Please cite this article as: Glasscock et al., Engineered Protein Machines: Emergent Tools for Synthetic Biology, Cell Chemical Biology (2016), http://dx.doi.org/10.1016/j.chembiol.2015.12.004

Cell Chemical Biology
Review

Engineered Protein Machines: Emergent Tools for Synthetic Biology

Cameron J. Glasscock,[1] Julius B. Lucks,[1] and Matthew P. DeLisa[1,*]
[1]School of Chemical and Biomolecular Engineering, Cornell University, Ithaca, NY 14853, USA
*Correspondence: md255@cornell.edu
http://dx.doi.org/10.1016/j.chembiol.2015.12.004

Nature has evolved an array of intricate protein assemblies that work together to perform the chemistry that maintains life. These protein machines function with exquisite specificity and coordination to accomplish their tasks, from DNA and RNA synthesis to protein folding and post-translational modifications. Despite their complexity, synthetic biologists have succeeded in redesigning many aspects of these molecular machines. For example, natural DNA polymerases have now been engineered to catalyze the synthesis of alternative genetic polymers called XNAs, orthogonal RNA polymerases and ribosomes have been engineered to enable the construction of genetic logic gates, and protein biogenesis machinery such as chaperonins and protein translocons have been repurposed to improve folding and expression of recombinant proteins. In this Review, we highlight the progress made in understanding, engineering, and repurposing bacterial protein machines for use in synthetic biology and biotechnology.

Introduction

Many fundamental biological processes are performed by complexes of proteins that work together in concert. These molecular machines often consist of 10 or more proteins, each with unique purpose, that must associate with spatial and temporal precision for biological activity, such as template-directed DNA, RNA, and protein synthesis (Alberts, 1998). A particularly familiar example is the bacterial ribosome, whereby 52 proteins and 3 rRNA molecules assemble into a ribonucleoprotein machine that orchestrates mRNA translation into nascent polypeptides (Dunkle et al., 2011; Fischer et al., 2015). Engineering such machines with new or enhanced functions is a daunting task but one that promises to catalyze forward progress in synthetic biology and biotechnology as a whole. For example, DNA/RNA polymerases and ribosomes have been reconfigured to synthesize alternative genetic polymers (Pinheiro and Holliger, 2012) and more efficiently incorporate unnatural amino acids (UAAs) into proteins (Lang and Chin, 2014). These synthetic biology feats promise to enable, and in many cases have enabled, the creation of genetically tailored nanostructures, cotranslational incorporation of bio-orthogonal probes of protein function, and the synthesis of new proteins with chemistries that are unattainable with the canonical set of amino acids.

Going beyond the manipulation of natural protein machines, this progress has invigorated the vision of designer cells made from engineered biomolecular components, or ''parts,'' which have well-defined functions that can be combined together with predictable composite function (Lucks et al., 2008). Such biomolecular systems would find utility in diverse applications, ranging from sustainable chemical production (Keasling, 2010) to environmental biosensing (Bereza-Malcolm et al., 2015). However, a key technical challenge to this vision is to create genetic regulatory layers that can reliably control the expression of the multi-component engineered protein machines so they form and function correctly (Lucks et al., 2008). Here again there has been substantial progress in creating libraries of robust and tunable RNA- and protein-based regulators (Chappell et al., 2015; Stanton et al., 2014) that, together with sets of well-characterized orthogonal protein machines such as RNA polymerases and ribosomes, lay the foundation for creating designer cellular systems.

To ultimately realize the full functionality of synthetic cellular systems, one of the practical challenges that many are currently tackling is overcoming the difficulty in efficiently expressing functional heterologous proteins as constituents of engineered synthesis machinery, metabolic pathways, signaling cascades, or other multi-protein systems. Fortunately, engineered machines themselves, such as those involved in protein folding and export, may provide a solution to the expression problem by supplying folding-optimized proteins for use in these and other related applications (DeLisa et al., 2003; Wang et al., 2002).

Overall, the continued engineering and repurposing of cellular machinery is likely to be of great value for a broad array of biotechnologies, especially in light of the ubiquity of protein machines in biology. However, given the sophistication of these systems, these efforts will almost certainly require methodologies not yet developed and contributions from chemistry, physics, and engineering, in addition to those from biology. Fortunately, synthetic biologists span these scientific disciplines and have succeeded in manipulating many of these systems for new biotechnological functions. In this review, we highlight the progress that has been made in understanding, engineering, and repurposing bacterial protein machines for use in synthetic biology and applications in biotechnology.

Engineered DNA Replication Machines for Novel Nucleic Acid Polymer Technologies

DNA polymerases are the molecular machines that govern genetic replication. They are at the core of any technologies that manipulate and evolve DNA sequences. As a result, DNA polymerases are perhaps among the most-engineered protein machines we know of that are routinely used in laboratories worldwide. Our currently available DNA polymerase tool box (Collins et al., 2003) includes enzymes that lack 3′-5′ exonuclease

Please cite this article as: Glasscock et al., Engineered Protein Machines: Emergent Tools for Synthetic Biology, Cell Chemical Biology (2016), http://dx.doi.org/10.1016/j.chembiol.2015.12.004

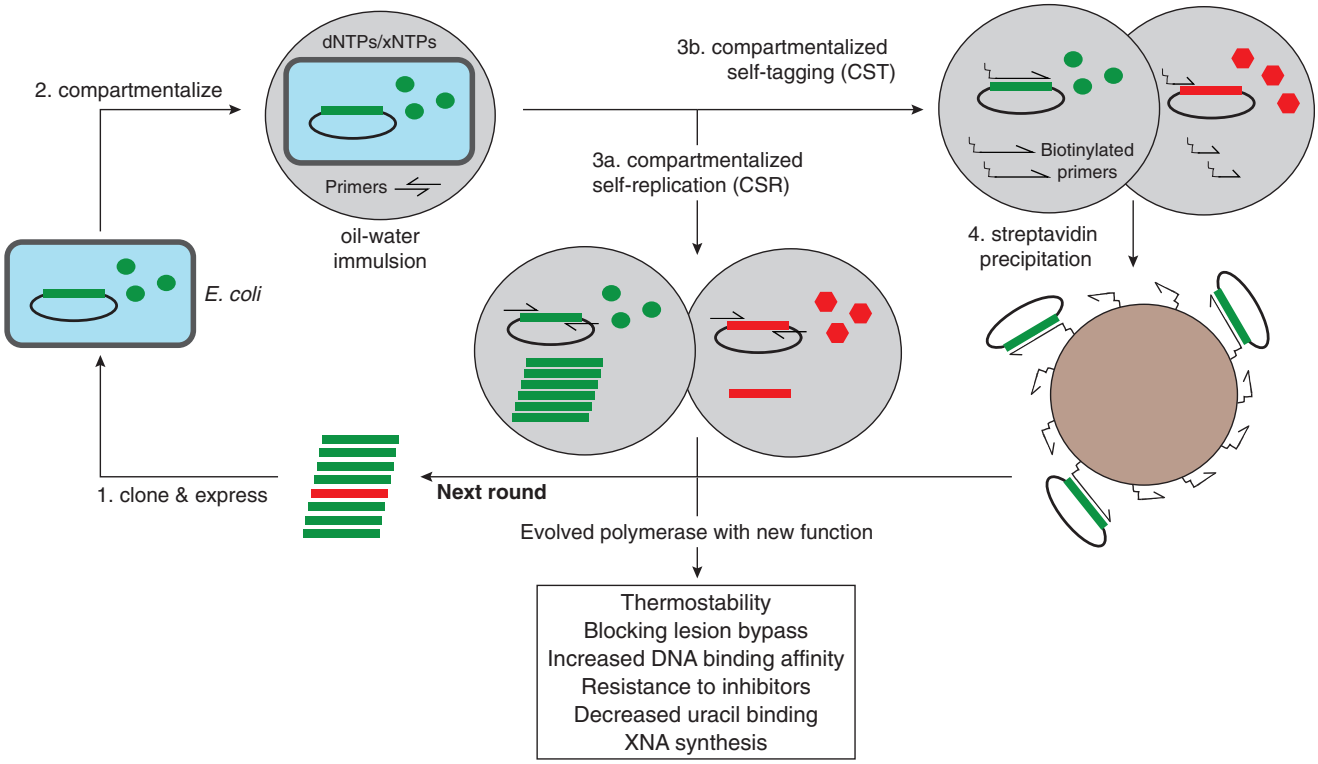

Figure 1. Schemes for Evolving DNA Polymerases with New Functionality

(1) DNA polymerase libraries are cloned and expressed in *E. coli*. (2) Cells expressing cloned polymerase genes are compartmentalized into oil-water emulsions containing dNTPs/xNTPs and primers. (3) Either (a) compartmentalized self-replication (CSR) is performed under selective conditions to preferentially amplify the genes of productive DNA polymerases (green circles) rather than unproductive DNA polymerases (red polygons) (Ghadessy et al., 2001) or (b) compartmentalized self-tagging (CST) is performed to tag productive polymerase genes with a biotin-labeled primer before (4) streptavidin precipitation (Pinheiro et al., 2012). These schemes have been used to generate new polymerases with increased thermostability, the ability to bypass blocking lesions, increased DNA binding affinity, resistance to inhibitors, decreased uracil binding, and, ultimately, synthesis of unnatural genetic polymers (Arezi et al., 2014; Ghadessy et al., 2001, 2004; Pinheiro et al., 2012; Tubeleviciute and Skirgaila, 2010). Green bars represent genes for productive polymerase variants; red bars represent genes for unproductive polymerase variants.

activity for use in error-prone PCR (Tindall and Kunkel, 1988), those that are fused with processivity-enhancing domains to improve speed and fidelity (Wang et al., 2004), and enzymes that possess broad substrate specificity to allow incorporation of modified bases, such as fluorescently labeled dNTPs (Glick et al., 2002). One technology that has greatly facilitated efforts in engineering DNA polymerases is compartmentalized self-replication (CSR) (Ghadessy et al., 2001), which provides a generalizable selection strategy for evolving new polymerase functionality (Figure 1). This technology, invented by Holliger and coworkers, takes advantage of the inherent ability of a DNA polymerase to amplify its own DNA sequence, thereby ensuring a linkage between phenotype and genotype. Specifically, combinatorial libraries of a DNA polymerase are overexpressed in *Escherichia coli* cells, and the resulting cell libraries are transferred into oil-water emulsions. Following heat treatment to degrade the components of the cell, only the polymerase and its template gene remain in the emulsion, and a PCR reaction can be initiated by polymerase variants that are functional under conditions that can be customized to select for numerous desired phenotypes. The CSR scheme has been used to generate novel DNA polymerases with a number of unique features, including increased thermostability, the ability to bypass blocking lesions, increased DNA binding affinity, resistance to inhibitors, and decreased uracil binding (Arezi et al., 2014; Ghadessy et al., 2001, 2004; Pinheiro et al., 2012; Tubeleviciute and Skirgaila, 2010).

Building off of CSR, synthetic biologists have made further, foundational, progress in polymerase engineering. Pinheiro et al. used a next-generation adaptation of CSR called compartmentalized self-tagging (CST) (Figure 1) to engineer polymerases that catalyze the synthesis of various synthetic genetic polymers made of xeno nucleic acids (XNAs) from a DNA template (Pinheiro et al., 2012). This was an important demonstration, because XNA-based genetic polymers could enable fine-tuning of chemical properties, such as helical confirmation, duplex stability, and altered base-pairing preferences (e.g., G·U wobble pairs). Such polymer chemistries could be optimized for many applications, ranging from next-generation data storage to nucleic acid-based therapeutics (Pinheiro and Holliger, 2012).

The CST selection strategy functions by compartmentalizing libraries of polymerase variants in oil-water emulsions, as in CSR, but uses a biotinylated XNA primer to initiate the PCR reaction. Primer extension by an active polymerase variant stabilizes binding to the plasmid template and allows isolation of active

Please cite this article as: Glasscock et al., Engineered Protein Machines: Emergent Tools for Synthetic Biology, Cell Chemical Biology (2016), http://dx.doi.org/10.1016/j.chembiol.2015.12.004

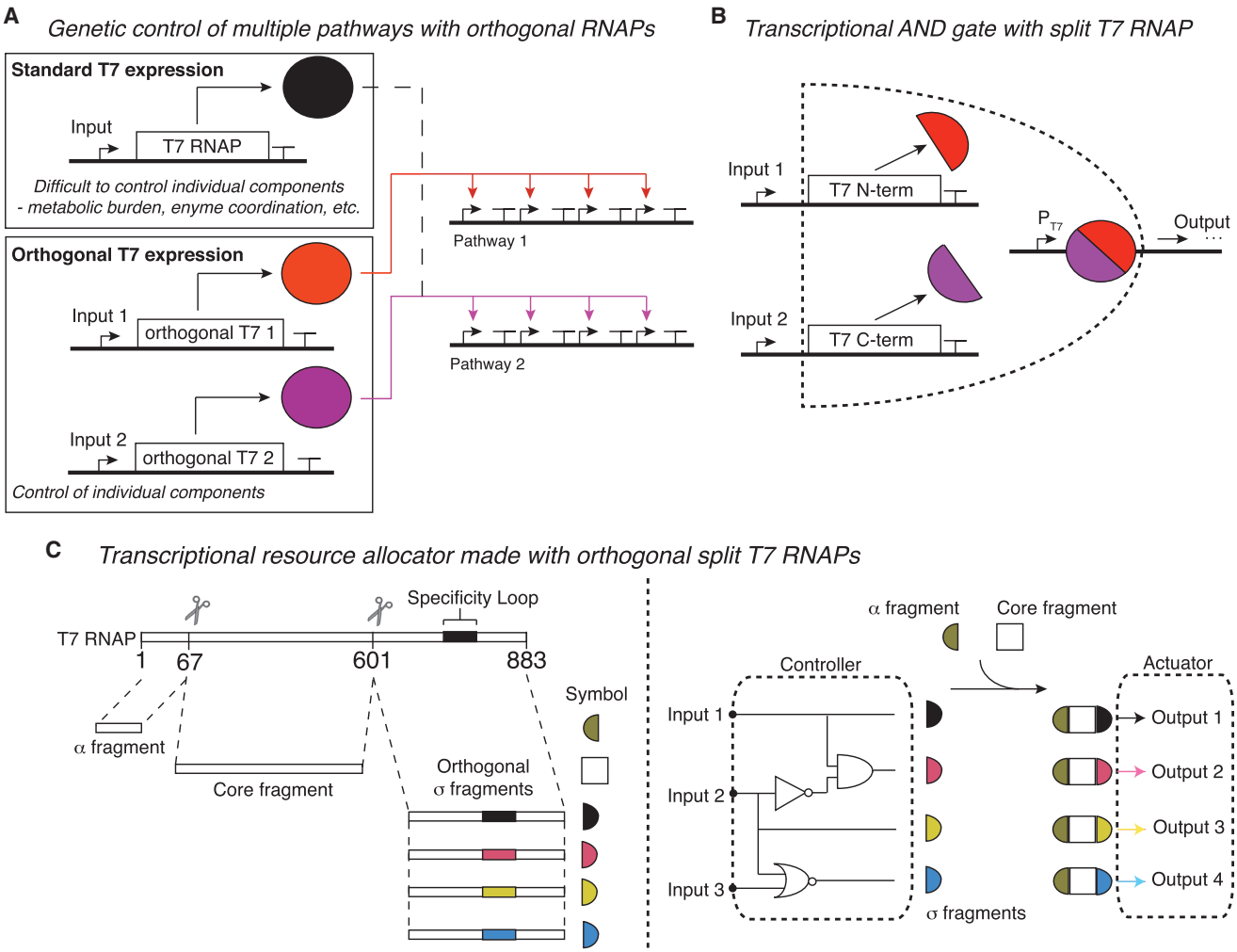

Figure 2. Controlling Gene Expression with Orthogonal and Split T7 RNAPs

(A) Orthogonal T7 RNAPs enable genetic control of multiple by-pathways and can be used to alleviate burden in cell factories and coordinate the expression of multiple genetic components (Temme et al., 2012).

(B) Transcriptional AND gate that produces output only when both N- and C-terminal split T7 components are induced by their respective inputs (Shis and Bennett, 2013).

(C) Split T7 RNAP-based resource allocator that controls output by the expression of orthogonal σ fragments in response to different inputs (Segall-Shapiro et al., 2014). An additional α fragment can also be used to modulate total output in response to the presence of other proteins in the system. Colored circles and lines in (A), (B), and (C) denote orthogonal RNAP variants.

polymerase genes by streptavidin precipitation. Successive rounds of CST were used to evolve and isolate variants of the TgoT DNA polymerase from *Thermococcus gorgonarius* that synthesized six different genetic polymers that formed base pairs with DNA but were chemically distinct. These included cyclohexenyl nucleic acids (CeNAs), arabinonucleic acids (ANAs), 2'-fluoro-arabinonucleic acids (FANAs), α-L-threofuranosyl nucleic acids (TNAs), 1,5-anhydrohexitol nucleic acids (HNAs), and 2′-*O*,4′-*C*-methylene-β-*D*-ribonucleic acids (locked nucleic acids [LNAs]). The ability to enzymatically catalyze the template-driven synthesis of LNA-based genetic polymers is of particular biotechnological importance because the ribofuranose ring is "locked" in a single conformation by a methylene bridge between the 2′-O and C4, contributing to considerable duplex stabilization over DNA (Singh and Koshkin, 1998). Further optimization of these engineered XNA polymerases could expand the toolbox for synthesis and chemical tuning of sequence-controlled synthetic polymers for a variety of applications in medicine, agriculture, and the environment (Lutz et al., 2013).

Engineered RNA Synthesis Machinery for Coordinating Multi-component Genetic Systems

RNA synthesis is the first step in gene expression, and controlling this process is essential in the aim to reliably manipulate cell behavior and implement genetic circuits. Specifically, in order to coordinate the activity of multiple heterologous components within a cell, it is necessary to develop tools for orthogonal control over gene expression so that components can be controlled individually with minimal cross-talk (Figure 2A). In an effort to develop these new tools, synthetic biologists have

Please cite this article as: Glasscock et al., Engineered Protein Machines: Emergent Tools for Synthetic Biology, Cell Chemical Biology (2016), http://dx.doi.org/10.1016/j.chembiol.2015.12.004

begun to engineer RNA synthesis machinery using both rational design and laboratory evolution.

The phage T7 RNA polymerase (RNAP) has long been a standard tool for protein expression that enables tight regulation of RNA synthesis and minimal cross-talk with host gene expression. However, standard T7 RNAP-based expression systems suffer from toxicity to the host and are limited in the number of orthogonal variants that can be used simultaneously within a cell. In order to address these issues, Temme et al. used rational techniques to design less-toxic phage T7 RNAP variants, which then served as scaffolds to create a panel of four highly orthogonal T7 RNAPs (Temme et al., 2012). The authors first determined mutations that altered expression strength and activity to reduce the toxicity associated with expression of the T7 RNAP in *E. coli*. Next, the authors used a bioinformatics strategy called part mining (Bayer et al., 2009) to identify homologous phage RNAPs. In part-mining strategies, protein homologs are found from bioinformatics searches of sequenced genomes. The DNA sequences of these homologs are then chemically synthesized so that they can be used in cellular systems. In this case, the authors found variations in the β hairpin of each identified polymerase homolog. These variations were exchanged into the mutated T7 RNAP backbone, yielding a panel of 43 synthetic T7 RNAPs. From this panel, four highly active and orthogonal variants were isolated after screening for function. To demonstrate the potential of these orthogonal transcription systems, the authors constructed a synthetic AND gate and used it to toggle between two metabolic pathways, depending on the inducers present. Systems such as this could be used to reduce burden in complex multi-enzyme pathways by streamlining our ability to individually express different components at different times and in response to different inducers.

In an effort to further simplify T7-based transcriptional control, Shis et al. showed that T7 RNAP can be split into separate subdomains while maintaining functionality in vivo when both domains are expressed together (Shis and Bennett, 2013). Interestingly, split T7 RNAP was stably expressed and did not cause cellular toxicity upon expression. The authors incorporated known mutations (Chelliserrykattil et al., 2001; Raskin et al., 1992, 1993) into the specificity loop of the split T7 RNAP to create a set of three highly orthogonal and active split T7 RNAP variants. These split T7 RNAPs were used to build two-input transcriptional logic gates, where the N- and C-terminal domains were expressed separately under different inducible promoters. Expression of active T7 RNAP only occurred under the addition of both inducers, whereas addition of individual inducers resulted in expression of a single inactive fragment (Figure 2B). In subsequent work, Schaerli et al. used a split-intein approach to construct T7 RNAP-based transcriptional AND logic gates that had greater activity than the split T7 RNAP-based approach (Schaerli et al., 2014).

Orthogonal split T7 RNAPs have recently been combined to create a transcription device that acts as a resource allocator for budgeting the limited transcriptional machinery provided by T7 RNAP (Segall-Shapiro et al., 2014) (Figure 2C). In this system, analogous to the regulation provided by sigma factors in *E. coli*, Segal-Shapiro et al. used transposon mutagenesis and a screen for transcription activity to identify two locations within T7 RNAP where the polymerase could be divided and yet remain functional. These locations, amino acids 67 and 601, were used as a guide to fragment T7 RNAP into a "core fragment," a DNA-binding fragment σ, and an α fragment. The core fragment was used to set the total level of polymerase within the cell, and orthogonal σ fragments, which were specific to different promoter sequences, competed for binding with the core fragment. By modulating the expression of individual σ fragments, different genes were assigned varying amounts of total T7 transcription capacity. Additionally, since the split polymerase was non-functional without the α fragment being present, the authors were able to make T7 transcription activity responsive to the level of α fragment produced in the cell. By further introducing an inactive (or null) α fragment that competes with active α fragment binding, this feature was used as a regulator that maintained constant gene-expression levels independent of plasmid backbone copy number. This striking feature could solve a major challenge in synthetic biology, where it is difficult to move genetic constructs from one genetic context (e.g., high-copy plasmid) to another (e.g., low-copy plasmid or chromosome) without the laborious effort of rebalancing expression levels. One could envision using such a system to screen and optimize individual components of a genetic system with the ease of plasmid-based expression and subsequent chromosomal integration of the final construct for stable, plasmid-free expression.

T7 RNAP-based technologies for transcriptional logic, resource allocation, and expression-level control have proven to be powerful tools for synthetic biology; however, additional orthogonal RNAP variants are required to implement large-scale genetic systems with many components. Fortunately, new laboratory evolution technologies have provided convenient methods for engineering these additional variants. For example, Ellefson et al. described another adaptation of CSR, termed compartmentalized partner replication (CPR) (Figure 3A), that enabled laboratory evolution of a phage T7-based RNAP variant capable of recognizing an orthogonal promoter sequence (Ellefson et al., 2013). CPR involves placing expression of Taq DNA polymerase under control of an orthogonal promoter sequence. In this way, mutants of T7 RNAP that efficiently transcribe the orthogonal promoter sequence will produce a greater amount of Taq DNA polymerase within the cell prior to a subsequent in vitro oil-water emulsion step. At this point, a PCR program is initiated so that Taq DNA polymerase will preferentially amplify the most efficient RNAP variant, or "partner gene." Building on this work, Meyer et al. created a library of six independently orthogonal T7 RNAP•promoter pairs that should further enable synthetic biologists to design and build large-scale genetic constructs (Meyer et al., 2014).

An alternative selection system for generating engineered RNAPs is phage-assisted continuous evolution (PACE) (Esvelt et al., 2011) (Figure 3B). PACE leverages filamentous phage and its dependence on the gene encoding minor coat protein pIII to create a link between RNAP activity and phage infection. As a result of this dependence, only those *E. coli* cells carrying a functional orthogonal RNAP will allow efficient phage infection. By combining a phage library containing RNAP variants, a mutation-prone plasmid, and an accessory plasmid encoding pIII, PACE allows continuous evolution of any biomolecule that can be coupled to production of pIII. Esvelt et al. used PACE to evolve a T7 RNAP that initiated transcription from the T3 RNAP

Please cite this article as: Glasscock et al., Engineered Protein Machines: Emergent Tools for Synthetic Biology, Cell Chemical Biology (2016), http://dx.doi.org/10.1016/j.chembiol.2015.12.004

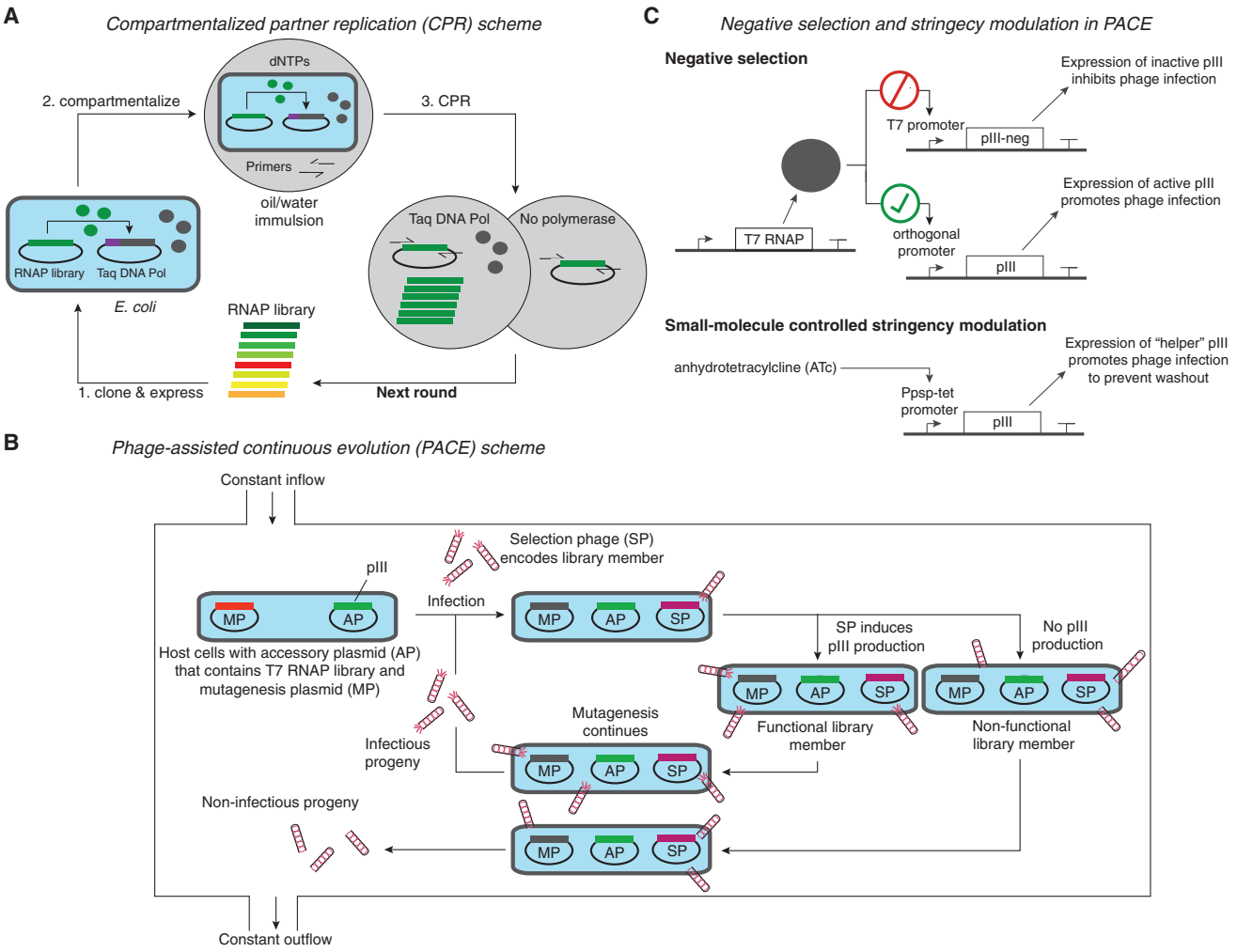

Figure 3. Schemes for Evolving RNA Polymerases with New Functionality

(A) Compartmentalized partner replication (CPR), a derivative of CSR, couples production of Taq DNA polymerase with RNAP activity. (1) A library of RNAPs (green circles) is cloned and expressed in *E. coli*. Taq DNA polymerase production is dependent on RNAP activity (Ellefson et al., 2013). (2) Cells are compartmentalized into oil-water emulsions containing dNTPs and primers. (3) Compartments containing Taq DNA polymerase (grey circles) preferentially amplify functional RNAP genes for subsequent rounds of evolution. Colored bars represent genes encoding RNAP variants, and purple/gray bars represent a promoter sequence (purple) and the gene for Taq DNA polymerase (grey).

(B) Phage-assisted continuous evolution (PACE) couples RNAP activity with phage infection (Esvelt et al., 2011). Production of protein pIII from an accessory plasmid is made dependent on RNAP functionality so that only phages containing functional RNAPs are capable of infection during continuous mutagenesis.

(C) Negative selection can be incorporated into PACE by placing a non-functional pIII mutant (pIII-neg) under control of an undesired promoter sequence. Furthermore, selection stringency can be modulated by including an anhydrotetracycline (ATc)-controlled component that expresses "helper" pIII to prevent washout of early mutations (Carlson et al., 2014).

promoter (although it maintained specificity for the native T7 RNAP promoter), as well as synthesis of transcripts beginning with ATP and CTP, rather than GTP, which is the native initiation nucleotide of a T7 RNAP synthesized transcript. Carlson et al. expanded the PACE method by incorporating a small-molecule controlled component to modulate selection stringency by controlling the levels of "helper" pIII that prevents washout of early mutations (Carlson et al., 2014) (Figure 3C). Furthermore, Carlson et al. introduced a negative-selection step for simultaneously excluding variants with reactivity toward undesired substrates. This step functions by placing a non-functional pIII mutant (pIII-neg) under control of the undesired promoter (e.g., the wild-type T7 promoter). By combining these advances, the authors successfully evolved T7 RNAP enzymes with significantly altered, rather than just broadened, specificity toward the T3 RNAP promoter. Overall, the use of strategies like CPR and PACE have enabled the creation of a larger repertoire of orthogonal T7 RNAP variants for synthetic biology and provide potentials paths toward evolution of other biomolecules that are important to the field, such as binding proteins, proteases, and optimized genome engineering enzymes.

Engineered Orthogonal Ribosomes for Genetic Regulation and Unnatural Protein Synthesis

In addition to engineered polymerases for customized control over DNA replication, as well as XNA and RNA synthesis, recent

Please cite this article as: Glasscock et al., Engineered Protein Machines: Emergent Tools for Synthetic Biology, Cell Chemical Biology (2016), http://dx.doi.org/10.1016/j.chembiol.2015.12.004

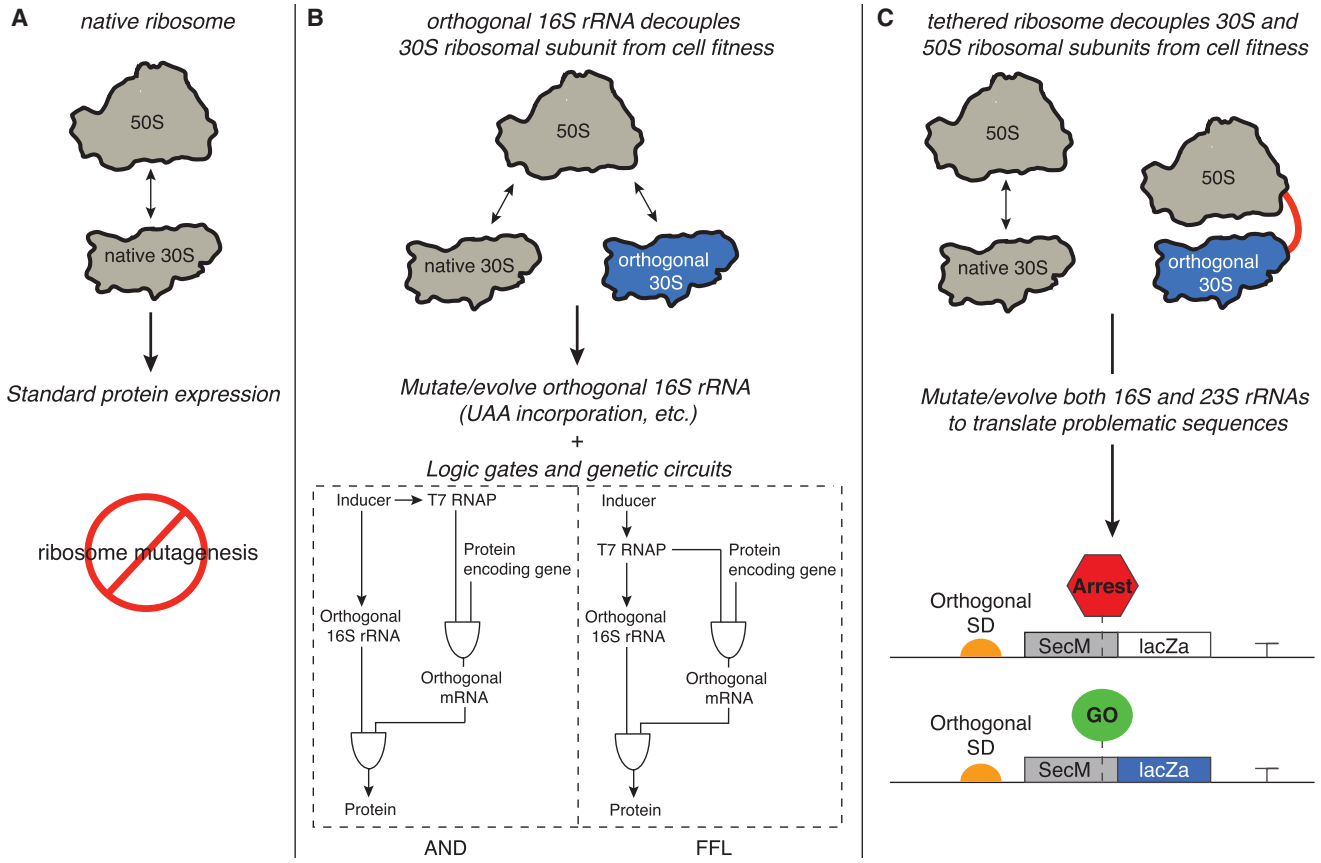

Figure 4. Engineering Orthogonal Ribosomes

(A) Engineering or evolving the native ribosome for new function is difficult because the cell depends on natural translation activity to maintain cell viability.

(B) Engineered orthogonal 16S rRNAs allow decoupling of the 30S ribosomal subunit from cell fitness. This enables evolution of the 16S rRNA for new function as well as new modes of synthetic regulation, such as the transcription-translation AND gate and feedforward loop (FFL), that direct transcription of an orthogonal mRNA transcript (An and Chin, 2009).

(C) Tethering of an orthogonal 16S rRNA to the 23S rRNA enables evolution of the entire ribosome to translate problematic amino acid sequences, such as SecM. In this experiment, mutagenesis of an entire orthogonal tethered ribosome was performed to select functional variants that translate through the SecM leader peptide and direct synthesis of β-galactosidase encoded by *lacZα* (Orelle et al., 2015).

efforts in the field have focused on the ribonucleoprotein ribosomal machinery underlying protein synthesis. Such ribosome engineering efforts open the door to new modes of synthetic regulation and enable further engineering of unique capabilities, such as efficient incorporation of unnatural amino acids into proteins. However, engineering ribosomes is difficult because they are essential for the cell to perform natural gene expression and maintain viability (Figure 4A). Hence, most attempts at ribosome engineering cannot be performed without the availability of orthogonal variants that can be supplied in the cell simultaneously with the host ribosome.

The bacterial ribosome, a 2.5-MDa complex of protein and rRNA, translates mRNA sequences into protein. In bacteria, this process is initiated, in part, through an interaction between the 16S small subunit of the ribosome and the ribosome-binding site (RBS) of the mRNA transcript (Liljas and Ehrenberg, 2011). This interaction occurs through mRNA-rRNA base pairing between the mRNA Shine-Dalgarno (SD) sequence and the anti-Shine-Dalgarno (ASD) sequence of the 16S rRNA (Steitz and Jakes, 1975). An early example of partially orthogonal ribosome activity showed that three mutations in the SD sequence of an mRNA transcript resulted in very low translation of human growth hormone (hGH) by the endogenous ribosome (Hui and de Boer, 1987). However, when compensatory mutations were made in an exogenous 16S rRNA, translation of hGH was recovered. This system allowed specialized translation from a mutated SD sequence; however, significant translation of endogenous protein was detected, and expression of the specialized ribosome resulted in toxicity and cell death.

A major step toward a fully orthogonal ribosome came when Rackham et al. used a dual-selection scheme to generate 16S rRNA•mRNA pairs that were orthogonal to each other and to the host (Rackham and Chin, 2005a). In this experiment, a negative selection was first performed to remove mRNA sequences that were substrates for the endogenous ribosome. This was accomplished by growing cells harboring a library of RBS variants in the presence of 5-fluorouracil (5-FU). In this negative-selection scheme, active RBS variants directed the synthesis of a chloramphenicol acetyltransferase (CAT)-uracil phosphoribosyltransferase (UPRT) fusion protein. Because UPRT converts 5-FU into a toxic product that results in cell death, RBS sequences that were a substrate for the endogenous ribosome were

Please cite this article as: Glasscock et al., Engineered Protein Machines: Emergent Tools for Synthetic Biology, Cell Chemical Biology (2016), http://dx.doi.org/10.1016/j.chembiol.2015.12.004

removed. In contrast, mRNAs with RBS sequences that were not translated by the endogenous ribosome survived the selection and comprised an enriched library of orthogonal mRNAs. Cells containing the enriched mRNAs were transformed with a second library of mutant 16S rRNAs. These cells were grown in the presence of chloramphenicol so that survival only occurred when orthogonal mRNA was efficiently translated into the CAT-UPRT fusion. Ten orthogonal 16S rRNA•mRNA pairs were identified in this work.

An immediate use of orthogonal 16S rRNAs is the orthogonal translation of different mRNA pairs (Figure 4B), similar in spirit to the orthogonal transcription of different promoters offered by the engineered T7 RNAP systems discussed above. To demonstrate this, Rackham et al. constructed a simple logic gate whose output was controlled by an orthogonal ribosome (Rackham and Chin, 2005a). For example, by combining an orthogonal ribosome, a gene encoding the α fragment of β-galactosidase (β-gal) on the corresponding orthogonal SD sequence, and a gene encoding the ω fragment of β-gal on a wild-type SD sequence, an AND gate was created in which the assembly of functional β-gal was dependent on the presence of each of the three components. Likewise, a Boolean OR gate was constructed whereby the ω fragment of β-gal was constitutively produced from a wild-type RBS while the α fragment was synthesized by either of two orthogonal ribosome•mRNA pairs (Rackham and Chin, 2005b), resulting in the production of functional β-gal.

Another advantage of these engineered translational systems is that they can be further combined with engineered transcriptional systems for more sophisticated control of gene expression. Evolved orthogonal ribosomes can be combined with T7 RNAP to compose dynamic transcription-translation regulatory networks such as the feedforward loop (FFL) (Figure 4B). An and Chin constructed AND gate behavior by placing an orthogonal mRNA encoding GFP under control of a T7 RNAP (An and Chin, 2009). This resulted in GFP synthesis that was dependent on both the T7 RNAP and the orthogonal ribosome. The authors went on to place synthesis of the orthogonal 16S rRNA under control of the T7 RNAP, such that addition of T7 RNAP allowed synthesis of both the orthogonal mRNA transcript and the orthogonal 16S rRNA. Subsequent synthesis and processing of the 16S rRNA then enabled translation of GFP, creating a functional FFL that could not be constructed with only the cell's endogenous machinery. This dynamic regulation allowed tunable time delay between T7 RNAP induction and GFP synthesis by altering the expression architecture of the orthogonal ribosome. This type of regulation could become useful in genetic systems where two components need to be activated in response to single input, but in a temporally staged fashion.

Engineered protein machines are also useful for studying fundamental biological phenomena. Several studies have utilized the development of orthogonal 16S rRNAs to probe mechanisms in translation. Weissman and coworkers used an orthogonal ribosome to evaluate the impact of SD-like sequences within a coding frame on translational pausing (Li et al., 2012). In other work, Rackham et al., utilized orthogonal 16S rRNA•mRNA pairs to functionally assess the importance of individual nucleotides in the 16S rRNA (Rackham et al., 2006). A key feature of their system that allowed this work was the orthogonality of the 16s rRNAs in the first place. By using orthogonal 16S rRNAs, mutants could be decoupled from cell fitness, which would not be possible if the wild-type ribosome was mutated. Functional mutations were selected by growth in the presence of chloramphenicol when the *cat* gene, coding for chloramphenicol acetyltransferase that allows bacteria to grow in the presence of this antibiotic, was encoded by the orthogonal mRNA. This study defined a subset of mutations that may have measurable effects on individual steps in the translation cycle.

In addition to creating whole new methods to controlling gene expression, orthogonal ribosomes that can be decoupled from cell fitness for further engineering have enabled new technologies for efficient incorporation of UAAs into proteins. The natural *E. coli* ribosome can often be ~20%–30% efficient at incorporating UAAs when supplied with an evolved aminoacyl-tRNA synthetase (Ryu and Schultz, 2006). However, because the ribosome is coupled to cell viability, it is difficult to screen combinatorial libraries of ribosomal components (e.g., 16S rRNA) with the goal of improving UAA incorporation. To overcome this challenge, Wang et al. developed a selection strategy to evolve orthogonal ribosomes for efficient decoding of the amber stop codon (Wang et al., 2007). Specifically, an orthogonal 16S rRNA•mRNA pair that encoded a mutated *cat* gene containing an amber stop codon was used to transform cells carrying a combinatorial library of 16S rRNAs and the above-mentioned machinery. By supplying an evolved orthogonal aminoacyl-tRNA synthetase for the amber stop codon, functional 16S rRNA variants capable of translating the mutated *cat* sequence were isolated based on their ability to confer growth in the presence of chloramphenicol. These 16S rRNAs not only decoded the amber stop codon but also efficiently incorporated the photocrosslinking amino acid *p*-benzoyl-L-phenylalanine (Bpa) into proteins.

Along similar lines, Neumann et al. evolved orthogonal ribosomes that efficiently decoded quadruplet codons (Neumann et al., 2010). These were combined with orthogonal aminoacyl-tRNA synthetase pairs to efficiently and site-specifically incorporate multiple UAAs into a single protein. The ability to efficiently decode amber stop codons and quadruplet codons has resulted in several technologies for interrogating protein structure and function. For instance, Sachdeva et al. and Wang et al. used these features for site-specific dual labeling of proteins with FRET probes, which in turn enabled visualization of protein conformational changes in response to ligand binding (Sachdeva et al., 2014; Wang et al., 2014).

Although evolving the 16S rRNA subunit has resulted in partially orthogonal translation systems, true insulation from the host requires independence from the endogenous 50S subunit. It has been assumed that functional translation requires independent subunit synthesis and detachment between the 30S and 50S subunits, making it difficult to apply mutagenesis approaches to the entire ribosome. Recently, however, two reports showed that the 30S and 50S subunits can be physically tethered by RNA linkers to create a truly orthogonal ribosome (Fried et al., 2015; Orelle et al., 2015) (Figure 4C). In Orelle et al., this was accomplished by using circular permutation to identify residues in the 23S rRNA that served as new 5′ and 3′ termini. Three circular permutations were identified that resulted in functional

Please cite this article as: Glasscock et al., Engineered Protein Machines: Emergent Tools for Synthetic Biology, Cell Chemical Biology (2016), http://dx.doi.org/10.1016/j.chembiol.2015.12.004

ribosome assembly, one of which had termini that were in close proximity to an apex loop of the 16S rRNA. Using this permutation as a building block for tethering of the 16S and 23S rRNAs with two RNA linkers, a functional tethered ribosome was produced that could maintain cell viability in the absence of endogenous ribosomes. Similar strategies were also used in Fried et al. to construct a ''single-subunit'' ribosome. As proof of utility, Orelle et al. mutated the ASD sequence of the 16S rRNA and its cognate mRNA SD sequence to create a truly orthogonal ribosome•mRNA pair. Directed evolution was subsequently applied to create a tethered ribosome capable of translating the SecM leader peptide, a sequence that stalls the endogenous ribosome (Figure 4C). Since this involved mutating regions of the 23S rRNA that are important for cell survival, these results would have been difficult or impossible to achieve in other cell-based systems with native ribosomes.

Engineered Chaperonins for Improved Protein Folding

Although one way to achieve efficient translation of difficult heterologous proteins, such as those with numerous UAA modifications, is by engineering orthogonal ribosomes, translation is not the end of protein biosynthesis—protein structure and function requires proper folding, a process that begins co-translationally during synthesis (Netzer and Hartl, 1997). Therefore, engineering protein machinery that influences folding is of technical importance for expression and stability of recombinant proteins. Folding of many proteins can be heavily influenced by the action of protein machines called chaperonins that act during and after translation (Ying et al., 2005). For example, the GroEL/ES system of *E. coli* is a group I chaperonin complex that interacts with ~250 proteins in wild-type cells, of which ~85 exhibit an obligate dependence on GroEL for folding under normal growth conditions (Kerner et al., 2005). GroEL/ES is composed of the chaperone GroEL, a tetradecamer of 57-kDa subunits arranged as two stacked, seven-membered rings that form a hollow tubular structure, and co-chaperone GroES, a smaller ring structure that binds to the ends of GroEL (Saibil, 2013). GroEL functions to capture substrate proteins within its central cavity as they exit from the ribosome. GroES then binds to both ends of GroEL and encapsulate the substrate, inducing a major conformational change that causes the central cavity of GroEL to double in size. This creates a polar environment in which the substrate folds in isolation from other cytoplasmic proteins before being released (Saibil, 2013).

While GroEL/ES is capable of folding many endogenous proteins and even some heterologous proteins, further engineering of the chaperonin has been reported to improve its ability to fold heterologous proteins that do not naturally encounter the complex. Wang et al. used rounds of selection and DNA shuffling to evolve mutants of GroEL/ES that more efficiently folded a single substrate, namely GFP (Wang et al., 2002). Interestingly, the results highlighted an inherent conflict between specificity and generality of cellular chaperonins. That is, GroEL/ES mutants that were more efficient at folding GFP lost their ability to fold many endogenous substrates. Hence, stringent substrate specificity appears to come at the expense of the ability of the chaperonin to fold its normal repertoire of proteins within the cell. Even so, engineering of GroEL/ES and other chaperonins could provide a general strategy for improving heterologous protein expression. Conversely, computational models have been developed to engineer substrate proteins to become better substrates for GroEL/ES (Kumar et al., 2012).

Repurposing Translocation Machinery for Improved Protein Folding

In addition to chaperonins, bacterial protein translocators have also been used to improve intracellular folding and stability of recombinant proteins. Notably, a protein translocation system in *E. coli* known as the twin-arginine translocation (Tat) pathway has been used to develop genetic screens and selections for efficient folding of recombinant proteins. In nature, the Tat system is one of two distinct pathways for the export of proteins across the inner membrane of *E. coli*. The bulk of proteins in *E. coli* are exported by the general secretory (Sec) pathway, which transports primarily unfolded polypeptides, often as they are still being translated. The Tat system, which involves a translocase comprised of the Tat(A/E)BC proteins, transports proteins that have been completely synthesized and folded in the cytoplasm prior to export. Inherent in this mechanism is a quality-control feature that allows Tat to discriminate between substrates based on the condition of their fold and ensures that only properly folded proteins are exported (DeLisa et al., 2003). Using a directed-evolution approach, Rocco et al. evolved Tat translocase variants carrying mutations that suppressed folding quality control and permitted export of misfolded substrate proteins (Rocco et al., 2012). These studies provided evidence for the direct participation of the Tat translocase in structural proofreading of substrate proteins and reveals epitopes in the translocase that are important for this process.

Tat-based selections for protein folding take advantage of the built-in quality-control feature of Tat to link protein folding with drug resistance (Fisher et al., 2006). This involved creating a tripartite fusion in which the protein of interest was sandwiched between an N-terminal Tat signal peptide and a C-terminal TEM1 β-lactamase (Bla) reporter protein. Only proteins that were well folded and efficiently exported by Tat conferred resistance to ampicillin (Amp), whereas proteins that misfolded in the cytoplasm were excluded from export and thus did not confer Amp resistance. This system and its derivatives (Karlsson et al., 2012; Lee et al., 2009) have been used to generate solubility-enhanced variants of several heterologous proteins, including Alzheimer's Aβ42 peptide (Fisher et al., 2006), intracellular antibodies in the single-chain Fv (scFv) format (Fisher and DeLisa, 2009) and heavy-chain variable domain (V_H) format (Kim et al., 2014), and the endoglucanase Cel5A from the plant-pathogenic fungus *Fusarium graminearum* (Boock et al., 2015).

Applying Synthetic Protein Machinery in Biotechnology

A motivating factor for the engineering of naturally occurring protein machinery is the opportunity for developing new biotechnological applications. Many of the engineered protein machines we discussed here may have applications in higher-level biosynthetic processes, such as disulfide bond formation or asparagine-linked (*N*-linked) protein glycosylation. These processes are among many in biology that require the spatially and temporally coordinated action of multiple enzymes and protein assemblies to produce their complex products.

Please cite this article as: Glasscock et al., Engineered Protein Machines: Emergent Tools for Synthetic Biology, Cell Chemical Biology (2016), http://dx.doi.org/10.1016/j.chembiol.2015.12.004

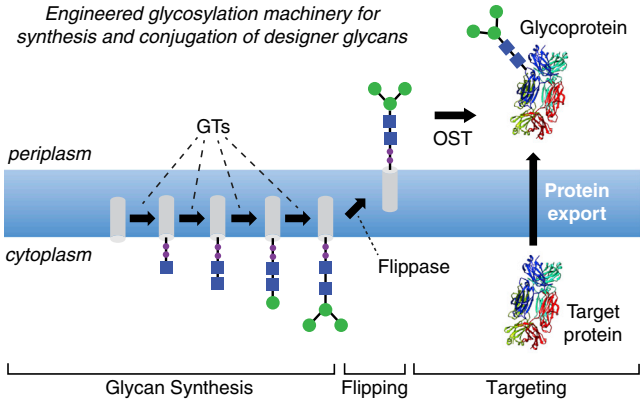

Figure 5. Engineering Bacteria with Protein Glycosylation Machinery

Schematic of a bottom-up engineered pathway for the biosynthesis and conjugation of designer glycans in living *E. coli* cells (Baker et al., 2013; Merritt et al., 2013; Valderrama-Rincon et al., 2012). Glycosyltransferase (GT) enzymes are directed to catalyze the synthesis of desired glycan structures in the form of lipid-linked oligosaccharides (LLOs) on the cytoplasmic side of the inner membrane. Native or heterologous flippase enzymes then translocate the LLO across in the inner membrane into the periplasm, where an oligosaccharyltransferase (OST) enzyme conjugates the glycan onto a target protein (Davies et al., 2014).

In the case of *N*-linked glycosylation, engineering synthetic pathways of glycosyltransferases (GTs) and oligosaccharyltransferases (OSTs) in *E. coli* has enabled the production of numerous medically relevant glycoconjugates (Baker et al., 2013; Cuccui and Wren, 2015; Merritt et al., 2013; Terra et al., 2012). These engineered pathways involve complex, multi-step processes that involve the expression of heterologous enzymes from multiple organisms and can be divided into distinct stages of glycan synthesis, glycan flipping, and glycan targeting (Figure 5). In the first stage, glycans are assembled in the form of lipid-linked oligosaccharides (LLOs) on the cytoplasmic face of the inner membrane. This involves functional co-expression of heterologous GT enzymes that catalyze the sequential addition of nucleotide-activated sugars. In the subsequent glycan-flipping stage, native *E. coli* or heterologous LLO flippase enzymes are utilized to flip newly synthesized LLOs to the periplasmic face of the inner membrane. Finally, in the glycan-targeting stage, an OST enzyme transfers the glycan from its lipid anchor to an asparagine residue in a periplasmically targeted substrate protein.

While bacterial glycoengineering efforts have clear implications for the development of new therapeutics and vaccines, there are many obstacles to efficient large-scale production of glycans and glycoproteins. For example, many required heterologous enzymes, such as sialyltransferases, are poorly expressed in wild-type *E. coli*. Engineering of protein chaperonin machines or genetic selection using the Tat translocation machinery (Fisher et al., 2006; Wang et al., 2002) could be used to optimize the folding and expression of these recalcitrant enzymes. In addition, overexpression of multiple heterologous enzymes often imposes a large burden on the host cell. Systems such as the split T7 RNAP-based resource allocator (Segall-Shapiro et al., 2014) might permit balancing of pathway enzyme production while minimizing the burden on native transcription machinery. In some cases, enzymes need to act sequentially to synthesize and target the correct glycan structure. Implementing mechanisms for dynamic regulation, such as genetic toggle-switches (Gardner et al., 2000), oscillators (Stricker et al., 2008), and single-input modules (Takahashi et al., 2015), could be an effective strategy to coordinate the expression or activity of individual pathway enzymes in time. Constructing these dynamic regulatory devices would benefit from the use of well-characterized parts, including orthogonal RNAPs and ribosomes, as well as other gene-expression elements from synthetic biology, such as engineered sigma factors (Stanton et al., 2014) and RNA regulators of transcription and translation (Chappell et al., 2015).

Furthermore, highly efficient assembly of engineered glycans and their site-specific conjugation to target proteins of interest, especially those of non-bacterial origin, may require reprogramming of GT and OST function (Aharoni et al., 2006; Ollis et al., 2014). Likewise, engineered glycosylation machinery may be needed to modify microbially produced glycans with bio-orthogonal chemical reporters, such as azido sugars (e.g., GlcNAz, ManNAz, SiaNAz), as a means of introducing non-natural chemical functionality into glycoproteins (Mahal et al., 1997). One possible way to address these and other glycoengineering challenges is to develop an in vitro compartmentalization strategy, akin to the ones described above for DNA and RNA synthesis machinery, for the machinery involved in glycosylation. Such a strategy might be made possible by the recently described ability to perform cell-free protein glycosylation reactions using lysates from glycoengineered *E. coli* (Guarino and DeLisa, 2012). Finally, alternative strategies for covalently attaching glycans onto proteins might become useful when it is difficult to flip and/or target certain glycan structures in vivo. In this vein, orthogonal ribosomes that efficiently incorporate alkyne or alkene UAAs could be a route for attaching glycan structures onto proteins via click chemistry or other conjugation techniques (Kaya et al., 2009). Further modeling of glycan structure would then be enabled by the array of chemo-enzymatic glycan synthesis technologies that are now available (Wang and Amin, 2014; Wang and Davis, 2013; Wang et al., 2013).

Conclusion

Synthetic biologists have made great strides in engineering the protein machines responsible for gene expression and protein synthesis. In the future, one might expect that protein machines at all these levels will be implemented in the design of new biotechnological processes. For example, we expect that many biosynthetic pathways will be enhanced by the use of tools for orthogonal control over gene expression and protein synthesis. In this realm, protein machines such as engineered T7 RNAP variants and orthogonal ribosomes can be used to create cellular logic gates that control decision making in cells, possibly enabling the timed expression of metabolic pathway enzymes in response to intra- and extra-cellular cues. Metabolic pathways often require balancing of enzyme expression levels to achieve proper stoichiometry (Pfleger et al., 2006). In this case, there may be clusters of enzymes that need to be balanced in a concerted fashion but have dynamically changing expression levels over time. The use of T7 RNAP-based resource allocators could enable this type of regulation by expressing enzyme clusters under the control of different

T7 RNAP σ fragments. Finally, synthetic protein machines such as orthogonal ribosomes and chaperonins are among many other engineered protein machines, such as orthogonal tRNA synthetases (Chin et al., 2003; Furter, 1998; Wang et al., 2001), that will be useful to efficiently express well-folded proteins that contain multiple UAAs or are encoded by sequences that are not amenable to the native host machinery. The successful implementation of many higher-level biosynthetic processes in bacteria, such as disulfide-bond formation, *N*-linked glycosylation, and synthesis of natural products, will likely rely on solutions to all of the above challenges. Future engineering of protein machinery will make these tasks possible and pave the way for many new possibilities in biotechnology.

AUTHOR CONTRIBUTIONS

C.J.G, J.B.L, and M.P.D. conceptualized and wrote the manuscript.

ACKNOWLEDGMENTS

We would like to thank the DeLisa and Lucks research groups at Cornell University for helpful comments and advice in preparation of this manuscript. This work was supported by the National Science Foundation grant CBET-1402843 (to M.P.D. and J.B.L.) and a National Science Foundation Graduate Research Fellowship (to C.J.G.). M.P.D. is a founder of Glycobia, Inc. and is a member of its scientific advisory board.

REFERENCES

Aharoni, A., Thieme, K., Chiu, C.P.C., Buchini, S., Lairson, L.L., Chen, H., Strynadka, N.C.J., Wakarchuk, W.W., and Withers, S.G. (2006). High-throughput screening methodology for the directed evolution of glycosyltransferases. Nat. Methods *3*, 609–614.

Alberts, B. (1998). The cell as a collection of protein machines: preparing the next generation of molecular biologists. Cell *92*, 291–294.

An, W., and Chin, J.W. (2009). Synthesis of orthogonal transcription-translation networks. Proc. Natl. Acad. Sci. USA *106*, 8477–8482.

Arezi, B., McKinney, N., Hansen, C., Cayouette, M., Fox, J., Chen, K., Lapira, J., Hamilton, S., and Hogrefe, H. (2014). Compartmentalized self-replication under fast PCR cycling conditions yields Taq DNA polymerase mutants with increased DNA-binding affinity and blood resistance. Front. Microbiol. *5*, 408.

Baker, J.L., Çelik, E., and DeLisa, M.P. (2013). Expanding the glycoengineering toolbox: the rise of bacterial N-linked protein glycosylation. Trends. Biotechnol. *31*, 313–323.

Bayer, T.S., Widmaier, D.M., Temme, K., Mirsky, E.A., Santi, D.V., and Voigt, C.A. (2009). Synthesis of methyl halides from biomass using engineered microbes. J. Am. Chem. Soc. *131*, 6508–6515.

Bereza-Malcolm, L.T., Mann, G., and Franks, A.E. (2015). Environmental sensing of heavy metals through whole cell microbial biosensors: a synthetic biology approach. ACS Synth. Biol. *4*, 535–546.

Boock, J.T., King, B.C., Taw, M.N., Conrado, R.J., Siu, K.-H., Stark, J.C., Walker, L.P., Gibson, D.M., and DeLisa, M.P. (2015). Repurposing a bacterial quality control mechanism to enhance enzyme production in living cells. J. Mol. Biol. *427*, 1451–1463.

Carlson, J.C., Badran, A.H., Guggiana-Nilo, D.A., and Liu, D.R. (2014). Negative selection and stringency modulation in phage-assisted continuous evolution. Nat. Chem. Biol. *10*, 216–222.

Chappell, J., Watters, K.E., Takahashi, M.K., and Lucks, J.B. (2015). A renaissance in RNA synthetic biology: new mechanisms, applications and tools for the future. Curr. Opin. Chem. Biol. *28*, 47–56.

Chelliserrykattil, J., Cai, G., and Ellington, A.D. (2001). A combined in vitro/in vivo selection for polymerases with novel promoter specificities. BMC Biotechnol. *1*, 13.

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.

Collins, C.H., Yokobayashi, Y., Umeno, D., and Arnold, F.H. (2003). Engineering proteins that bind, move, make and break DNA. Curr. Opin. Biotechnol. *14*, 665, 378.

Cuccui, J., and Wren, B. (2015). Hijacking bacterial glycosylation for the production of glycoconjugates, from vaccines to humanised glycoproteins. J. Pharm. Pharmacol. *67*, 338–350.

Davies, A.M., Jefferis, R., and Sutton, B.J. (2014). Crystal structure of deglycosylated human IgG4-Fc. Mol. Immunol. *62*, 46–53.

DeLisa, M.P., Tullman, D., and Georgiou, G. (2003). Folding quality control in the export of proteins by the bacterial twin-arginine translocation pathway. Proc. Natl. Acad. Sci. USA *100*, 6115–6120.

Dunkle, J.A., Wang, L., Feldman, M.B., Pulk, A., Chen, V.B., Kapral, G.J., Noeske, J., Richardson, J.S., Blanchard, S.C., and Cate, J.H.D. (2011). Structures of the bacterial ribosome in classical and hybrid states of tRNA binding. Science *332*, 981–984.

Ellefson, J.W., Meyer, A.J., Hughes, R.A., Cannon, J.R., Brodbelt, J.S., and Ellington, A.D. (2013). Directed evolution of genetic parts and circuits by compartmentalized partnered replication. Nat. Biotechnol. *32*, 97–101.

Esvelt, K.M., Carlson, J.C., and Liu, D.R. (2011). A system for the continuous directed evolution of biomolecules. Nature *472*, 499–503.

Fisher, A.C., and DeLisa, M.P. (2009). Efficient isolation of soluble intracellular single-chain antibodies using the twin-arginine translocation machinery. J. Mol. Biol. *385*, 299–311.

Fisher, A.C., Kim, W., and DeLisa, M.P. (2006). Genetic selection for protein solubility enabled by the folding quality control feature of the twin-arginine translocation pathway. Protein Sci. *15*, 449–458.

Fischer, N., Neumann, P., Konevega, A.L., Bock, L.V., Ficner, R., Rodnina, M.V., and Stark, H. (2015). Structure of the E. coli ribosome-EF-Tu complex at <3 Å resolution by Cs-corrected cryo-EM. Nature *520*, 567–570.

Fried, S.D., Schmied, W.H., Uttamapinant, C., and Chin, J.W. (2015). Ribosome subunit stapling for orthogonal translation in E. coli. Angew. Chem. Int. Ed. Engl. *54*, 12791–12794.

Furter, R. (1998). Expansion of the genetic code: site-directed p-fluoro-phenylalanine incorporation in Escherichia coli. Protein Sci. *7*, 419–426.

Gardner, T.S., Cantor, C.R., and Collins, J.J. (2000). Construction of a genetic toggle switch in Escherichia coli. Nature *403*, 339–342.

Ghadessy, F.J., Ong, J.L., and Holliger, P. (2001). Directed evolution of polymerase function by compartmentalized self-replication. Proc. Natl. Acad. Sci. USA *98*, 4552–4557.

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.

Glick, E., Anderson, J.P., and Loeb, L.A. (2002). In vitro production and screening of DNA polymerase eta mutants for catalytic diversity. Biotechniques *33*, 1136–1142, 1144.

Guarino, C., and DeLisa, M.P. (2012). A prokaryote-based cell-free translation system that efficiently synthesizes glycoproteins. Glycobiology *22*, 596–601.

Hui, A., and de Boer, H.A. (1987). Specialized ribosome system: preferential translation of a single mRNA species by a subpopulation of mutated ribosomes in Escherichia coli. Proc. Natl. Acad. Sci. USA *84*, 4762–4766.

Karlsson, A.J., Lim, H.K., Xu, H., and Rocco, M.A. (2012). Engineering antibody fitness and function using membrane-anchored display of correctly folded proteins. J. Mol. Biol. *416*, 94–107.

Kaya, E., Gutsmiedl, K., Vrabel, M., Müller, M., Thumbs, P., and Carell, T. (2009). Synthesis of threefold glycosylated proteins using click chemistry and genetically encoded unnatural amino acids. Chembiochem *10*, 2858–2861.

Keasling, J.D. (2010). Manufacturing molecules through metabolic engineering. Science *330*, 1355–1358.

Kerner, M.J., Naylor, D.J., Ishihama, Y., Maier, T., Chang, H.-C., Stines, A.P., Georgopoulos, C., Frishman, D., Hayer-Hartl, M., Mann, M., et al. (2005). Proteome-wide analysis of chaperonin-dependent protein folding in Escherichia coli. Cell *122*, 209–220.

Kim, D.-S., Song, H.-N., Nam, H.J., Kim, S.-G., Park, Y.-S., Park, J.-C., Woo, E.-J., and Lim, H.-K. (2014). Directed evolution of human heavy chain variable domain (VH) using in vivo protein fitness filter. PLoS One *9*, e98178.

Kumar, V., Punetha, A., Sundar, D., and Chaudhuri, T.K. (2012). In silico engineering of aggregation-prone recombinant proteins for substrate recognition by the chaperonin GroEL. BMC Genomics *13* (*Suppl 7*), S22.

Lang, K., and Chin, J.W. (2014). Cellular incorporation of unnatural amino acids and bioorthogonal labeling of proteins. Chem. Rev. *114*, 4764–4806.

Lee, L.L., Ha, H., Chang, Y.T., and DeLisa, M.P. (2009). Discovery of amyloid-beta aggregation inhibitors using an engineered assay for intracellular protein folding and solubility. Protein Sci. *18*, 277–286.

Li, G.-W., Oh, E., and Weissman, J.S. (2012). The anti-Shine–Dalgarno sequence drives translational pausing and codon choice in bacteria. Nature *484*, 538–541.

Liljas, A., and Ehrenberg, M. (2011). Structural Aspects of Protein Synthesis, Second Edition (World Scientific).

Lucks, J.B., Qi, L., Whitaker, W.R., and Arkin, A.P. (2008). Toward scalable parts families for predictable design of biological circuits. Curr. Opin. Microbiol. *11*, 567–573.

Lutz, J.-F., Ouchi, M., Liu, D.R., and Sawamoto, M. (2013). Sequence-controlled polymers. Science *341*, 1238149.

Mahal, L.K., Yarema, K.J., and Bertozzi, C.R. (1997). Engineering chemical reactivity on cell surfaces through oligosaccharide biosynthesis. Science *276*, 1125–1128.

Merritt, J.H., Ollis, A.A., Fisher, A.C., and DeLisa, M.P. (2013). Glycans-by-design: engineering bacteria for the biosynthesis of complex glycans and glycoconjugates. Biotechnol. Bioeng. *110*, 1550–1564.

Meyer, A.J., Ellefson, J.W., and Ellington, A.D. (2014). Directed evolution of a panel of orthogonal T7 RNA polymerase variants for in vivo or in vitro synthetic circuitry. ACS Synth. Biol. *4*, 1070–1076.

Netzer, W.J., and Hartl, F.U. (1997). Recombination of protein domains facilitated by co-translational folding in eukaryotes. Nature *388*, 343–349.

Neumann, H., Wang, K., Davis, L., Garcia-Alai, M., and Chin, J.W. (2010). Encoding multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome. Nature *464*, 441–444.

Ollis, A.A., Zhang, S., Fisher, A.C., and DeLisa, M.P. (2014). Engineered oligosaccharyltransferases with greatly relaxed acceptor-site specificity. Nat. Chem. Biol. *10*, 816–822.

Orelle, C., Carlson, E.D., Szal, T., Florin, T., Jewett, M.C., and Mankin, A.S. (2015). Protein synthesis by ribosomes with tethered subunits. Nature *524*, 119–124.

Pfleger, B.F., Pitera, D.J., Smolke, C.D., and Keasling, J.D. (2006). Combinatorial engineering of intergenic regions in operons tunes expression of multiple genes. Nat. Biotechnol. *24*, 1027–1032.

Pinheiro, V.B., and Holliger, P. (2012). The XNA world: progress towards replication and evolution of synthetic genetic polymers. Curr. Opin. Chem. Biol. *16*, 245–252.

Pinheiro, V.B., Taylor, A.I., Cozens, C., Abramov, M., Renders, M., Zhang, S., Chaput, J.C., Wengel, J., Peak-Chew, S.-Y., McLaughlin, S.H., et al. (2012). Synthetic genetic polymers capable of heredity and evolution. Science *336*, 341–344.

Rackham, O., and Chin, J.W. (2005a). A network of orthogonal ribosome x mRNA pairs. Nat. Chem. Biol. *1*, 159–166.

Rackham, O., and Chin, J.W. (2005b). Cellular logic with orthogonal ribosomes. J. Am. Chem. Soc. *127*, 17584–17585.

Rackham, O., Wang, K., and Chin, J.W. (2006). Functional epitopes at the ribosome subunit interface. Nat. Chem. Biol. *2*, 254–258.

Raskin, C.A., Diaz, G., Joho, K., and McAllister, W.T. (1992). Substitution of a single bacteriophage T3 residue in bacteriophage T7 RNA polymerase at position 748 results in a switch in promoter specificity. J. Mol. Biol. *228*, 506–515.

Raskin, C.A., Diaz, G.A., and McAllister, W.T. (1993). T7 RNA-polymerase mutants with altered promoter specificities. Proc. Natl. Acad. Sci. USA *90*, 3147–3151.

Rocco, M.A., Waraho-Zhmayev, D., and DeLisa, M.P. (2012). Twin-arginine translocase mutations that suppress folding quality control and permit export of misfolded substrate proteins. Proc. Natl. Acad. Sci. USA *109*, 13392–13397.

Ryu, Y., and Schultz, P.G. (2006). Efficient incorporation of unnatural amino acids into proteins in Escherichia coli. Nat. Methods *3*, 263–265.

Sachdeva, A., Wang, K., Elliott, T., and Chin, J.W. (2014). Concerted, Rapid, Quantitative, and site-specific dual labeling of proteins. J. Am. Chem. Soc. *136*, 7785–7788.

Saibil, H. (2013). Chaperone machines for protein folding, unfolding and disaggregation. Nat. Rev. Mol. Cell. Biol. *14*, 630–642.

Schaerli, Y., Gili, M., and Isalan, M. (2014). A split intein T7 RNA polymerase for transcriptional AND-logic. Nucleic Acids Res. *42*, 12322–12328.

Segall-Shapiro, T.H., Meyer, A.J., Ellington, A.D., Sontag, E.D., and Voigt, C.A. (2014). A "resource allocator" for transcription based on a highly fragmented T7 RNA polymerase. Mol. Syst. Biol. *10*, 742.

Shis, D.L., and Bennett, M.R. (2013). Library of synthetic transcriptional AND gates built with split T7 RNA polymerase mutants. Proc. Natl. Acad. Sci. USA *110*, 5028–5033.

Singh, S.K., and Koshkin, A.A. (1998). LNA (locked nucleic acids): synthesis and high-affinity nucleic acid recognition. Chem. Commun. *4*, 455–456.

Stanton, B.C., Nielsen, A.A.K., Tamsir, A., Clancy, K., Peterson, T., and Voigt, C.A. (2014). Genomic mining of prokaryotic repressors for orthogonal logic gates. Nat. Chem. Biol. *10*, 99–105.

Steitz, J.A., and Jakes, K. (1975). How ribosomes select initiator regions in mRNA: base pair formation between the 3' terminus of 16S rRNA and the mRNA during initiation of protein synthesis in Escherichia coli. Proc. Natl. Acad. Sci. USA *72*, 4734–4738.

Stricker, J., Cookson, S., Bennett, M.R., Mather, W.H., Tsimring, L.S., and Hasty, J. (2008). A fast, robust and tunable synthetic gene oscillator. Nature *456*, 516–519.

Takahashi, M.K., Chappell, J., Hayes, C.A., Sun, Z.Z., Kim, J., Singhal, V., Spring, K.J., Al-Khabouri, S., Fall, C.P., Noireaux, V., et al. (2015). Rapidly characterizing the fast dynamics of RNA genetic circuitry with cell-free transcription-translation (TX-TL) systems. ACS Synth. Biol. *4*, 503–515.

Temme, K., Hill, R., Segall-Shapiro, T.H., Moser, F., and Voigt, C.A. (2012). Modular control of multiple pathways using engineered orthogonal T7 polymerases. Nucleic Acids Res. *40*, 8773–8781.

Terra, V.S., Mills, D.C., Yates, L.E., Abouelhadid, S., Cuccui, J., and Wren, B.W. (2012). Recent developments in bacterial protein glycan coupling technology and glycoconjugate vaccine design. J. Med. Microbiol. *61*, 919–926.

Tindall, K.R., and Kunkel, T.A. (1988). Fidelity of DNA synthesis by the Thermus aquaticus DNA polymerase. Biochemistry *27*, 6008–6601.

Tubeleviciute, A., and Skirgaila, R. (2010). Compartmentalized self-replication (CSR) selection of Thermococcus litoralis Sh1B DNA polymerase for diminished uracil binding. Protein Eng. Des. Sel. *23*, 589–597.

Valderrama-Rincon, J.D., Fisher, A.C., Merritt, J.H., Fan, Y.-Y., Reading, C.A., Chhiba, K., Heiss, C., Azadi, P., Aebi, M., and DeLisa, M.P. (2012). An engineered eukaryotic protein glycosylation pathway in Escherichia coli. Nat. Chem. Biol. *8*, 434–436.

Wang, L.-X., and Davis, B.G. (2013). Realizing the promise of chemical glycobiology. Chem. Sci. *4*, 3381–3394.

Wang, L.-X., and Amin, M.N. (2014). Chemical and chemoenzymatic synthesis of glycoproteins for deciphering functions. Chem. Biol. *21*, 51–66.

Wang, L., Brock, A., Herberich, B., and Schultz, P.G. (2001). Expanding the genetic code of Escherichia coli. Science *292*, 498–500.

Wang, J.D., Herman, C., Tipton, K.A., Gross, C.A., and Weissman, J.S. (2002). Directed evolution of substrate-optimized GroEL/S chaperonins. Cell *111*, 1027–1039.

Wang, Y., Prosen, D.E., Mei, L., Sullivan, J.C., Finney, M., and Vander Horn, P.B. (2004). A novel strategy to engineer DNA polymerases for enhanced processivity and improved performance in vitro. Nucleic Acids Res. *32*, 1197–1207.

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.

Wang, Z., Chinoy, Z.S., Ambre, S.G., Peng, W., McBride, R., de Vries, R.P., Glushka, J., Paulson, J.C., and Boons, G.-J. (2013). A general strategy for the chemoenzymatic synthesis of asymmetrically branched N-glycans. Science *341*, 379–383.

Wang, K., Sachdeva, A., Cox, D.J., Wilf, N.M., Wilf, N.W., Lang, K., Wallace, S., Mehl, R.A., and Chin, J.W. (2014). Optimized orthogonal translation of unnatural amino acids enables spontaneous protein double-labelling and FRET. Nat. Chem. *6*, 393–403.

Ying, B.-W., Taguchi, H., Kondo, M., and Ueda, T. (2005). Co-translational involvement of the chaperonin GroEL in the folding of newly translated polypeptides. J. Biol. Chem. *280*, 12035–12040.

