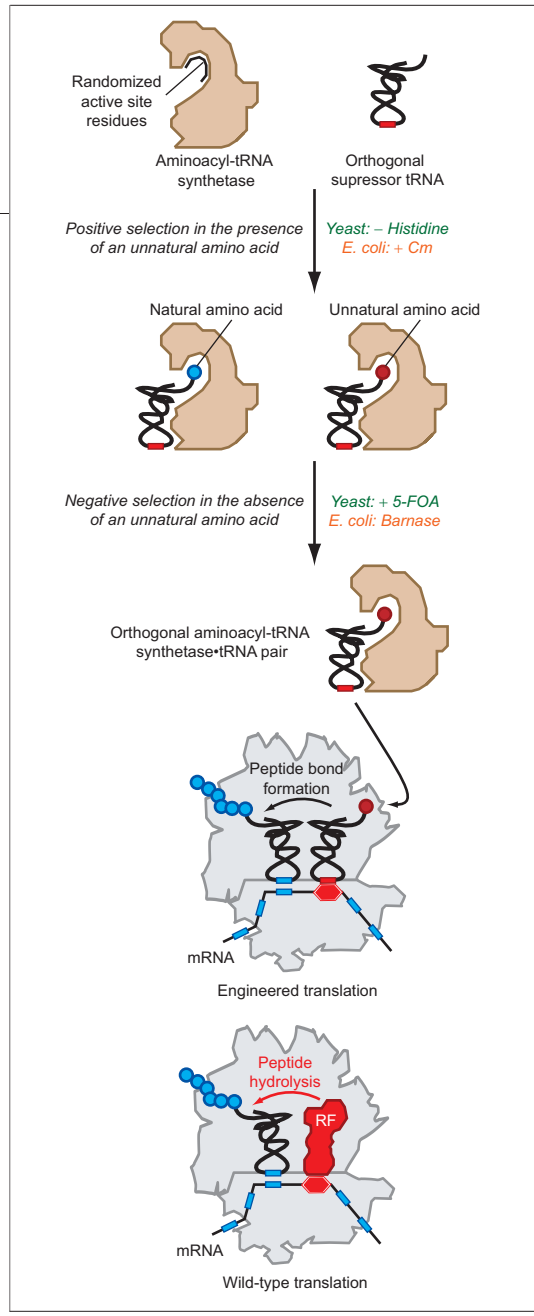

Figure 5. Expanding the genetic code. The specificity of an orthogonal aminoacyl-tRNA synthetase can be reassigned to allow the incorporation of an unnatural amino acid. A library of active site variants are generated by targeted mutagenesis. Mutant synthetases that charge their cognate suppressor tRNA and amino acid are selected in the presence of the unnatural amino acid of interest, using positive selection pressure. In a second step, only those synthetases that use the unnatural amino acid but not any other amino acid are selected, using negative selection pressure. The engineered synthetase-tRNA pair enables unnatural amino acids to be incorporated at UAG stop codons where the translational machinery would usually terminate protein synthesis.

vide additional codons to encode unnatural monomers. The addition of new codons has been achieved *in vitro* using codons containing unnatural bases. Initially, it was shown that the unnatural base isocytosine could be incorporated into a codon and specifically decoded by a tRNA with a complementary isoguanine-containing anticodon (*88*). More recently, these observations have been extended to the unnatural base pair of 2-amino-6-(2-thienyl)purine and pyridin-2-one (*89*). The development of additional unnatural base pairs (*90*), as well as mutant polymerases for the replication of nucleic acids containing these base pairs (*91*), may allow the engineering of cells that adhere to a distinct orthogonal genetic code.

Conclusions and Future Perspectives. The diversity of naturally occurring species and their vast metabolic diversity provide inspiration for new approaches and designs that combine natural elements in unnatural networks and stimulate our imagination to engineer new modules that are not found in nature. However, our ability to reengineer cellular metabolism is still in its infancy. The ability to create useful model cellular systems or cell-based technologies requires well-characterized components that can function predictably in a variety of contexts. A powerful approach to achieve this goal is to discover and engineer orthogonal modules whose function is not compromised in the presence of endogenous pathways. The principles and methods are now in place to begin this work in earnest. Sourcing orthogonal components from nature has been vital to the early establishment of synthetic biology, but to fully realize its potential directed evolution must be used to generate new orthogonal modules. The selection approaches that have been most successful to date all use genetic markers in two-step experiments. This approach can be generically applied in the engineering of almost any molecular interaction that can be linked to gene expression, in bacteria and yeast. This approach provides a powerful route to create new cellular modules with defined connections to and separation from other cellular functions. As more modules are engineered and characterized, it will become easier to link multiple orthogonal components in pathways of increasing complexity. When combined with observation and analysis of natural systems, future work will fulfill synthetic biology's promise to aid in the understanding of complex biological systems. Furthermore, the tools developed in these exploits will enable the fabrication of useful new cell-based technologies. Already, synthetic pathways have been created that allow cells to act as environmental sensors (*92*), simple computers (*50, 67*), drug and

nanomaterial factories (*93–95*), and tools for bioremediation (*96*). In the future, these types of systems may form the basis of important new technologies such as cell-based biofuel fermentors (*97, 98*). Given that the intricate and highly complex features of metazoan organ systems and body plans are thought to have been made possible by the duplication and neofunctionalization of only a small number of control genes (*99*), this provides an indication of the potential power of these approaches to engineer remarkable new functions in living organisms.

Acknowledgment: A.F. is a Howard Florey Centenary Fellow, and O.R. is a Peter Doherty Fellow. The authors are supported by the Australian Research Council and the National Health and Medical Research Council. We thank the anonymous reviewers for their helpful and constructive comments and G. J. Chin for the inspiration for the title of this Review.

REFERENCES

1. Crick, F. H. C. (1966) *Of molecules and men*, University of Washington Press, Seattle, WA.
2. Pandey, A., and Mann, M. (2000) Proteomics to study genes and genomes, *Nature 405*, 837–846.
3. Lockhart, D. J., and Winzeler, E. A. (2000) Genomics, gene expression and DNA arrays, *Nature 405*, 827–836.
4. Papin, J. A., Hunter, T., Palsson, B. O., and Subramaniam, S. (2005) Reconstruction of cellular signalling networks and analysis of their properties, *Nat. Rev. Mol. Cell Biol. 6*, 99–111.
5. Mattick, J. S. (2004) RNA regulation: a new genetics? *Nat. Rev. Genet. 5*, 316–323.
6. Martindale, M. Q. (2005) The evolution of metazoan axial properties, *Nat. Rev. Genet. 6*, 917–927.
7. Endy, D. (2005) Foundations for engineering biology, *Nature 438*, 449–453.
8. Benner, S. A., and Sismour, A. M. (2005) Synthetic biology, *Nat. Rev. Genet. 6*, 533–543.
9. Guido, N. J., Wang, X., Adalsteinsson, D., McMillen, D., Hasty, J., Cantor, C. R., Elston, T. C., and Collins, J. J. (2006) A bottom-up approach to gene regulation, *Nature 439*, 856–860.
10. Voigt, C. A., and Keasling, J. D. (2005) Programming cellular function, *Nat. Chem. Biol. 1*, 304–307.
11. Szalay, M. S., Kovacs, I. A., Korcsmaros, T., Bode, C., and Csermely, P. (2007) Stress-induced rearrangements of cellular networks: consequences for protection and drug design, *FEBS Lett. 581*, 3675–3680.
12. Han, J. D., Bertin, N., Hao, T., Goldberg, D. S., Berriz, G. F., Zhang, L. V., Dupuy, D., Walhout, A. J., Cusick, M. E., Roth, F. P., and Vidal, M. (2004) Evidence for dynamically organized modularity in the yeast protein–protein interaction network, *Nature 430*, 88–93.
13. Tu, B. P., Kudlicki, A., Rowicka, M., and McKnight, S. L. (2005) Logic of the yeast metabolic cycle: temporal compartmentalization of cellular processes, *Science 310*, 1152–1158.
14. Woese, C. R., Olsen, G. J., Ibba, M., and Soll, D. (2000) Aminoacyl-tRNA synthetases, the genetic code, and the evolutionary process, *Microbiol. Mol. Biol. Rev. 64*, 202–236.
15. Bhattacharyya, R. P., Remenyi, A., Yeh, B. J., and Lim, W. A. (2006) Domains, motifs, and scaffolds: the role of modular interactions in the evolution and wiring of cell signaling circuits, *Annu. Rev. Biochem. 75*, 655–680.
16. Kobayashi, H., Kaern, M., Araki, M., Chung, K., Gardner, T. S., Cantor, C. R., and Collins, J. J. (2004) Programmable cells: interfacing natural and engineered gene networks, *Proc. Natl. Acad. Sci. U.S.A. 101*, 8414–8419.
17. Cropp, T. A., and Chin, J. W. (2006) Expanding nucleic acid function *in vitro* and *in vivo*, *Curr. Opin. Chem. Biol. 10*, 601–606.
18. Chin, J. W. (2006) Modular approaches to expanding the functions of living matter, *Nat. Chem. Biol. 2*, 304–311.
19. Lynch, M., and Conery, J. S. (2000) The evolutionary fate and consequences of duplicate genes, *Science 290*, 1151–1155.
20. Neidhardt, F. C., Curtiss, III, R., Ingraham, J. L., Lin, E. C. C., Low, K. B., Magasanik, B., Reznikoff, W. S., Riley, M., Schaechter, M., and Umbarger, H. E. (1996) *Escherichia coli and Salmonella: Cellular and Molecular Biology*, American Society for Microbiology, Washington, DC.
21. Guthrie, C., and Fink, G. R. (2002) *Guide to Yeast Genetics and Molecular Cell Biology, Part B*, Vol. 350, Academic Press, San Diego, CA.
22. Bloom, J. D., Meyer, M. M., Meinhold, P., Otey, C. R., MacMillan, D., and Arnold, F. H. (2005) Evolving strategies for enzyme engineering, *Curr. Opin. Struct. Biol. 15*, 447–452.
23. Yoshikuni, Y., Ferrin, T. E., and Keasling, J. D. (2006) Designed divergent evolution of enzyme function, *Nature 440*, 1078–1082.
24. Chica, R. A., Doucet, N., and Pelletier, J. N. (2005) Semi-rational approaches to engineering enzyme activity: combining the benefits of directed evolution and rational design, *Curr. Opin. Biotechnol. 16*, 378–384.
25. Fetrow, J. S., Giammona, A., Kolinski, A., and Skolnick, J. (2002) The protein folding problem: a biophysical enigma, *Curr. Pharm. Biotechnol. 3*, 329–347.
26. Ashworth, J., Havranek, J. J., Duarte, C. M., Sussman, D., Monnat, R. J., Jr., Stoddard, B. L., and Baker, D. (2006) Computational redesign of endonuclease DNA binding and cleavage specificity, *Nature 441*, 656–659.
27. Dwyer, M. A., Looger, L. L., and Hellinga, H. W. (2004) Computational design of a biologically active enzyme, *Science 304*, 1967–1971.
28. Neylon, C. (2004) Chemical and biochemical strategies for the randomization of protein encoding DNA sequences: library construction methods for directed evolution, *Nucleic Acids Res. 32*, 1448–1459.
29. Morley, K. L., and Kazlauskas, R. J. (2005) Improving enzyme properties: when are closer mutations better? *Trends Biotechnol. 23*, 231–237.
30. Kikuchi, M., and Harayama, S. (2002) DNA shuffling and family shuffling for *in vitro* gene evolution, *Methods Mol. Biol. 182*, 243–257.
31. Toscano, M. D., Woycechowsky, K. J., and Hilvert, D. (2007) Minimalist active-site redesign: teaching old enzymes new tricks, *Angew. Chem., Int. Ed. 46*, 3212–3236.
32. Joyce, G. F. (2004) Directed evolution of nucleic acid enzymes, *Annu. Rev. Biochem. 73*, 791–836.
33. Hwang, Y. W., and Miller, D. L. (1987) A mutation that alters the nucleotide specificity of elongation factor Tu, a GTP regulatory protein, *J. Biol. Chem. 262*, 13081–13085.
34. Weijland, A., Parlato, G., and Parmeggiani, A. (1994) Elongation factor Tu D138N, a mutant with modified substrate specificity, as a tool to study energy consumption in protein biosynthesis, *Biochemistry 33*, 10711–10717.
35. Bishop, A., Buzko, O., Heyeck-Dumas, S., Jung, I., Kraybill, B., Liu, Y., Shah, K., Ulrich, S., Witucki, L., Yang, F., Zhang, C., and Shokat, K. M. (2000) Unnatural ligands for engineered proteins: new tools for chemical genetics, *Annu. Rev. Biophys. Biomol. Struct. 29*, 577–606.

36. Shah, K., Liu, Y., Deirmengian, C., and Shokat, K. M. (1997) Engineering unnatural nucleotide specificity for Rous sarcoma virus tyrosine kinase to uniquely label its direct substrates, *Proc. Natl. Acad. Sci. U.S.A. 94*, 3565–3570.
37. Allen, J. J., Li, M., Brinkworth, C. S., Paulson, J. L., Wang, D., Hubner, A., Chou, W. H., Davis, R. J., Burlingame, A. L., Messing, R. O., Katayama, C. D., Hedrick, S. M., and Shokat, K. M. (2007) A semisynthetic epitope for kinase substrates, *Nat. Methods 4*, 511–516.
38. Ubersax, J. A., Woodbury, E. L., Quang, P. N., Paraz, M., Blethrow, J. D., Shah, K., Shokat, K. M., and Morgan, D. O. (2003) Targets of the cyclin-dependent kinase Cdk1, *Nature 425*, 859–864.
39. Locker, J. (2001) *Transcription factors*, Academic Press, San Diego, CA.
40. Elowitz, M. B., and Leibler, S. (2000) A synthetic oscillatory network of transcriptional regulators, *Nature 403*, 335–338.
41. Becskei, A., and Serrano, L. (2000) Engineering stability in gene networks by autoregulation, *Nature 405*, 590–593.
42. Gardner, T. S., Cantor, C. R., and Collins, J. J. (2000) Construction of a genetic toggle switch in Escherichia coli, *Nature 403*, 339–342.
43. Fields, S., and Song, O. (1989) A novel genetic system to detect protein–protein interactions, *Nature 340*, 245–246.
44. Biggins, J. B., and Koh, J. T. (2007) Chemical biology of steroid and nuclear hormone receptors, *Curr. Opin. Chem. Biol. 11*, 99–110.
45. Chockalingam, K., and Zhao, H. (2005) Creating new specific ligand-receptor pairs for transgene regulation, *Trends Biotechnol. 23*, 333–335.
46. Doyle, D. F., Braasch, D. A., Jackson, L. K., Weiss, H. E., Boehm, M. F., Mangelsdorf, D. J., and Corey, D. R. (2001) Engineering orthogonal ligand-receptor pairs from "near drugs", *J. Am. Chem. Soc. 123*, 11367–11371.
47. Schwimmer, L. J., Rohatgi, P., Azizi, B., Seley, K. L., and Doyle, D. F. (2004) Creation and discovery of ligand-receptor pairs for transcriptional control with small molecules, *Proc. Natl. Acad. Sci. U.S.A. 101*, 14707–14712.
48. Chockalingam, K., Chen, Z., Katzenellenbogen, J. A., and Zhao, H. (2005) Directed evolution of specific receptor-ligand pairs for use in the creation of gene switches, *Proc. Natl. Acad. Sci. U.S.A. 102*, 5691–5696.
49. Sprinzak, D., and Elowitz, M. B. (2005) Reconstruction of genetic circuits, *Nature 438*, 443–448.
50. Drubin, D. A., Way, J. C., and Silver, P. A. (2007) Designing biological systems, *Genes Dev. 21*, 242–254.
51. Sayut, D. J., Niu, Y., and Sun, L. (2006) Construction and engineering of positive feedback loops, *ACS Chem. Biol. 1*, 692–696.
52. Nomura, W., and Sugiura, Y. (2007) Design and synthesis of artificial zinc finger proteins, *Methods Mol. Biol. 352*, 83–93.
53. Karig, D., and Weiss, R. (2005) Signal-amplifying genetic circuit enables *in vivo* observation of weak promoter activation in the Rhl quorum sensing system, *Biotechnol. Bioeng. 89*, 709–718.
54. Collins, C. H., Leadbetter, J. R., and Arnold, F. H. (2006) Dual selection enhances the signaling specificity of a variant of the quorum-sensing transcriptional activator LuxR, *Nat. Biotechnol. 24*, 708–712.
55. Collins, C. H., Arnold, F. H., and Leadbetter, J. R. (2005) Directed evolution of *Vibrio fischeri* LuxR for increased sensitivity to a broad spectrum of acyl-homoserine lactones, *Mol. Microbiol. 55*, 712–723.
56. Aharoni, A., Gaidukov, L., Khersonsky, O., Mc, Q. G. S., Roodveldt, C., and Tawfik, D. S. (2005) The 'evolvability' of promiscuous protein functions, *Nat. Genet. 37*, 73–76.
57. Isaacs, F. J., Dwyer, D. J., and Collins, J. J. (2006) RNA synthetic biology, *Nat. Biotechnol. 24*, 545–554.
58. Davidson, E. A., and Ellington, A. D. (2007) Synthetic RNA circuits, *Nat. Chem. Biol. 3*, 23–28.
59. Isaacs, F. J., Dwyer, D. J., Ding, C., Pervouchine, D. D., Cantor, C. R., and Collins, J. J. (2004) Engineered riboregulators enable post-transcriptional control of gene expression, *Nat. Biotechnol. 22*, 841–847.
60. Ellington, A. D. (2007) What's so great about RNA? *ACS Chem. Biol. 2*, 445–448.
61. Tuerk, C., and Gold, L. (1990) Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase, *Science 249*, 505–510.
62. Ellington, A. D., and Szostak, J. W. (1990) *In vitro* selection of RNA molecules that bind specific ligands, *Nature 346*, 818–822.
63. Bayer, T. S., and Smolke, C. D. (2005) Programmable ligand-controlled riboregulators of eukaryotic gene expression, *Nat. Biotechnol. 23*, 337–343.
64. Rackham, O., and Chin, J. W. (2006) Synthesizing cellular networks from evolved ribosome-mRNA pairs, *Biochem. Soc. Trans. 34*, 328–329.
65. Ramakrishnan, V. (2002) Ribosome structure and the mechanism of translation, *Cell 108*, 557–572.
66. Rackham, O., and Chin, J. W. (2005) A network of orthogonal ribosome·mRNA pairs, *Nat. Chem. Biol. 1*, 159–166.
67. Rackham, O., and Chin, J. W. (2005) Cellular logic with orthogonal ribosomes, *J. Am. Chem. Soc. 127*, 17584–17585.
68. Rackham, O., Wang, K., and Chin, J. W. (2006) Functional epitopes at the ribosome subunit interface, *Nat. Chem. Biol. 2*, 254–258.
69. Kimmerlin, T., and Seebach, D. (2005) '100 years of peptide synthesis': ligation methods for peptide and protein synthesis with applications to beta-peptide assemblies, *J. Pept. Res. 65*, 229–260.
70. Blaschke, U. K., Silberstein, J., and Muir, T. W. (2000) Protein engineering by expressed protein ligation, *Methods Enzymol. 328*, 478–496.
71. Tan, Z., Forster, A. C., Blacklow, S. C., and Cornish, V. W. (2004) Amino acid backbone specificity of the *Escherichia coli* translation machinery, *J. Am. Chem. Soc. 126*, 12752–12753.
72. Ambrogelly, A., Palioura, S., and Soll, D. (2007) Natural expansion of the genetic code, *Nat. Chem. Biol. 3*, 29–35.
73. Saks, M. E., Sampson, J. R., Nowak, M. W., Kearney, P. C., Du, F., Abelson, J. N., Lester, H. A., and Dougherty, D. A. (1996) An engineered *Tetrahymena* tRNAGln for *in vivo* incorporation of unnatural amino acids into proteins by nonsense suppression, *J. Biol. Chem. 271*, 23169–23175.
74. Hohsaka, T., and Sisido, M. (2002) Incorporation of non-natural amino acids into proteins, *Curr. Opin. Chem. Biol. 6*, 809–815.
75. Link, A. J., Mock, M. L., and Tirrell, D. A. (2003) Non-canonical amino acids in protein engineering, *Curr. Opin. Biotechnol. 14*, 603–609.
76. Wang, L., Brock, A., Herberich, B., and Schultz, P. G. (2001) Expanding the genetic code of *Escherichia coli*, *Science 292*, 498–500.
77. Chin, J. W., Cropp, T. A., Anderson, J. C., Mukherji, M., Zhang, Z., and Schultz, P. G. (2003) An expanded eukaryotic genetic code, *Science 301*, 964–967.
78. Kisselev, L. L., and Buckingham, R. H. (2000) Translational termination comes of age, *Trends Biochem. Sci. 25*, 561–566.
79. Wang, L., Xie, J., and Schultz, P. G. (2006) Expanding the genetic code, *Annu. Rev. Biophys. Biomol. Struct. 35*, 225–249.
80. Bose, M., Groff, D., Xie, J., Brustad, E., and Schultz, P. G. (2006) The incorporation of a photoisomerizable amino acid into proteins in *E. coli*, *J. Am. Chem. Soc. 128*, 388–389.
81. Wu, N., Deiters, A., Cropp, T. A., King, D., and Schultz, P. G. (2004) A genetically encoded photocaged amino acid, *J. Am. Chem. Soc. 126*, 14306–14307.
82. Deiters, A., Cropp, T. A., Mukherji, M., Chin, J. W., Anderson, J. C., and Schultz, P. G. (2003) Adding amino acids with novel reactivity to the genetic code of *Saccharomyces cerevisiae*, *J. Am. Chem. Soc. 125*, 11782–11783.
83. Tsao, M. L., Tian, F., and Schultz, P. G. (2005) Selective Staudinger modification of proteins containing p-azidophenylalanine, *ChemBioChem 6*, 2147–2149.

84. Deiters, A., and Schultz, P. G. (2005) *In vivo* incorporation of an alkyne into proteins in *Escherichia coli*, *Bioorg. Med. Chem. Lett. 15*, 1521–1524.
85. Agard, N. J., Baskin, J. M., Prescher, J. A., Lo, A., and Bertozzi, C. R. (2006) A comparative study of bioorthogonal reactions with azides, *ACS Chem. Biol. 1*, 644–648.
86. Prescher, J. A., and Bertozzi, C. R. (2005) Chemistry in living systems, *Nat. Chem. Biol. 1*, 13–21.
87. Wang, K., Neumann, H., Peak-Chew, S. Y., and Chin, J. W. (2007) Evolved orthogonal ribosomes enhance the efficiency of synthetic genetic code expansion, *Nat. Biotechnol. 25*, 770–777.
88. Bain, J. D., Switzer, C., Chamberlin, A. R., and Benner, S. A. (1992) Ribosome-mediated incorporation of a non-standard amino acid into a peptide through expansion of the genetic code, *Nature 356*, 537–539.
89. Hirao, I., Ohtsuki, T., Fujiwara, T., Mitsui, T., Yokogawa, T., Okuni, T., Nakayama, H., Takio, K., Yabuki, T., Kigawa, T., Kodama, K., Yokogawa, T., Nishikawa, K., and Yokoyama, S. (2002) An unnatural base pair for incorporating amino acid analogs into proteins, *Nat. Biotechnol. 20*, 177–182.
90. Matsuda, S., Fillo, J. D., Henry, A. A., Rai, P., Wilkens, S. J., Dwyer, T. J., Geierstanger, B. H., Wemmer, D. E., Schultz, P. G., Spraggon, G., and Romesberg, F. E. (2007) Efforts toward expansion of the genetic alphabet: structure and replication of unnatural base pairs, *J. Am. Chem. Soc. 129*, 10466–10473.
91. Ghadessy, F. J., Ramsay, N., Boudsocq, F., Loakes, D., Brown, A., Iwai, S., Vaisman, A., Woodgate, R., and Holliger, P. (2004) Generic expansion of the substrate spectrum of a DNA polymerase by directed evolution, *Nat. Biotechnol. 22*, 755–759.
92. Radhika, V., Proikas-Cezanne, T., Jayaraman, M., Onesime, D., Ha, J. H., and Dhanasekaran, D. N. (2007) Chemical sensing of DNT by engineered olfactory yeast strain, *Nat. Chem. Biol. 3*, 325–330.
93. Deiters, A., Cropp, T. A., Summerer, D., Mukherji, M., and Schultz, P. G. (2004) Site-specific PEGylation of proteins containing unnatural amino acids, *Bioorg. Med. Chem. Lett. 14*, 5743–5745.
94. Chang, M. C., Eachus, R. A., Trieu, W., Ro, D. K., and Keasling, J. D. (2007) Engineering *Escherichia coli* for production of functionalized terpenoids using plant P450s, *Nat. Chem. Biol. 3*, 274–277.
95. Baltz, R. H. (2006) Molecular engineering approaches to peptide, polyketide and other antibiotics, *Nat. Biotechnol. 24*, 1533–1540.
96. Lu, T. K., and Collins, J. J. (2007) Dispersing biofilms with engineered enzymatic bacteriophage, *Proc. Natl. Acad. Sci. U.S.A. 104*, 11197–11202.
97. Stephanopoulos, G. (2007) Challenges in engineering microbes for biofuels production, *Science 315*, 801–804.
98. Alper, H., Moxley, J., Nevoigt, E., Fink, G. R., and Stephanopoulos, G. (2006) Engineering yeast transcription machinery for improved ethanol tolerance and production, *Science 314*, 1565–1568.
99. Holland, P. W., and Garcia-Fernandez, J. (1996) Hox genes and chordate evolution, *Dev. Biol. 173*, 382–395.